TEXAS STATE UNIVERSITY

The rising STAR *of Texas*

TxDOT Report 0-7144-1
Develop a Real-Time Decision Support Tool for Urban Roadway Safety Improvements

Texas State University
Subasish Das (PI)
Jinli Liu
David Mills
Mahmuda Sultana Mimi

Texas A&M Transportation Institute
Lingtao Wu
Srinivas Geedipally
Richard Dzinyela

Submitted August 2025; Published January 2026

Report No: 0-7144-R1

TEXAS STATE UNIVERSITY

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

Final Report 0-7144-R1

Project performed in cooperation with
the Texas Department of Transportation

Texas Department of Transportation

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA/TX-26/0-7144-R1		

4. Title and Subtitle	5. Report Date
Develop a Real-Time Decision Support Tool for Urban Roadway Safety Improvements	January 2026
	6. Performing Organization Code

7. Author(s)	8. Performing Organization Report No.
Subasish Das, Jinli Liu, David Mills, Mahmuda Sultana Mimi, Lingtao Wu, Srinivas Geedipally, and Richard Dzinyela	Report 0-7144-R1

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Texas State University, San Marcos, Texas 78666 Texas A&M Transportation Institute (TTI) College Station, Texas 77843-3135	11. Contract or Grant No. Project 0-7144

12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
Texas Department of Transportation Research and Technology Implementation Office 125 E. 11th Street, Austin, Texas 78701-2483	Technical Report: January 2023–August 2025
	14. Sponsoring Agency Code

15. Supplementary Notes

Project performed in cooperation with the Texas Department of Transportation
Project Title: Develop a Real-Time Decision Support Tool for Urban Roadway Safety Improvements
URL:

16. Abstract

Speed-related crashes remain a critical concern for Texas, identified as one of the seven key emphasis areas in the Texas Strategic Highway Safety Plan (SHSP) for 2018–2022. Traditional crash risk models often overlook essential real-time factors such as operating speed, traffic variability, and weather, limiting their effectiveness. To address this, TxDOT Project 0-7144 integrated data from diverse sources, including NPMRDS and INRIX XD (real-time speeds), TMAS (traffic volumes), CRIS (crashes), RHiNO (roadway inventory), and NOAA (weather), to develop both annual and short-duration Safety Performance Functions (SPFs) for a wide range of urban facility types. Annual SPFs were developed for urban facilities, including freeways, multi-lane divided and undivided highways, continuous left-turn lane roads, and two-lane highways, incorporating geometric and operational features such as lane and shoulder width, truck proportion, driveway density, average speed, and precipitation. This project also employed a Negative Binomial-Lindley (NBL) modeling framework to produce monthly SPFs from 2019 to 2022. These short-duration models captured the dynamic influence of seasonal factors, monthly average daily traffic (MADT), speed and travel time variability, and weather patterns on crash frequency across facility types. A key outcome of the project is the development of an interactive, GIS-based decision support tool using the Shiny platform. This tool enables users to visualize, filter, and download crash prediction outputs at both annual and short-duration levels. It offers dynamic data uploading capabilities, allowing local engineers to input updated roadway and traffic data for customized analyses. The project concludes with recommendations to update Texas SPFs for urban roadways by integrating operating speed and weather data, emphasizing speed consistency in guidelines, and expanding the tool's capacity for daily-level modeling and facility-specific diagnostics.

17. Key Words	18. Distribution Statement
Urban facilities, traffic safety, safety performance functions, speed CMF, short duration model, 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		

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

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

by

Subasish Das, Ph.D.
Research Supervisor, Assistant Professor
Texas State University

Jinli Liu
Post-doctoral Associate
Texas State University

David Mills
Ph.D. Student
Texas State University

Mahmuda Sultana Mimi
M.S. Student
Texas State University

Lingtao Wu, Ph.D.
Associate Research Engineer
Texas A&M Transportation Institute

Srinivas Geedipally, Ph.D.
Research Engineer
Texas A&M Transportation Institute

Richard Dzinyela
Ph.D. Student
Texas A&M Transportation Institute

Report 0-7144-R1
Project 0-7144
Project Title: Develop a Real-Time Decision Support Tool for Urban Roadway Safety Improvements

Performed in cooperation with the
Texas Department of Transportation

January 2026
TEXAS STATE UNIVERSITY
San Marcos, Texas 78666

DISCLAIMER

This research was performed in cooperation with the Texas Department of Transportation (TxDOT). 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.

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. The researchers would like to thank TxDOT panel members (Alberto Guevera, Hui Wu, Epigmenio Gonzalez, Jenny Li, Jose Madrid Jr., Mark Johnson, Tamara Gart, and Temi Abayomi) for providing valuable help, information, data, and advice throughout this project. The researchers also gratefully acknowledge the support and assistance provided by TxDOT project manager Katelyn Kasberg.

Project team members met with numerous other individuals at TxDOT to gather and/or complement data and information needed for the analysis. The researchers would also like to thank Darrell Anderson, Syed Aaqib Javed, Rohit Chakraborty, Swastika Barua for gathering and processing various datasets that were analyzed in this study.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

Page

LIST OF ACRONYMS, ABBREVIATIONS, AND TERMS

AADT	Annual Average Daily Traffic
2U	Two-Lane Undivided Arterial
3T	Three-Lane Arterials
4D	Four-Lane Divided Arterials
4U	Four-Lane Undivided Arterials
5T	Five-Lane Arterials Including a Center TWLTL
AASHTO	American Association of State Highway Transportation Officials
ADAS	Advanced Driver Assistance Systems
ADT	Average Daily Traffic
AEB	Automatic Emergency Braking
ASC	Automatic Speed Control
C3S	Copernicus Climate Change Service
CAMS	Copernicus Atmospheric Monitoring Service
CART	Classification And Regression Trees
CDS	Copernicus Climate Data Store
CMF	Crash Modification Factor
CNN	Convolutional Neural Network
CRIS	Crash Records Information System
CRS	Coordinate Reference System
DOT	Department of Transportation
DT	Decision Tree
DUSL	Default Urban Speed Limit
EB	Empirical Bayes
ECMWF	European Centre for Medium-Range Weather Forecasts
EDC - 5	Everyday Counts
EV	Electric Vehicles
FCS	Functional Categorization System
FHWA	Federal Highway Administration
FI	Fatal and Injury
FICrE	Fatal/Injury Crashes
FMCSA	Federal Motor Carrier Safety Administration
GIS	Geographic Information System
GRIB	Gridded Binary
HSM	Highway Safety Manual
IRI	International Roughness Index
ITS	Intelligent Transportation System

KDT	Knowledge Discovery In Text
KE	Kinetic Energy
KNN	K-Nearest-Neighbor
MCSSD	Cross-Section Speed Standard Deviation
MNL	Multinomial Logit
MVPLN	Multivariate Poisson-Lognormal Model
NB	Negative Binomial
NCHRP	National Cooperative Highway Research Program
NHS	National Highway System
NHTSA	National Highway Traffic Safety Administration
NOAA	National Oceanic and Atmospheric Administration
NPMRDS	National Performance Management Research Data Set
PDO	Property Damage Only
PDOCrE	Property Damage Only Crashes
PHBs	Pedestrian Hybrid Beacons
PLN	Poisson-Lognormal
PSL	Posted Speed Limit
RF	Random Forest
RHiNO	Road-Highway Inventory Network Offload
RRFBs	Rectangular Rapid Flashing Beacons
SDCSM	Standard Deviation of the Cross-Sectional Speed Mean
SEM	Structural Equation Modeling
SHSP	Strategic Highway Safety Plan
SPF	Safety Performance Function
SSA	Safe System Approach
SS	Safe System
STM	Structural Topic Modeling
SVM	Support Vector Machines
TASP-CNN	Traffic Accident's Severity Prediction-Convolutional Neural Network
TMAS	Travel Monitoring Analysis System
TotalCrE	Total Crashes
TWLTL	Two-Way Left-Turn Lane
TxDOT	Texas Department of Transportation
VSL	Variable Speed Limit
WHO	World Health Organization

CHAPTER 1:
INTRODUCTION

1.1 BACKGROUND

The identification of appropriate correlations between operating speed, roadway geometry, and traffic exposure can significantly contribute to the enhancement of roadway safety. A comprehensive understanding and characterization of these associations has the potential to advance existing safety improvement procedures leading to a reduction in both the frequency and severity of crashes. Therefore, it is imperative to employ data-driven methodologies to gain a better comprehension of these relationships. Traditional assessments primarily rely on corridor traffic volume and physical site characteristics. However, the lack of reliable data on operating speeds has been a major obstacle in developing robust models that depict this relationship, particularly for urban roadways.

Despite the significance of including speed in safety evaluations, it is noteworthy that the first edition of the Highway Safety Manual (HSM) (as well as the upcoming second edition) does not directly employ speed as a measure in the safety performance functions (SPFs) for various types of road facilities (AASHTO, 2010). This omission mainly arises from the intricate interplay between roadway geometry, speed, and crash frequency. Furthermore, research findings have not provided conclusive evidence in determining appropriate speed measures. While some studies indicate that higher operating speeds result in more frequent and severe crashes, other studies have observed the opposite effect. Overall, it is evident that the existing models may not adequately account for the interconnectedness between speed measures, roadway factors, and safety.

Traditional crash risk analysis methods often overlook crucial factors such as real-time speed, real-time volume, and weather conditions, resulting in limited effectiveness in predicting safety risks. To address this research gap, it is proposed to enhance existing state-specific traffic crash data by incorporating three national databases:

- **The NPMRDS:** This database contains comprehensive travel time data for both passenger and freight vehicles, encompassing the National Highway System (NHS) as well as other roadways. By integrating NPMRDS data, researchers can gain valuable insights into the performance of transportation networks. INRIX XD data provides additional details on non-NHS roadways.
- **The Travel Monitoring Analysis System (TMAS):** TMAS provides traffic volume data through temporary traffic counting and continuous traffic counting programs. By incorporating TMAS data, researchers can access information regarding the flow of vehicles, enabling a more comprehensive analysis of crash risks.

- **Real-time weather data from the National Oceanic and Atmospheric Administration (NOAA):** Weather conditions significantly influence road safety, and the inclusion of real-time weather data from NOAA enhances the accuracy of crash risk analysis. Factors such as precipitation, temperature, and visibility can now be considered, enabling a more holistic assessment.

The 0-7144 research project aims to address these limitations by developing updated SPFs specifically tailored for urban roadways. Additionally, a decision support tool was developed to facilitate the exploration of segment-based crash risks for various types of urban roadway facilities. By leveraging these integrated datasets and tools, transportation professionals will have access to improved methodologies for assessing crash risks, allowing for more effective safety planning and decision-making processes.

1.2 PROJECT GOAL AND RESEARCH TASKS

The objective of this study is to develop annual and short duration SPF development for urban facilities. This project aims to answer the following three research questions:

- To what extent can operational speed and weather variables be interpreted as directly affecting crash risk or crash severity on urban roadways?
- To what extent can the decision criteria be developed in assigning the risk measures of the urban roadways for the decision support tool?
- Can the decision support tool provide updated risk scores based on the new influx of data?

To achieve the project goals, the Project Team has conducted five major tasks, and they are summarized as follows:

- **Task 2 – Literature Review:** The Project Team conducted a broader overview of speed and safety on urban roadway networks.
- **Task 3 – Data Collection and Data Preparation:** The Project Team conflated several data sources to develop the database for the safety evaluation of urban roadways. For traffic crash data, the Project Team used five years (2018-2022) of traffic crash data from Crash Records Information System (CRIS).
- **Task 4 – Safety Evaluation of Urban Roadway Networks:** The Project Team conducted a safety evaluation of different urban facility types in this task. The Project Team developed models at two temporal levels:
    - Long duration analysis (e.g., multi-year, annual).
    - Short duration analysis (e.g., daily, hourly).
- **Task 5 – Urban Decision Support Tool Development:** The Project Team developed an interactive map-based web application to analyze and visualize the risk on the urban roadway

network. The risk predicted using modeling for each segment of the roadway network was illustrated on interactive GIS maps. The decision support tool platform has three major components, a cloud-based data warehouse, a cloud-based computational platform, and a web server. The Project Team developed a dynamic version of the tool in which risk scores will be updated with the influx of new data.

- **Task 6 – Guideline and Workshop Presentation File Development:** The Project Team developed a guidance document on the tool usage and developed workshop-ready presentation slides.

1.3 REPORT ORGANIZATION

The remaining chapters of this report include the following:

- **Chapter 2: Literature Review:** This chapter provides an overview of the methods that can be used to perform highway safety evaluations.
- **Chapter 3: Data Preparation:** This chapter provides a brief overview of the data and the data conflation framework
- **Chapter 4: Model Development:** This chapter presents the development of annual and short-duration safety performance functions, including model specification, estimation results, and interpretation of key covariates across facility types.
- **Chapter 5: Decision Support Tool**: This chapter presents the safety decision tool that the research group developed for this study.
- **Chapter 6: Conclusions and Recommendations**: This chapter summarizes the most important research findings and provides a list of implementation recommendations stemming from the work performed and lessons learned throughout this project.

CHAPTER 2:
LITERATURE REVIEW

2.1 INTRODUCTION

This chapter provides a synthesis of methods that will help readers gain a better understanding of key aspects related to urban safety. These aspects include predictive methods for urban facilities, the consideration of non-motorists in overall safety calculations, the relationship among operating speed, geometric variables, posted speed limits, and crash outcomes, as well as the impact of weather on urban safety. To develop this synthesis, the Project Team gathered and reviewed relevant documentation, such as journal articles, research reports, guidebooks, and handbooks.

2.2 MODELING FRAMEWORKS

2.2.1 HSM Predictive Methods for Urban Non-Freeway Segments

The HSM AASHTO (2010), as described in Chapter 12, provides useful predictive methods for evaluating urban and suburban arterial systems. This document provides a well-organized technique for calculating the typical crash frequency, crash severity, and collision types anticipated in facilities with given characteristics. Except for collisions between bicycles and pedestrians, it covers all kinds of crashes involving automobiles, bicycles, and pedestrians. This predictive technique may be used for a variety of scenarios, such as current sites, design choices for existing sites, new sites, or alternate traffic volume forecasts.

Definition of Facility Types

The HSM provides precise definitions for facility types as well as prediction models for each facility type. The manual classifies facility types into two main categories: roadway segments and intersections. For this literature review, the Project Team solely focused on the facility types related to roadway segments. Below is a list of the specific site types related to roadway segments as presented in the HSM (AASHTO, 2010):

- **Two-lane undivided arterial (2U):** This refers to a road with two lanes that share a continuous cross-section, allowing for travel in two directions without any physical separation or barriers between the lanes.
- **Three-lane arterials (3T):** This depicts a road that has three lanes and a continuous cross-section that allows for two-way traffic. In this arrangement, the middle lane functions as a two-way left-turn lane (TWLTL).
- **Four-lane undivided arterials (4U):** This pertains to a road consisting of four lanes, without any physical separation or barriers between them, enabling travel in two directions on a continuous cross-section.

- **Four-lane divided arterials (4D):** This describes a continuous cross-sectional roadway with two lanes in each direction. Physical barriers, such as a raised or depressed median, are used to physically divide the lanes.
- **Five-lane arterials including a center TWLTL (5T):** This designates a two-way highway with five lanes that share a continuous cross-section. In this arrangement, the middle lane functions as a TWLTL.

Overview of the Predictive Method

The HSM uses a technique to calculate the expected average crash frequency for a particular location, and these calculations may be used to get the frequency for an entire facility or network. This estimation is done for a predetermined amount of time (in years), during which the geometric layout and traffic-controlling characteristics remain the same and the traffic volumes are either known or anticipated. Predictive models are used to provide these estimations, and the Empirical Bayes (EB) Method is used to integrate the results of the models with the data from the observed crashes. HSM's SPFs focus on two primary crash severity levels: crashes involving fatalities and injuries and crashes involving merely property damage. The term 'fatal-and-injury crashes' refers to collisions that result in injuries of any severity, including fatalities, incapacitating injuries, non-incapacitating injuries, and potential injuries. Equation (1) lists the prediction models used in the HSM for arterial facilities in urban and suburban areas to anticipate the average crash frequency (AASHTO, 2010).

$$N_{predicted} = \left(N_{spf\,x} \times \left(CMF_{1x} \times CMF_{2x} \times \ldots \times CMF_{yx}\right) + N_{pedx} + N_{bikex}\right) \times C_x \quad (1)$$

Where,

$N_{predicted}$ = projected annual average crash frequency for site type x;

$N_{spf\,x}$ = projected average crash frequency as estimated by the SPF generated for site type x under base circumstances;

N_{pedx} = anticipated annual average number of crashes between a vehicle and a pedestrian for site type x;

N_{bikex} = expected annual average number of vehicle-bicycle crashes for site type x;

CMF_{yx} = crash modification factor (CMF) unique to site type x; and

C_x = the calibration factor to modify SPF for the site type x's local circumstances.

Predictive Method Framework

Predictive models can be used to anticipate the average frequency of certain crash severity categories or types, or to estimate the total average crashes across all severities and crash types. The predictive model includes the SPF, CMFs, and a calibration factor when applied to a specific highway segment or junction. The predictive models are created to assess the predicted average crash frequency for crashes unrelated to junctions in the setting of highway segments. This covers crashes that take place near a junction but are unrelated to the intersection itself. In other

words, estimates of crashes that would occur regardless of whether an intersection is present are given by the predictive models for roadway segments. Equations (2) and (3) outline the specific predictive models for roadway segments (AASHTO, 2010).

$$N_{predicted\,rs} = C_r \times \left(N_{br} + N_{pedr} + N_{biker}\right) \quad (2)$$

$$N_{predicted} = N_{spf\,rs} \times (CMF_{1r} \times CMF_{2r} \times \ldots \times CMF_{nr}) \quad (3)$$

Where,

$N_{predicted\,rs}$ = projected average crash frequency of a specific roadway segment for the chosen year;

N_{br} = projected average crash frequency of a specific roadway segment;

$N_{spf\,rs}$ = projected total average crash frequency of a specific roadway segment for baseline conditions;

N_{pedr} = projected average crash frequency of vehicle-pedestrian collisions;

N_{biker} = projected average crash frequency of car-bicycle crashes;

$CMF_{1r} \ldots CMF_{nr}$= CMFs for roadway segments; and

C_r = calibration factor for a particular kind of road segment that was created for usage in a certain region.

Equation (2) demonstrates that the estimation of crash frequency for roadway segments comprises three components: N_{br}, N_{pedr}, N_{biker}. Moreover, the SPF portion of N_{br}, denoted as $N_{spf\,rs}$, can be broken down into three components based on collision type, as depicted in Equation (4).

$$N_{spf\,rs} = N_{brmv} + N_{brsv} + N_{brdwy} \quad (4)$$

Where,

N_{brmv} = projected average crash frequency of multiple vehicle non-driveway collisions for base conditions;

N_{brsv} = projected average crash frequency of single vehicle crashes for base conditions; and

N_{brdwy} = projected average multiple-vehicle driveway crash frequency.

Predictive Method Steps

The prediction approach consists of eighteen phases as listed in the HSM. For an urban or suburban arterial facility, the predictive method's application results in an estimate of the predicted average crash frequency. Steps 9, 10, and 11 of the predictive technique involve determining and applying the elements of the HSM predictive models. In some circumstances, some actions won't even be necessary. For instance, a new facility won't have observed crash data, therefore the EB method procedures don't need to be taken. The prediction techniques' steps are mentioned below (AASHTO, 2010).

- Step 1: Establish the borders of the different types of facilities and roads that make up the research network, site, or facility where we need to calculate the predicted average crash frequency, severity, and collision types.

- Step 2: Select the time frame.
- Step 3: Evaluate the availability of yearly average daily traffic volumes, pedestrian crossing volumes, and, if appropriate, witnessed collision data for an existing highway network.
- Step 4: List the geometric design aspects, traffic control features, and site characteristics for each site in the network of interest.
- Step 5: Segment the roadway network or facility into distinct uniform roadway sections and intersections, known as sites.
- Step 6: Allocate any observed collisions to the corresponding individual sites (if applicable).
- Step 7: Initiate the process with the first or subsequent individual site in the study network. If no further sites require assessment, proceed to Step 15.
- Step 8: Choose the first or subsequent year within the specified period of interest for the selected site. If there are no more years to evaluate that site, proceed to Step 14.
- Step 9: Determine and implement the suitable SPF for the type of facility and traffic control characteristics at the selected site.
- Step 10: Adjust the outcome from Step 9 by multiplying it with the relevant CMFs to account for specific geometric design and traffic control features of the site.
- Step 11: Multiply the result from Step 10 by the appropriate calibration factor.
- Step 12: If there are additional years to evaluate the chosen site within the study period, return to Step 8. Otherwise, proceed to Step 13.
- Step 13: Employ the site-specific EB Method if applicable.
- Step 14: If there are more sites to assess, return to Step 7. Otherwise, proceed to Step 15.
- Step 15: Utilize the project-level EB Method if the site-specific EB Method is not applicable.
- Step 16: Calculate the total collision frequency by aggregating all sites and years in the study.
- Step 17: Determine if there are alternative designs, treatments, or projected annual average daily traffic (AADT) to assess.
- Step 18: Evaluate and compare the results obtained.

Definition of Roadway Segments

The roadway is separated into discrete sites in step 5 of the predictive technique, comprising homogeneous roadway segments and junctions. The collection of these locations is referred to as a facility, and a highway network is made up of several facilities. A road segment begins at the center of an intersection and ends either at the center of the next junction or at the point where it changes to another homogeneous segment. The length of a highway segment that starts or ends at a junction is calculated from the intersection's center.

The approach of segmenting involves dividing the roadway network or facility into distinct road sections and intersections that share common characteristics, such as traffic volumes, important aspects of highway design, and traffic control measures. It is important to note that between two intersections, there may be multiple road segments that exhibit uniformity. A new and separate segment is initiated at each intersection or whenever there is a change in factors such as the

AADT, number of through lanes, presence of medians, the existence of TWLTL, type of on-street parking, density of roadside fixed objects, presence of lighting, speed category based on actual traffic speed or posted speed limit, and implementation of automated enforcement measures. Table 1 provides the recommended rounded widths for medians without barriers, which should be considered when determining 'homogeneous' segments (AASHTO, 2010).

Table 1. Rounded Widths for Medians, Reproduced from the HSM (AASHTO, 2010).

Measured Median Width	Rounded Median Width
1 ft to 14 ft	10 ft
15 ft to 24 ft	20 ft
25 ft to 34 ft	30 ft
35 ft to 44 ft	40 ft
45 ft to 54 ft	50 ft
55 ft to 64 ft	60 ft
65 ft to 74 ft	70 ft
75 ft to 84 ft	80 ft
85 ft to 94 ft	90 ft
95 ft or more	100 ft

SPFs for Urban Non-Freeway Segments

During Step 9 of the predictive approach, SPFs are employed to anticipate the frequency of crashes under specific baseline conditions. SPFs are regression models designed to estimate the average predicted crash frequency for individual roadway segments. These SPFs are constructed using observed crash data collected from comparable sites. Similar to other regression models, SPFs estimate the value of a dependent variable based on a set of independent variables. In the SPFs developed for the HSM, the dependent variable estimated is the average predicted crash frequency for a roadway segment or intersection under baseline conditions, while the independent variables are the AADT values for the roadway segment or intersection legs. This literature review primarily focuses on the SPFs relevant to roadway segments. Additionally, each SPF is accompanied by an overdispersion parameter, denoted as k, which indicates the statistical reliability of the SPF. A lower value of the overdispersion parameter signifies a greater level of statistical reliability. The SPFs in HSM for roadway segments are summarized in Table 2.

Table 2. SPFs for Urban and Suburban Arterials, Reproduced from the HSM (AASHTO, 2010).

SPFs for Urban and Suburban Arterials	SPF Components by Collision Type
Roadway segments (Non-Freeway)	multiple-vehicle nondriveway collisions
	single-vehicle crashes

	multiple-vehicle driveway-related collisions
	vehicle-pedestrian collisions
	vehicle-bicycle collisions

According to the HSM, the method for calculating the average collision frequency for a certain urban or suburban arterial route section is described in Equation (2) of the predictive model. While the influence of geometric design and traffic control elements is taken into account using CMFs, the impact of AADT on crash frequency is taken into consideration by incorporating the SPF. On urban and suburban arterials, five different types of highway segments have SPFs and adjustment factors available (AASHTO, 2010):

- Two-lane undivided arterials (referred to as 2U)
- Three-lane arterials with a center TWLTL (referred to as 3T)
- Four-lane undivided arterials (referred to as 4U)
- Four-lane divided arterials, which have a raised or depressed median (referred to as 4D)
- Five-lane arterials with a center TWLTL (referred to as 5T)

In the third step of the prediction method, the calculation of traffic volumes for the road segments included in the SPFs is explained. These SPFs are specifically developed for individual road segments on urban and suburban arterials, and they are designed to be applicable within certain ranges of AADT. However, it is important to note that deviating significantly from these defined AADT ranges, as indicated in the HSM, may result in inaccurate. The AADT ranges for each arterial type are as follows (AASHTO, 2010):

- 2U have an AADT range of 0 to 32,600 vehicles per day.
- 3T can accommodate traffic volumes ranging from 0 to 32,900 vehicles per day.
- 4U have an AADT range of 0 to 40,100 vehicles per day.
- 4D, which include a raised or depressed median, can handle traffic volumes ranging from 0 to 66,000 vehicles per day.
- 5Ts have an AADT range of 0 to 53,800 vehicles per day.

<u>Multiple-Vehicle Nondriveway Collisions</u>

The SPF for multiple-vehicle nondriveway collisions can be developed using Equation (5).

$$N_{brmv} = exp(a + b \times \ln(AADT) + \ln L) \quad (5)$$

Where,

$AADT$ = average annual daily traffic volume (vehicles/day) on roadway segment;

L = length of roadway segment (mi); and

a, b = regression coefficients.

Table 3 displays the coefficients a and b, as well as the overdispersion parameter k, used in the application of the Equation (5).

Table 3. SPF Coefficients for Multiple-Vehicle Nondriveway Collisions on Roadway Segments, Reproduced from the HSM (AASHTO, 2010).

Road Type	Coefficients Used in Equation (5)		Overdispersion Parameter (k)
	Intercept (a)	AADT (b)	
Total crashes			
2U	15.22	1.68	0.84
3T	12.40	1.41	0.66
4U	11.63	1.33	1.01
4D	12.34	1.36	1.32
Fatal-and-injury crashes			
2U	16.22	1.66	0.65
3T	16.45	1.69	0.59
4U	12.08	1.25	0.99
4D	12.76	1.28	1.31
5T	10.47	1.12	0.62
Property-damage-only crashes			
2U	15.62	1.69	0.87
3T	11.95	1.33	0.59
4U	12.53	1.38	1.08
4D	12.81	1.38	1.34
5T	9.97	1.17	0.88

First, Equation (5) is utilized to calculate N_{brmv} using the coefficients provided in Table 3 for total crashes. N_{brmv} is then dividied into two components based on severity, $N_{brmv(FI)}$ for fatal and injury (FI) crashes and $N_{brmv(PDO)}$ for property damage only (PDO) crashes. These preliminary values of $N_{brmv(FI)}$ and $N_{brmv(PDO)}$, desiganatec as $N'_{brmv(FI)}$ and $N'_{brmv(PDO)}$ in Equation (6), are determined with Equation (5) using the coefficients for FI and PDO crashes, respectively, in Table 3. To ensure that $N_{brmv(FI)}$ and $N_{brmv(PDO)}$ add up to N_{brmv}, the following Equations (6) and (7) are used. The ratios provided in Table 4 are utilized to divide $N_{brmv(FI)}$ and $N_{brmv(PDO)}$ into various collision categories (AASHTO, 2010) (see Table 4.

$$N_{brmv(FI)} = N_{brmv(total)} \left(\frac{N'_{brmv(FI)}}{N'_{brmv(FI)} + N'_{brmv(PDO)}} \right) \quad (6)$$

$$N_{brmv(PDO)} = N_{brmv(total)} - N_{brmv(FI)} \quad (7)$$

Table 4. Distribution of Multiple-Vehicle Nondriveway Collisions for Roadway Segments by Manner of Collision Type, Reproduced from the HSM (AASHTO, 2010).

Collision Type	Proportion of Crashes by Severity Level for Specific Road Types									
	2U		3T		4U		4D		5T	
	FI	PDO	FI	PDO	FI	PDO	FI	PDO	FI	PDO
Rear-end collision	0.730	0.778	0.845	0.842	0.511	0.506	0.832	0.662	0.846	0.651
Head-on collision	0.068	0.004	0.034	0.020	0.077	0.004	0.020	0.007	0.021	0.004
Angle collision	0.085	0.079	0.069	0.020	0.181	0.130	0.040	0.036	0.050	0.059
Sideswipe, same direction	0.015	0.031	0.001	0.078	0.093	0.249	0.050	0.223	0.061	0.248
Sideswipe, opposite direction	0.073	0.055	0.017	0.020	0.082	0.031	0.010	0.001	0.004	0.009
Other multiple-vehicle collisions	0.029	0.053	0.034	0.020	0.056	0.080	0.048	0.071	0.018	0.029

Single-Vehicle Crashes

According to the HSM, the SPF for multiple-vehicle non-driveway collisions can be developed using Equation (8) (AASHTO, 2010).

$$N_{brsv} = exp(a + b \times \ln(AADT) + \ln(L)) \quad (8)$$

The coefficients and factors used in the Equation (8) for each roadway type are presented in Table 5. Equation (8) is employed to calculate N_{brsv}. Subsequently, N_{brsv} is divided into two components, namely $N_{brsv(FI)}$ for FI crashes and $N_{brsv(PDO)}$ for PDO crashes. These initial values of $N_{brsv(FI)}$ and $N_{brsv(PDO)}$, denoted as $N'_{brsv(FI)}$ and $N'_{brsv(PDO)}$ in Equation (9), are determined using Equation (8) with the coefficients for FI and PDO crashes. To ensure that $N_{brsv(FI)}$ and $N_{brsv(PDO)}$ collectively add up to N_{brsv}, adjustments are made using the following equations. The ratios provided in Table 6 are utilized to divide $N_{brsv(FI)}$ and $N_{brsv(PDO)}$ into different components based on the type of collision (AASHTO, 2010).

$$N_{brsv(FI)} = N_{brsv(total)} \left(\frac{N'_{brsv(FI)}}{N'_{brsv(FI)} + N'_{brsv(PDO)}} \right) \quad (9)$$

$$N_{brsv(PDO)} = N_{brsv(total)} - N_{brsv(FI)} \quad (10)$$

Table 5. SPF Coefficients for Single-Vehicle Crashes on Roadway Segments, Reproduced from the HSM (AASHTO, 2010).

Road Type	Coefficients Used in Equation 12-11		Overdispersion Parameter (k)
	Intercept (a)	AADT (b)	
Total crashes			
2U	-5.47	0.56	0.81
3T	-5.74	0.54	1.37
4U	-7.99	0.81	0.91
4D	-5.05	0.47	0.86
5T	-4.82	0.54	0.52
Fatal-and-injury crashes			

Road Type	Coefficients Used in Equation 12-11		Overdispersion Parameter (k)
	Intercept (a)	AADT (b)	
2U	-3.96	0.23	0.50
3T	-6.37	0.47	1.06
4U	-7.37	0.61	0.54
4D	-8.71	0.66	0.28
5T	-4.43	0.35	0.36
Property-damage-only crashes			
2U	-6.51	0.64	0.87
3T	-6.29	0.56	1.93
4U	-8.50	0.84	0.97
4D	-5.04	0.45	1.06
5T	-5.83	0.61	0.55

Table 6. Distribution of Single-Vehicle Crashes for Roadway Segments by Collision Type, Reproduced from the HSM (AASHTO, 2010).

Collision Type	Proportion of Crashes by Severity Level for Specific Road Types									
	2U		3T		4U		4D		5T	
	FI	PDO	FI	PDO	FI	PDO	FI	PDO	FI	PDO
Collision with animals	0.026	0.066	0.001	0.001	0.001	0.001	0.001	0.063	0.016	0.049
Collision with fixed objects	0.723	0.759	0.688	0.963	0.612	0.809	0.500	0.813	0.398	0.768
Collision with other objects	0.010	0.013	0.001	0.001	0.020	0.029	0.028	0.016	0.005	0.061
Other single-vehicle collision	0.241	0.162	0.310	0.035	0.367	0.161	0.471	0.108	0.581	0.122

Multiple-Vehicle Driveway-Related Collisions

The prior model described collisions involving numerous vehicles but did not include those involving driveways. Due to the number and kind of driveways, crashes involving several cars that are caused by driveways are handled differently. Equation (11) may be used to calculate the total number of multiple-vehicle driveway-related crashes on a route stretch.

$$N_{brdwy} = \sum_{all\ driveway\ ypes} n_i \times N_j \times \left(\frac{AADT}{15{,}000}\right)^{(t)} \quad (11)$$

Where N represents the annual number of collisions per driveway for a specific driveway type, denoted as j. n_i corresponds to the count of driveways within the roadway segment of the same driveway type, encompassing both sides of the road. The coefficient t is utilized for adjusting the impact of traffic volume.

To compute the quantity n_i, which represents the number of driveways of a specific type, the count is obtained by summing the number of driveways of that type of present on both sides of the road. The calculation of driveway counts is performed independently for each side of the

road, and the results are combined. In the modeling process, the HSM considers seven different types of driveways, as outlined in Table 7 (AASHTO, 2010).

Table 7. SPF Coefficients for Multiple-Vehicle Driveway Related Collisions, Reproduced from the HSM (AASHTO, 2010).

Driveway Type (j)	Coefficients for Specific Roadway Types				
	2U	3T	4U	4D	5T
Number of Driveway-Related Collisions per Driveway per Year (N_j)					
Major commercial	0.158	0.102	0.182	0.033	0.165
Minor commercial	0.050	0.032	0.058	0.011	0.053
Major industrial/institutional	0.172	0.110	0.198	0.036	0.181
Minor industrial/institutional	0.023	0.015	0.026	0.005	0.024
Major residential	0.083	0.053	0.096	0.018	0.087
Minor residential	0.016	0.010	0.018	0.003	0.016
Other	0.025	0.016	0.029	0.005	0.027
Regression Coefficient for AADT (t)					
All driveways	1.000	1.000	1.172	1.106	1.172
Overdispersion Parameter (k)					
All driveways	0.81	1.10	0.81	1.39	0.10
Proportion of Fatal-and-Injury Crashes (f_{lwy})					
All driveways	0.323	0.243	0.342	0.284	0.269
Proportion of Property-Damage-Only Crashes					
All driveways	0.677	0.757	0.658	0.716	0.731

Vehicle-Pedestrian Collisions

According to the HSM, Equation 14 can be utilized to estimate the annual count of vehicle-pedestrian collisions for a roadway segment (AASHTO, 2010).

$$N_{pedr} = N_{br} \times f_{pedr} \quad (12)$$

Where,

f_{pedr} = crash adjustment factor.

Equation (3) is applied to compute the value of N_{br}, which is then used in the Equation (12). The coefficients for f_{pedr} utilized in Equation (12) are listed in Table 8.

Table 8. Pedestrian Crash Adjustment Factor for Roadway Segments, Reproduced from the HSM (AASHTO, 2010).

Road type	Bicycle Crash Adjustment Factor (f_{biker})	
	Posted Speed 30 mph or Lower	Posted Speed Greater than 30 mph
2U	0.036	0.005
3T	0.041	0.013
4U	0.022	0.009
4D	0.067	0.019
5T	0.030	0.023

Vehicle-Bicycle Collisions

According to the HSM, Equation 14 can be utilized to estimate the annual count of vehicle-bicycle collisions for a roadway segment (AASHTO, 2010).

$$N_{biker} = N_{br} \times f_{biker} \quad (13)$$

Where,

f_{biker} = crash adjustment factor.

The value of N_{br} utilized in Equation (13) is calculated with the Equation (3). The coefficients for f_{biker} utilized in Equation (13) are listed in Table 9.

Table 9. Bicycle Crash Adjustment Factors for Roadway Segments, Reproduced from the HSM (AASHTO, 2010).

Road type	Bicycle Crash Adjustment Factor (f_{biker})	
	Posted Speed 30 mph or Lower	Posted Speed Greater than 30 mph
2U	0.018	0.004
3T	0.027	0.007
4U	0.011	0.002
4D	0.013	0.005
5T	0.050	0.012

CMFs for Roadway Segments

According to the HSM, the selected SPF chosen in Step 9 is subjected to CMFs during Step 10 of the prediction technique. CMFs are used to modify the projected average crash frequency based on certain geometric features and traffic control factors, as stated in the general predictive model. Each feature has a CMF of 1.00 for its base condition. The CMF of a feature exceeds 1.00 if it is linked to a higher crash frequency than the baseline condition. A feature's CMF is smaller than 1.00 if, on the other hand, it is linked to a lower crash frequency (AASHTO, 2010) (see Table 10).

Table 10. Summary of CMFs in HSM and the Corresponding SPFs for Roadway Segments, Reproduced from the HSM (AASHTO, 2010).

Applicable SPF	CMF	CMF Description
Roadway Segments	CMF_{1r}	On-Street Parking
	CMF_{2r}	Roadside Fixed Objects
	CMF_{3r}	Median Width
	CMF_{4r}	Lighting
	CMF_{5r}	Automated Speed Enforcement

2.2.2 HSM Predictive Methods for Urban Freeway Segments

The supplement to HSM, published in 2014, describes the predictive methods for freeways in Chapter 18. The following section offers a high-level overview of the freeway predictive methods, their associated SPFs, CMFs, and the steps involved in the predictive methods.

Overview of Predictive Methods

The HSM outlines an 18-step procedure for estimating the anticipated crash frequency of two types of facilities: freeway segments and freeway speed change lanes. According to the HSM, a speed-change lane is defined as an uncontrolled transition area between a ramp and a freeway. The predictive models follow a general format described by Equation (14). It is important to note that these predictive equations can generate estimated crash frequencies for overall crashes, specific crash types, and different severity levels (AASHTO, 2014).

$$N_{p,w,x,y,z} = N_{spf,\mathrm{w},x,y,z} \times \left(CMF_{1,\mathrm{w},x,y,z} \times CMF_{2,\mathrm{w},x,y,z} \times \ldots \times CMF_{m,w,x,y,z}\right) \times C_{w,x,y,z} \quad (14)$$

Where, w, x, y, z, and m represent the type of site, control type, crash, severity, and traffic control features, respectively; $N_{p,w,x,y,z}$ represents the projected annual average crash frequency; $N_{spf,\mathrm{w},x,y,z}$ represents projected average crash frequency as estimated by the SPF under base circumstances; $CMF_{m,w,x,y,z}$ represents CMF for specific geometric design and traffic control features; and $C_{w,x,y,z}$ represents the adjustment factors for local conditions.

Definition of Freeway Facility and Site Types

The HSM defines a freeway as a type of roadway that has complete access control and separation from intersecting roads. Access to freeways is limited to interchanges with grade separation. The classification of an area as urban, suburban, or rural takes into account various factors such as road characteristics, population density, and land use. Urban areas, according to Federal Highway Administration (FHWA) guidelines, are designated as locations within urban boundaries with a population exceeding 5,000. On the other hand, rural areas encompass regions outside urban centers where the population is below 5,000. The HSM also uses the term suburban to refer to areas on the outskirts of urban regions. Note that this literature review only discussed the freeway prediction models related to urban facilities. To predict crash frequencies, HSM employs specific methods developed for three types of facilities: rural freeway segments with four to eight lanes, urban freeway segments with four to ten lanes, and freeway speed-change lanes associated with entrance and exit ramps. Additionally, the HSM further classifies freeway segments into four categories based on the number of lanes: four-lane, six-lane, eight-lane, and ten-lane segments (AASHTO, 2014).

SPFs for Urban Freeways

The HSM includes a set of sixteen SPFs for urban freeway segments, corresponding to four specific facility types as mentioned in the previous section. These SPFs are listed in Table 11. Furthermore, the HSM provides an additional set of sixteen SPFs for speed-change lanes, corresponding to eight distinct facility types, as listed in Table 12 (AASHTO, 2014).

Table 11. SPFs for Urban Freeway Segments (fs), Reproduced from the HSM (AASHTO, 2014).

Cross Section (x)	Crash Type (y)	Crash Severity (z)	SPF
Four-lane divided (4)	Multiple vehicle (mv)	Fatal and injury (fi)	$N_{spf,fs,4,mv,fi}$
		Property damage only (pdo)	$N_{spf,fs,4,mv,pdo}$
	Single vehicle (mv)	Fatal and injury (fi)	$N_{spf,fs,4,sv,fi}$
		Property damage only (pdo)	$N_{spf,fs,4,sv,pdo}$
Six-lane divided (6)	Multiple vehicle (mv)	Fatal and injury (fi)	$N_{spf,fs,6,mv,fi}$
		Property damage only (pdo)	$N_{spf,fs,6,mv,pdo}$
	Single vehicle (mv)	Fatal and injury (fi)	$N_{spf,fs,6,sv,fi}$
		Property damage only (pdo)	$N_{spf,fs,6,sv,pdo}$
Eight-lane divided (8)	Multiple vehicle (mv)	Fatal and injury (fi)	$N_{spf,fs,8,mv,fi}$
		Property damage only (pdo)	$N_{spf,fs,8,mv,pdo}$
	Single vehicle (mv)	Fatal and injury (fi)	$N_{spf,fs,8,sv,fi}$
		Property damage only (pdo)	$N_{spf,fs,8,sv,pdo}$
Ten-lane divided (10)	Multiple vehicle (mv)	Fatal and injury (fi)	$N_{spf,fs,10,mv,fi}$
		Property damage only (pdo)	$N_{spf,fs,10,mv,pdo}$
	Single vehicle (mv)	Fatal and injury (fi)	$N_{spf,fs,10,sv,fi}$
		Property damage only (pdo)	$N_{spf,fs,10,sv,pdo}$

Table 12. SPFs for Urban Freeway Speed-Change Lane (sc), Reproduced from the HSM (AASHTO, 2014).

Cross Section (x)	Crash Type (y)	Crash Severity (z)	SPF
Ramp entrance to four-lane divided ($4EN$)	All types (at)	Fatal and injury (FI)	$N_{spf,sc,4EN,at,fi}$
		Property damage only (PDO)	$N_{spf,sc,4EN,at,pdo}$
Ramp entrance to six-lane divided ($6EN$)	All types (at)	Fatal and injury (FI)	$N_{spf,sc,6EN,at,fi}$
		Property damage only (PDO)	$N_{spf,sc,6EN,at,pdo}$
Ramp entrance to eight-lane divided ($8EN$)	All types (at)	Fatal and injury (FI)	$N_{spf,sc,8EN,at,fi}$
		Property damage only (PDO)	$N_{spf,sc,8EN,at,pdo}$
Ramp entrance to ten-lane divided ($10EN$)	All types (at)	Fatal and injury (FI)	$N_{spf,sc,10EN,at,fi}$
		Property damage only (PDO)	$N_{spf,sc,10EN,at,pdo}$
Ramp exit from four-lane divided ($4EX$)	All types (at)	Fatal and injury (FI)	$N_{spf,sc,4EX,at,fi}$
		Property damage only (PDO)	$N_{spf,sc,4EX,at,pdo}$
Ramp exit from six-lane divided ($6EX$)	All types (at)	Fatal and injury (FI)	$N_{spf,sc,6EX,at,fi}$
		Property damage only (PDO)	$N_{spf,sc,6EX,at,pdo}$
Ramp exit from eight-lane divided ($8EX$)	All types (at)	Fatal and injury (FI)	$N_{spf,sc,8EX,at,fi}$
		Property damage only (PDO)	$N_{spf,sc,8EX,at,pdo}$
	All types (at)	Fatal and injury (FI)	$N_{spf,sc,10EX,at,fi}$

Cross Section (x)	Crash Type (y)	Crash Severity (z)	SPF
Ramp exit from ten-lane divided ($10EX$)		Property damage only (PDO)	$N_{spf,sc,10EX,at,pdo}$

Predictive Model for Freeway Segments

The predictive model for freeway segments described in the HSM can be defined using Equations (15) to (19) (AASHTO, 2014).

$$N_{p,fs,n,at,as} = N_{p,fs,n,mv,fi} + N_{p,fs,n,sv,fi} + N_{p,fs,n,mv,pdo} + N_{p,fs,n,sv,pdo} \quad (15)$$

$$N_{p,fs,n,mv,fi} = C_{fs,ac,mv,fi} \times N_{spf,fs,n,mv,fi} \times (CMF_{1,fs,ac,mv,fi} \times .. \times CMF_{m,fs,ac,mv,fi}) \times (CMF_{1,fs,ac,at,fi} \times .. \times CMF_{m,fs,ac,at,fi}) \quad (16)$$

$$N_{p,fs,n,sv,fi} = C_{fs,ac,sv,fi} \times N_{spf,fs,n,sv,fi} \times (CMF_{1,fs,ac,sv,fi} \times .. \times CMF_{m,fs,ac,sv,fi}) \times (CMF_{1,fs,ac,at,fi} \times .. \times CMF_{m,fs,ac,at,fi}) \quad (17)$$

$$N_{p,fs,n,mv,pdo} = C_{fs,ac,mv,pdo} \times N_{spf,fs,n,mv,pdo} \times (CMF_{1,fs,ac,mv,pdo} \times .. \times CMF_{m,fs,ac,mv,pdo}) \times (CMF_{1,fs,ac,at,pdo} \times .. \times CMF_{m,fs,ac,at,pdo}) \quad (18)$$

$$N_{p,fs,n,sv,pdo} = C_{fs,ac,sv,fi} \times N_{spf,fs,n,sv,pdo} \times (CMF_{1,fs,ac,sv,pdo} \times .. \times CMF_{m,fs,ac,sv,pdo}) \times (CMF_{1,fs,ac,at,pdo} \times .. \times CMF_{m,fs,ac,at,pdo}) \quad (19)$$

Where, $N_{p,fs,n,y,z}$ represents the projected average crash frequency of a freeway segment with n lanes, crash type y ($y = sv$: single vehicle, mv: multiple vehicle, at: all types), and severity z; $N_{spf,fs,n,y,z}$ represents the projected average crash frequency of a freeway segment with base conditions; $CMF_{n,fs,ac,y,z}$ represents CMF for a freeway segment with any cross-section ac, feature m, crash type y, and severity z; and $C_{fs,ac,y,z}$ represents calibration factors for freeway segments.

Predictive Model for Urban Freeway Speed-Change Lanes

The predictive model for freeway speed-change lanes described in the HSM can be defined using Equations (20) to (22) (AASHTO, 2014).

$$N_{p,sc,nEN,at,as} = N_{p,sc,nEN,at,fi} + N_{p,sc,nEN,at,pdo} \quad (20)$$

$$N_{p,sc,nEN,at,fi} = C_{sc,EN,at,fi} \times N_{spf,sc,nEN,at,fi} \times (CMF_{1,sc,nEN,at,fi} \times .. \times CMF_{m,sc,nEN,at,fi}) \times (CMF_{1,sc,ac,at,fi} \times .. \times CMF_{m,sc,ac,at,fi}) \quad (21)$$

$$N_{p,sc,nEN,at,pdo} = C_{sc,EN,at,pdo} \times N_{spf,sc,nEN,at,pdo} \times (CMF_{1,sc,nEN,at,pdo} \times .. \times CMF_{m,sc,nEN,at,pdo}) \times (CMF_{1,sc,ac,at,pdo} \times .. \times CMF_{m,sc,ac,at,pdo}) \quad (22)$$

Where, $N_{p,sc,nEN,at,as}$ represents the projected average annual crash frequency of ramp entrance speed-change lane on a freeway with n lanes, all crash types at, and severity z ($z = fi$: fatal and injury, pdo: property damage only, as: all severities); $N_{spf,sc,nEN,at,z}$ represents the projected average annual crash frequency of ramp entrance speed-change lane on a freeway considering base condition; $CMF_{n,sc,x,at,z}$ represents CMF for a speed-change

lane with feature m, cross-section x ($x = nEN$: ramp entrance adjacent to a freeway with n lanes, nEX: ramp exit adjacent to a freeway with n lanes, ac: any cross-section), all crash types at, and severity z; and; and $C_{sc,EN,at,z}$ represents the calibration factor for a ramp entrance speed-change lane with all crash types and severity.

Predictive Method Steps for Urban Freeway Segments

The HSM outlines a procedure consisting of eighteen steps for a predictive method on freeways. These steps are as follows (AASHTO, 2014).

- Step 1: Establish the project boundaries.
- Step 2: Determine the timeframe of interest.
- Step 3: AADT volumes and observed crash data for the study period, particularly for existing projects, to determine if the EB Method is applicable.
- Step 4: Identify the geometric design features, traffic control features, and site characteristics for all locations within the project boundaries.
- Step 5: Divide the roadway into distinct sites.
- Step 6: Assign recorded collisions to the individual sites (if relevant).
- Step 7: Select the initial or subsequent specific site within the project boundaries.
- Step 8: For the chosen site, designate the first or next year in the designated time frame. If there are no more years to evaluate that site, move to Step 13.
- Step 9: Identify the geometric design features, traffic control features, and site characteristics for all locations within the project boundaries.
- Step 10: Multiply the outcome from Step 9 by the appropriate CMFs.
  Step 11: Multiply the outcome from Step 10 by the suitable calibration factor.
- Step 12: If there are additional years to assess within the evaluation period for the selected site, return to Step 8. Otherwise, proceed to Step 13.
- Step 13: Apply the site-specific EB Method (if applicable) and incorporate SDFs.
- Step 14: If there are additional sites to evaluate, return to Step 7; otherwise, proceed to Step 15.
- Step 15: Implement the project-wide EB Method (if relevant) and incorporate SDFs.
- Step 16: Aggregate the data from all sites and years in the study to estimate the overall crash frequency.
- Step 17: Assess whether there are any alternative designs, treatments, or projected AADT to be examined.
- Step 18: Analyze and compare the outcomes.

CMFs for Urban Freeway Segments

The HSM has provided a comprehensive list of CMFs that are applicable to the SPFs discussed in the previous section.

Table 13 presents a detailed compilation of CMFs categorized by specific site types, cross sections, crash types, and crash severities (AASHTO, 2014).

Table 13. Freeway CMFs and their Corresponding SPFs (AASHTO, 2014).

Applicable SPF(s)	CMF Variable (z)	SPF
Freeway segments or speed-change lanes	$CMF_{1,w,x,y,z}$	Horizontal curve
	$CMF_{2,w,x,y,fi}$	Lane width
	$CMF_{3,w,x,y,z}$	Inside shoulder width
	$CMF_{4,w,x,y,z}$	Median width
	$CMF_{5,w,x,y,z}$	Median barrier
	$CMF_{6,w,x,y,z}$	High volume
Multiple-vehicle crashes on freeway segments	$CMF_{7,fs,ac,mv,z}$	Lane change
Single-vehicle crashes on freeway segments	$CMF_{8,fs,ac,sv,z}$	Outside shoulder width
	$CMF_{9,fs,ac,sv,fi}$	Shoulder rumble strip
	$CMF_{10,fs,ac,sv,fi}$	Outside clearance
	$CMF_{11,fs,ac,sv,z}$	Outside barrier
Ramp entrances	$CMF_{12,sc,nEN,at,z}$	Ramp entrance
Ramp exits	$CMF_{13,sc,nEX,at,z}$	Ramp exit

2.2.3 Safety Prediction Models

Researchers have utilized a wide variety of modeling techniques to explore the association between operating speed, posted speed, roadway geometry, traffic exposures, and safety. The primary modeling approaches used for safety prediction models up to this point may be generally divided into two categories: statistical approaches and machine learning/ data mining approaches. This section provides a brief overview of state-of-the-art modeling techniques.

Statistical Approaches

Poisson Regression

The Poisson regression model has been proposed as a viable method to examine the connection between factors affecting risk and the modeling of traffic crashes. The utilization of Poisson regression has been extensive in analyzing data related to transportation counts, particularly when studying the frequency of crashes. Since crash-frequency data consists of non-negative whole numbers, traditional ordinary least-squares regression, which assumes a continuous dependent variable, is not suitable. As a result, researchers have recently embraced the Poisson regression model as a framework for modeling crash-frequency data, considering its appropriateness for analyzing non-negative integer dependent variables. Presented below is a mathematical representation of the Poisson regression model (Lord and Mannering, 2010).

$$P(y_i) = \frac{EXP(-\lambda_i)\lambda_i^{y_i}}{y_i!} \quad (23)$$

Where the probability of a roadway entity, denoted as i, experiencing a specific number of crashes per period, represented as y_i, is denoted by $P(y_i)$. The Poisson parameter, λ_i, for roadway entity i is equal to the expected number of crashes per year, $E(y_i)$. To estimate Poisson regression models, the Poisson parameter λ_i, which signifies the expected number of crashes per period, is determined as a function of explanatory variables. The most used functional form is $\lambda_i = EXP(\beta X_i)$, where X_i represents a vector of explanatory variables and β represents a vector of estimable parameters.

Negative Binomial Regression

To account for potential excessive scattering in the data, the negative binomial (NB) model has been developed as an expansion of the Poisson model. The Poisson-Gamma approach serves as an alternative term for this technique. Within the negative binomial model, the Poisson parameter relies on a gamma probability distribution, leading to a closed-form equation. Manipulating the relationship between the mean and variance structures is straightforward in this model. By incorporating over-dispersion into crash data counts, this method relaxes the requirement of equal mean and variance. The negative binomial model accommodates possible over-dispersion in crash data counts through the utilization of a Gamma probability distribution. By introducing an error term to the expected number of crashes in the Poisson regression, the modified equation yields the negative binomial model (Lord and Mannering, 2010).

$$\lambda_i = EXP(\beta X_i + \varepsilon_i) \quad (24)$$

Where $EXP(\varepsilon_i)$ is a gamma-distributed error term with mean 1 and variance α. Negative binomial models have garnered considerable attention in the realm of crash frequency analysis. Khattak et al. (2021) conducted a study utilizing a negative binomial regression model to establish SPFs for urban intersections. The findings suggested that employing negative binomial models has the potential to enhance the accuracy of estimating SPFs for urban intersections. Similarly, Rista et al. (2018) employed negative binomial models to assess the safety implications of narrow lane widths on urban and suburban arterials. Kim and Washington, (2006) utilized this modeling approach to develop crash models specifically for intersections with a focus on left-turn lanes. Additionally, Daniels et al. (2010) employed negative binomial models to investigate safety performance at roundabouts.

Poisson-Lognormal Regression

To address the limitations associated with NB models, researchers have developed the Poisson-lognormal (PLN) model. In contrast to the NB model, the PLN model assumes a PLN error term instead of a gamma distribution, making it better suited for handling under-dispersed data counts. Although the PLN model shares similarities with the NB model, it introduces a lognormal-distributed the $EXP(\varepsilon_i)$ term in the model, resulting in increased flexibility. However, the PLN

model also presents certain drawbacks, such as the lack of a closed form for the PLN distribution, leading to more complex parameter estimation compared to the NB model (Abdulhafedh, 2017).

Several scholars have employed Poisson modeling methods to examine the correlation between crashes and relevant risk factors, particularly in urban settings. A study conducted by Park and Lord (2007) introduced a new approach called the multivariate Poisson-lognormal model (MVPLN) to simultaneously analyze crash count data based on severity. Additionally, Zhao et al., (2018) utilized the MVPLN model to investigate crashes occurring at signalized intersections in urban areas. The results of their study demonstrated that the MVPLN model provides a better fit when compared to the traditional univariate Poisson model.

Zero Inflated Poisson and Negative Binomial Regression

Two commonly utilized methodologies for modeling crash frequencies are the zero-inflated Poisson and zero-inflated negative binomial models. These models were developed to tackle the problem of over-dispersion arising from an abundance of zero counts in traffic data, which occurs when no crashes are observed at specific locations. By employing a zero-altered process, these models enable the representation of crash frequencies in two distinct states: the state of zero crashes and the state of non-zero crashes. The likelihood of a section being in either state can be determined through the utilization of a binary logit or probit model. The presence of a significant number of zero observations in crash data is typically attributed to underreporting of minor crashes, the presence of nearby hazardous crash sites that render the observed sites relatively safe, and the absence of certain types of crashes at these specific locations. The objective of zero-inflated models is to accommodate these excess zeros by assuming a dual-state crash system in which one state represents the absence of crashes, indicating virtual safety during the observation period, while the other state represents the occurrence of non-zero crashes (Abdulhafedh, 2017).

The zero-inflated model has gained considerable traction among transportation safety researchers, and its application has provided valuable insights into various aspects of road safety. For instance, in a comprehensive study conducted by Liu et al. (2018), the zero-inflated model was utilized to examine urban mid-block crashes, leading to improved predictions of crash frequency and severity compared to traditional models. Similarly, Raihan et al. (2019) employed the zero-inflated negative binomial model to develop CMFs for bicycle crashes in urban areas, identifying contributing factors and recommending safety countermeasures. Kumara and Chin (2003) also applied the zero-inflated negative binomial model to investigate crash occurrence at signalized tee intersections, successfully capturing the excess zeros and overdispersion commonly observed in crash data. Moreover, researchers such as Carson and Mannering (2001), Lee and Mannering (2002), Shankar et al., (2003) have utilized the zero-inflated model to analyze crash data and identify factors influencing crashes on different types of roadways.

Conway–Maxwell–Poisson

The Conway-Maxwell-Poisson distribution is an advanced statistical distribution that can be applied to model queues and service rates. Unlike the traditional Poisson distribution, it can handle both under-dispersed and over-dispersed data, thus providing a more flexible approach for various types of crash-frequency data. The distribution also encompasses several other probability density functions, such as the geometric, Bernoulli, and Poisson distributions, which allows its utilization in a broader range of applications beyond crash-frequency modeling. By utilizing the Conway-Maxwell-Poisson distribution, researchers can expand their analytical capabilities when examining crash-frequency data. The versatility and flexibility of this distribution make it an invaluable tool for transportation safety analysts who aim to accurately model and predict crash occurrences. Many studies have implemented this modeling technique to analyze crash data and develop safety prediction models (Lord et al., 2008; Sellers and Shmueli, 2010).

Logit and Probit Models

Logit and probit models are commonly employed to capture the severity of crash data, considering the multitude of factors associated with crash occurrences. Binary models, which accommodate two possible outcomes, are preferable when modeling crash severity since the dependent variable often comprises multiple outcome categories. While discriminant analysis is a viable alternative, logit and probit models are generally favored due to their flexibility and ability to handle non-linear relationships. The classification of traffic crash severity models can be categorized as either nominal or ordinal, with no consensus on the optimal approach as the choice of model relies heavily on the characteristics of the data. Some researchers favor nominal models due to the potential shared unobserved effects among adjacent injury categories, while others prefer ordinal models for their simplicity and overall performance. Additionally, multinomial models, accommodating three or more outcomes, can be employed to model crash severity (Abdulhafedh, 2017). Numerous studies have utilized these modeling techniques to investigate crash severity. For example, Chen and Fan (2019) conducted a study to identify significant contributing factors to pedestrian injury severity in pedestrian-vehicle crashes in both rural and urban areas of North Carolina, United States. Similarly, logit and probit models have been applied in freight crash analysis (Doustmohammadi, 2019) and the examination of intersection-related crashes (Tay, 2015). Other studies have utilized logit and/or probit models to analyze safety in urban networks (Asil and Bargegol, 2022; Haleem and Gan, 2013; Intini et al., 2020).

Random-Parameter Models

Random-parameter models represent an extension of random-effects models, which allow estimated parameters to vary across each individual observation in the dataset, rather than just affecting the model intercept. This modeling approach aims to capture the unobserved heterogeneity among different roadway sites. The primary motivation for using random-parameter models is to account for this heterogeneity that cannot be explained by observed variables alone. Such models assume that estimated parameters vary across observations

according to a particular distribution. Researchers have utilized this modeling technique to investigate various road safety issues, such as the influence of traffic, geometric, and context variables on urban crash types (Intini et al., 2020), assessing the impact of traffic signal performance on crash frequency for signalized intersections along urban (Kabir et al., 2021), examining the safety impacts of narrow lane widths on urban/suburban arterials (Rista et al., 2018), and investigating motorcyclist injury severities (Se et al., 2021).

Other Statistical Models

In addition to the modeling techniques, there are several other statistical models that are commonly used to investigate traffic safety. These include the random-parameters model, gamma model, generalized estimating equation, generalized additive model, negative multinomial model, hierarchical/multilevel model, finite mixture/Markov switching model, and Bayesian approach. Each of these models has its strengths and weaknesses, and the choice of model depends on the research question and the characteristics of the data (Lord and Mannering, 2010).

Machine Learning/ Data Mining Approaches

Classification and Regression Trees (CART)/ Decision Tree (DT)

DT or CART are widely used machine learning methods that construct classification or regression models in the form of a tree, where each node represents a predictor variable, and each branch corresponds to a decision rule or criterion. At each step, the explanatory variable that achieves the "best" split is chosen to predict the value of the response variable. The simplicity and interpretability of DTs make them a popular choice for modeling problems in various domains, including traffic safety. However, overfitting is a common issue with DTs, which can lead to poor generalization performance. The application of DTs in crash severity modeling dates back to early machine learning research. For instance, (Kuhnert et al., 2000) compared the performance of logistic regression, CART, and multivariate adaptive regression splines for modeling motor vehicle crashes and found that CART was capable of selecting significant independent variables. Similarly, Chang and Chien (2013) developed a CART model to investigate the relationship between injury severity outcomes and various driver/vehicle, highway geometric, environmental, and crash-related factors. Recent studies have applied variations of the DT technique to explore the relationship between crash severity and other factors. For example, Prati et al. (2017) used a modified DT method to identify the most significant factors contributing to bicycle crashes, including crash characteristics, infrastructure characteristics, cyclists' demographics, and environmental factors. Additionally, Arefkhani et al. (2021) applied CART to identify the most significant factors contributing to drivers' injury status.

K-Nearest-Neighbor (KNN)

KNN is a non-parametric statistical technique applicable to classification and regression tasks. Its mechanism involves identifying k neighboring values close to the response variable, based on

the calculated distance between this variable and other known observations. Subsequently, the predicted value of the response variable is computed. One of the advantages of KNN is its lack of assumptions about the data's functional form, though it may be sensitive to the local data structure (Iranitalab and Khattak, 2017). The study conducted by Zhang et al. (2018) involved the development of multiple machine learning models, including KNN, for the prediction of crash injury severity. The results obtained from these models were compared with traditional statistical models, and it was reported that KNN outperformed the other models in terms of prediction accuracy. KNN has been applied in various transportation safety studies, such as investigating speed violations (Kuşkapan et al., 2021), calibrating SPFs (Farid et al., 2018), and assessing crash risk at diverging areas (Xing et al., 2020).

Support Vector Machines (SVM)

SVM is a discriminative classifier that operates by identifying an optimal hyperplane in a space with high dimensions or even infinite dimensions. This hyperplane serves as a decisive boundary, effectively separating the data into two distinct classes. The margin, which refers to the distance between the hyperplane and the closest data point on either side, plays a crucial role in determining the performance of SVM. The primary objective of an SVM classifier is to select a hyperplane with the widest margin between the two classes, thereby enhancing the accuracy of classifying new data instances (Bambrick, 2016; Hastie et al., 2001). Researchers have utilized SVM in various applications, such as real-time crash risk assessment and the analysis of crash severity on mountainous highways, incorporating real-time traffic and weather data. For instance, Yu and Abdel-Aty (2013) conducted a study comparing SVM models with different kernel functions against logistic regression models to evaluate real-time crash risk. The findings highlighted the superior performance of SVM compared to logistic regression. In a subsequent study, the same authors employed SVM to examine crash severity on mountainous highways, incorporating real-time traffic and weather data (Yu and Abdel-Aty, 2014a). Sun et al. (2014) also utilized SVM models to predict real-time crash risk on urban expressways.

Random Forest

Random Forest (RF) is a supervised classification algorithm in machine learning that constructs a collection of Decision Trees using a random subset of training data. The predictions of each tree are then combined through voting to make the final prediction. This approach is commonly employed to create a forest of multiple classification trees, resulting in more precise predictions. RF offers several advantages including the ability to handle missing values, avoid overfitting by using random subsets of data, and provide information about the importance of different features in the dataset. This feature importance can aid in identifying significant contributing factors during model training. RF has been widely applied in various road safety domains to predict different outcomes. For example, Wang and Kim (2019) utilized RF to identify factors and predict crash severity, comparing its performance to multinomial logit (MNL). Their findings indicated that RF outperformed MNL in terms of prediction accuracy, as measured by precision, recall, and F1 score. Although the differences were not statistically significant, sensitivity analysis revealed that RF was less sensitive than MNL. Moreover, RF has the capability to

capture nonlinear effects of continuous variables and mitigate the impact of collinearity among explanatory variables. In another study, Wahab and Jiang (2019) developed three machine learning models, including Decision Trees, RF, and Instance-Based learning with parameter k (IBk), to model injury severity in motorcycle crashes. They observed that RF-based algorithms exhibited better agreement with experimental data compared to the other two algorithms, attributed to RF's global optimization and extrapolation capabilities. Dash et al. (2022) applied RF to investigate factors influencing bike crash severity in urban areas and found that RF demonstrated potential for greater explanatory accuracy. This finding was particularly noteworthy given the limited use of RF in bike safety studies. Similarly, R. Yu et al. (2019) employed RF to rank the most critical factors associated with crash risk on the expressway system.

eXtreme Gradient Boosting (XGBoost)

XGBoost is a popular and efficient implementation of Gradient Boosting, an ensemble learning method that utilizes a series of weak learners to build a predictive model. In contrast to other ensemble methods, Gradient Boosting constructs individual models in a sequential manner and adjusts their weights based on information from previous models. XGBoost improves upon Gradient Boosting by incorporating a more precise estimation method that includes information on the gradient direction and minimum loss function. It also uses regularized boosting to minimize overfitting and enhance performance. XGBoost is governed by several parameters, including shrinkage, boosting iterations, minimum loss reduction, and decision tree-related parameters (A. Das et al., 2022). Recent studies have demonstrated the versatility of XGBoost in the safety analysis. For example, Goswamy et al. (2023) utilized XGBoost to identify the factors that influence the severity of injuries at pedestrian crossing locations with flashing beacons. Similarly, Jiang and Ma (2021) employed XGBoost to investigate macro factors associated with traffic fatality rules. XGBoost has also been used in conjunction with SHapley Additive exPlanations, an explainable AI-based model, to examine various aspects of safety analysis. For example, Chang et al. (2022) investigated fatal pedestrian crashes, Parsa et al. (2020) conducted real-time crash detection and feature analysis, and Yang et al. (2021) examined factors in freight truck-related crashes.

Artificial Neural Network (ANN)

The Artificial Neural Network (ANN) is a computational system inspired by the structure of the human brain that comprises numerous interconnected neurons arranged in layers, including input, hidden, and output layers. The system operates in a coordinated manner to solve problems and learns from information provided by comparing its classification with known information classifications. A significant advantage of the ANN is that it does not require any assumptions or prior knowledge for problem-solving (Zheng et al., 2014). Recently, ANNs have gained popularity in the field of crash severity data analysis. This type of analysis involves treating crash severity modeling as a pattern recognition task, in which the ANN classifier assigns each input to a severity category. In previous studies, fully connected feed-forward neural networks were employed to predict the severity of vehicle crashes (Abdelwahab and Abdel-Aty, 2001;

Sohn and Shin, 2001). Subsequently, ANNs were used to assess the effects of electronic toll collection (ETC) systems at toll plazas on highway safety (Abdelwahab and Abdel-Aty, 2002). To improve the accuracy of individual ANN classifiers, a data fusion method was introduced to combine data from various sources (Sohn and Lee, 2003). More recently, researchers divided crash severity datasets into subsets using k-means clustering and applied ANNs to model them separately, leading to significant improvements in prediction performance when compared with ANNs trained on the entire dataset (Alkheder et al., 2017).

Data Mining

Data mining, which is a process of extracting knowledge from large datasets, has emerged as a powerful tool for analyzing and understanding complex road safety datasets. There are several advantages of data mining techniques in traffic safety analysis. Firstly, data mining can identify patterns and relationships between variables that are not easily detected using traditional statistical methods. In addition, data mining can handle large and complex datasets with a high level of accuracy and efficiency. This is particularly useful in traffic safety analysis, where datasets can contain hundreds of variables and thousands of records. Unlike most traditional statistical models, data mining usually does not have any predefined assumptions.

Data mining techniques, such as association rule mining, can identify contributing factors behind safety-critical events, including crashes, by investigating unclear and complex relationships among variables in big data. This may not be possible using traditional safety analysis methods that rely on traditional modeling. Association rule mining is a rules-based data mining method for investigating interesting associations of variables in large databases. The study conducted by Khan et al. provides an in-depth analysis of driver speed selection behavior in adverse weather utilizing telematics data coupled with association mining, with the aim of developing a human-in-the-loop variable speed limit (VSL) algorithm (Khan et al., 2020). Another study utilized this robust technique to discover vehicle-pedestrian crash patterns to help safety professionals understand significant patterns and relevant countermeasures to raise awareness and improve the potential reduction of pedestrian crashes (Das et al., 2019). In recent years, many researchers have been using clustering-based data mining methods including k-means clustering, k-medoids clustering, hierarchical clustering, density-based clustering, fuzzy clustering, and spectral clustering. One recent study applied several clustering methods to categorize drivers into aggressive, normal, and calm clusters, and subsequently proposed a driving score to assess driving performance (Mohammadnazar et al., 2021).

The use of text mining has also been leveraged in many recent traffic safety-related studies to get a deep understanding of the contributing factors leading to crashes and associated driver behaviors. With the rise of the internet and digital technology, publication media has become a valuable source of information that can be used to enhance collective knowledge. Das et al. conducted a study that demonstrated the potential of text mining in gathering information related to traffic safety. They utilized the knowledge discovery in text (KDT) approach to analyze over

15,000 research papers, revealing research trends and histories of development (Das et al., 2016). In another study, the same researchers utilized latent Dirichlet allocation (LDA) in conjunction with structural topic modeling (STM) to understand topical trends in the complex and evolving field of transportation engineering research, particularly within traffic safety (S. Das et al., 2017). A. Rakotonirainy et al. (2015) conducted a study that employed text mining to identify various factors that contribute to crashes, such as the use of phones and oversteering, particularly on curved sections.

Deep Learning

In recent times, the utilization of deep learning models has witnessed a growing trend in traffic safety analysis. Among these models, Convolutional Neural Network (CNN) has emerged as a popular option for developing prediction models due to its exceptional performance compared to alternative methods. In a recent investigation carried out by Zheng et al. (2019), they introduced a new model called Traffic Accident's Severity Prediction-Convolutional Neural Network (TASP-CNN) for predicting crash severity. The study compared the performance of TASP-CNN against various statistical and machine learning models, ultimately demonstrating the superiority of the TASP-CNN model. Similarly, another study proposed a deep learning framework known as 'DeepScooter' to predict the severity of motorcycle-involved crashes, achieving an impressive accuracy of 100% and 94% for the training and testing datasets, respectively (Das et al., 2018). A. Furthermore, Rahim and Hassan (2021) utilized four years of crash data from Louisiana and employed DeepInsight in conjunction with a pre-trained CNN called EfficientNet to establish a crash severity prediction framework. The study highlighted that the proposed framework exhibited significantly better performance compared to traditional machine learning models.

Other Machine Learning Models

Other machine learning and data mining-based models that have been used to conduct safety analysis include Naive Bayes, gradient boosting, Light GBM, recurrent neural network, AdaBoost, genetic algorithm, fuzzy adaptive neural network, conditional inference forest, pre-trained CNNs, and TabNet.

2.3 LAND AREA CONTEXT

2.3.1 Expanded Functional Categorization System

To evaluate the impact of land use on traffic safety, it's crucial to consider the latest land use categories presented in the NCHRP Report 855 (Stamatiadis et al., 2018). The report introduced an Expanded Functional Categorization System (FCS) comprising five distinct context categories, each requiring different geometric design practices in terms of operating speeds, accessibility, and user groups: Rural, Rural Town, Suburban, Urban, and Urban Core.

- Rural: Regions characterized by minimal population concentration, limited buildings or constructions (scattered or absence of residential, commercial, and industrial establishments), and typically significant distances between structures.

- Rural Town: Regions with low population density yet diverse land utilization, featuring a commercial main street ambiance, potential availability of on-street parking and sidewalks, and relatively small distances between structures.
- Suburban: Areas displaying moderate population density, a blend of land uses within and between buildings (including mixed-use town centers, commercial strips, and residential zones), and varying distances between structures.
- Urban: Localities marked by high population density, a mix of land uses and prominent destinations, potential provision for on-street parking and sidewalks, and a combination of different distances between structures.
- Urban Core: Districts showcasing the highest population density, diverse land uses within and among predominantly tall structures, and relatively minimal distances between structures.

The Expanded FCS's main goal is to give designers better information so that they can make better design decisions by considering the mobility and safety of all roadway users. Figure 1 shows the typical user priorities in the expanded FCS (freeways are not shown).

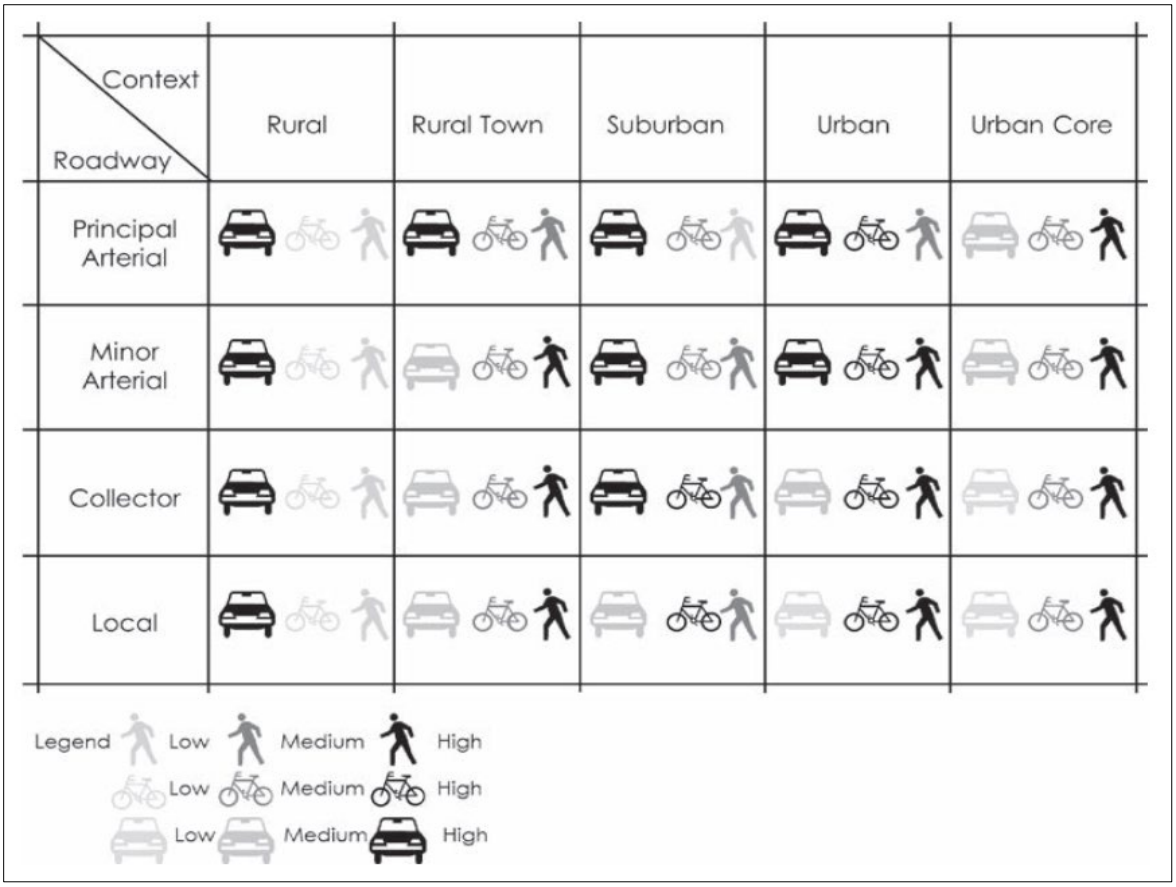


Figure 1. Typical User Priorities in the Expanded FCS (Stamatiadis et al., 2018).

2.3.2 Studies on Land Use and Traffic Safety

Xie et al. (2019) presented an innovative approach to evaluate the impact of land use conversions on traffic safety outcomes. The study specifically focused on examining the association between land use conversions and changes in severe crash occurrences. The results demonstrated that urban areas characterized by residential, commercial, and mixed residential-commercial land uses exhibited the highest levels of exposure to severe crash risk. In a similar vein, Mathew et al. (2022) conducted a study that explored the influence of road network characteristics, demographic variables, and land use characteristics on the occurrence of collisions involving minors. Through their research, they identified several significant explanatory factors that contribute to the incidence of teen crashes. These factors include AADT, the presence of light commercial land use, the number of occupied housing units in proximity, and the number of students enrolled in public or private high schools.

The association between several built environment elements on and around the university campus and pedestrian crashes was investigated by Dai et al. (2010). The study used network-based geospatial methods to pinpoint the locations of crash clusters in the study region. Data on pedestrian crashes from 2003 to 2007 were gathered from the Georgia Department of Transportation, and each road segment and intersection's environmental aspects of the streetscape, infrastructure, and pedestrians were assessed. According to the study, there were pedestrian crashes on more than 50% of roadways that were wider (more than 29 feet), two-way, and in good shape. Crashes involving pedestrians happened more frequently on road segments with mixed land uses and substantial street compactness, and they were notably concentrated in high-density areas. Mukoko and Pulugurtha (2020) investigate the impact of network, land use, and demographic factors on the estimation of bicycle-vehicle crashes on urban roads. The research findings indicate that network characteristics exhibit comparable or superior predictive capabilities compared to land use and demographic factors. Specifically, the study reveals that bicyclists have a higher probability of being involved in crashes on road segments lacking dedicated bicycle lanes, those with traffic lights, and where the speed limit is set at 45 mph. Furthermore, increased crash likelihood is observed on road segments situated in close proximity to commercial, research, institutional, multi-family residential, and heavy industrial areas. Conversely, the presence of single-family residential areas appears to have a relatively lower impact on the occurrence of bicycle-vehicle crashes.

In another study, Ding et al. (2020) examined the influence of infrastructure and land use on bicycle crash exposure and frequency using bike-sharing data. They identified factors such as higher road density, business areas, a larger percentage of older individuals, male and white populations, and higher median family income as positively associated with bicycle crash incidence. Pulugurtha et al. (2013) developed crash estimation models at the traffic analysis zone (TAZ) level, indicating the significance of land use factors such as mixed-use development, urban residential, single-family residential, multi-family residential, business, and office districts

in predicting crashes. The models also revealed a negative coefficient for single-family residential areas, suggesting a decrease in crash frequency with an increase in the extent of such development. Koloushani et al. (2022) conducted a study focusing on the spatial correlations between land use and pedestrian injury severity in non-intersection crashes in Northwest Florida. The findings of this research indicate that certain types of land use play a significant role in predicting the severity of pedestrian-involved crashes. Specifically, the likelihood of a severe pedestrian-involved crash increases in areas with commercial land use, such as retail stores or nightclubs, while it decreases near university campuses. Furthermore, regardless of the surrounding land use type, daylight conditions contribute to the severity of pedestrian-involved crashes, while average traffic volume is a statistically significant factor for crashes occurring in proximity to parking lots and office buildings. Pedestrian-involved crashes predominantly occur during daylight hours near office buildings, and pedestrians in these areas are less likely to sustain severe injuries.

2.4 PEDESTRIAN AND BIKE SAFETY

Each year, pedestrian and cyclist fatalities account for approximately 19% of all traffic-related deaths, resulting in approximately 6,000 pedestrian deaths and 850 cyclist deaths, along with numerous injuries (FHWA, 2022). While the overall number of traffic fatalities has decreased, the proportion of pedestrian and cyclist fatalities within the total has been increasing. However, the actual magnitude of the safety issue is likely greater than what is reflected in published statistics due to limited data on changes in pedestrian and cyclist travel patterns and exposure compared to other modes of transportation.

To tackle these safety concerns, jurisdictions are adopting more advanced approaches to assess the safety performance of transportation infrastructure. This often involves monitoring shifts in pedestrian and cyclist crash frequencies and exposure. Such monitoring helps identify potential risks and the most effective strategies for creating safer environments for walking and biking. For instance, if both crash occurrences and exposure decrease, it may indicate that people are avoiding walking or biking due to safety apprehensions. Conversely, if exposure increases while crashes decrease, it suggests that initiatives like Toward Zero Deaths or Vision Zero have been successful. Encouraging active modes of transportation while simultaneously reducing crash incidents necessitates collaborative efforts by state and local agencies. However, the limited availability of data restricts our understanding of the severity of pedestrian and cyclist crash risks and the locations where issues are prevalent. For instance, two intersections may report the same number of pedestrian crashes in a given year, but one intersection might have significantly higher pedestrian volumes, making it comparatively safer (Kristen Brookshire et al., 2016).

2.4.1 Common Factors Contributing to Pedestrian and Bicycle-Related Crashes

Although crash frequencies and rates can provide information on the severity and changes over time, they do not reveal the root causes of crashes or how to address them. To effectively combat

pedestrian and bicyclist crashes and injuries, it is crucial to fully comprehend the factors that contribute to these crashes. This will enable the development and application of the most cost-efficient and appropriate combination of countermeasures. Common issues that increase the risk of collisions and severe injury include (Kristen Brookshire et al., 2016):

- **Speeding:** Bicyclists and pedestrians are more likely to be struck by vehicles that go beyond the posted speed limit or drive too quickly for the conditions of the road. According to studies, the chance of a pedestrian dying increases as the impact speed of the motor vehicle rises. According to one study, the likelihood of a pedestrian dying in a collision with a moving car increases from 8% at 50 km/h (about 31 mph) to 50% at 75 km/h (about 47 mph) (Rosén and Sander, 2009).
- **Conflicts at crossing locations:** The danger of a collision may increase when the layout of an intersection or other crossing place causes conflicts between various road users. Drivers are less likely to expect pedestrians and bicycles to cross at places that are not intended for crossing, which increases the likelihood of crashes.
- **Inadequate conspicuity:** Drivers may not detect pedestrians and bicyclists if they are difficult to spot, especially in low-light situations, which increases the likelihood of a collision. Additionally, a lot of bikers are either unaware of or disobey the rules requiring them to wear reflectors and/or lights when riding at night.
- **Poor compliance with traffic laws and proper use of facilities:** The safety of all road users is impacted by motorists, pedestrians, and bicyclists who break the law. For instance, pedestrians and bicyclists who use the incorrect side of the road or disregard traffic signs and signals, as well as automobiles that refuse to yield to pedestrians at crosswalks. Sometimes, poor compliance can be attributed to missing or inadequately constructed facilities and crossings, as well as to a lack of awareness of the intended use of certain designs or traffic control systems.
- **Inadequate separation:** Bicyclists and pedestrians who are not segregated from fast-moving, heavy traffic are at risk. Drivers could not see pedestrians in time to prevent a collision if they were forced to cross the road while walking. Dense traffic, heavy vehicle traffic, darkness, and a small field of vision may also have an impact on bicyclists who share motor vehicle lanes in similar circumstances. The danger of collisions at junctions and driveways, as well as possible confrontations with pedestrians, might rise if riders choose to ride on sidewalks in an effort to feel safer.

A recent study has identified several risk factors associated with pedestrian crashes at both intersections and roadway segments. These factors include traffic volume, functional classes, pedestrian volume, presence of signals, and vehicle speed. The report discusses the potential roadway risk factors and their relationship to pedestrian crashes. The study found that most of the factors have a generally positive but not linear relationship to pedestrian crashes, except for high-turning volumes and the proportion of local streets at an intersection, which are still

unknown (see Table 14). The presence of a two-way left-turn lane was found to be positively related to pedestrian crashes on roadway segments, while vehicle speed and speed limits were found to be positively related to crash severity (Thomas et al., 2018).

Table 14. Potential Roadway Risk Factors and Relationship to Pedestrian Crashes Reproduced from Thomas et al. (2018).

Variable/Risk Factors	Intersections	Segments
Traffic volume	Positive (generally positive but not linear)	Positive (generally positive but not linear)
High-turning volumes	Unknown threshold	Unknown at present
Functional classes—arterials and collectors compared with local streets	Positive	Positive
Proportion of truck/bus traffic in traffic stream	Positive (crash severity)	Positive (crash severity)
Proportion of local streets at intersection	Negative	Unknown at present
Pedestrian volume	Positive (but not linear)	Positive (but not linear)
Number of legs > 3	Positive	Unknown at present
Total lanes on largest leg (5+)	Positive	Unknown at present
No median/median island	Positive (less certain than for segments)	Positive
Presence/number of transit stops	Positive	Positive
Presence of on-street parking	Positive	Positive
Presence/number of driveways	Positive	Unknown (theoretically yes)
Presence of signal	Positive with crash frequencies Negative with crash severity	Unknown at present
Lack of separate turning movements from walk phase (all red walk phase, or walk and restricted turn phase)	Positive	Unknown at present
Lack of leading pedestrian interval	Positive	Negative
Presence of four or more through lanes Higher numbers of total lanes	Theoretically yes	Positive
Presence of TWLTL	Unknown at present	Positive
Speed limit > 25 mph	Unknown at present	Positive with crash severity; positive with frequency in a few studies
Vehicle speed	Positive with severity	Positive with severity

2.4.2 Addressing Pedestrian and Bicycle Safety Concerns

The '3E' approach is frequently used by transportation professionals to increase the safety of cyclists and pedestrians. The three main 'Es' in this concept are engineering, education, and enforcement. Engineering describes the addition of sidewalks, bike lanes, or traffic lights, among other modifications to the environment or operations of the highway that impact how people on foot, on bicycles, and in other vehicle types navigate it. In order to encourage positive changes in attitudes or actions among drivers, pedestrians, cyclists, and other groups, education entails raising knowledge of certain safety regulations, concerns, or practices. Compliance with safety-related rules and regulations, such as speed limits or the use of crosswalks, is the main goal of enforcement activities.

There are more crucial aspects to consider in addition to these three mains 'Es.' These could include Emergency Response, Emerging Technologies (such applications that warn of vehicle-pedestrian collisions), and Encouragement/Engagement initiatives that encourage people to walk and bike and include them in safety-related discussions. Some professionals look at the 'Es' from the perspectives of evaluation and equity. While equity aims to equally distribute spending in safety measures across a community, evaluation focuses on assessing the impact of safety treatments or programs. According to research, combining the 'Es' (such as adopting engineering modifications with education and enforcement initiatives) is typically more effective than depending on a single strategy. The National Highway Traffic Safety Administration (NHTSA) recently conducted a comprehensive review of existing treatments and programs related to key pedestrian and bicycle safety concerns, as well as the related treatments and programs.
Table 15 summarizes the key pedestrian and bicycle safety concerns and the treatments and programs associated with them (Kristen Brookshire et al., 2016).

Table 15. Summary of Key Pedestrian and Bicycle Safety Concerns and Related Treatments and Programs.

Treatment or Program	Primary Mode Affected	Relation to Safety Concerns				
		Excessive vehicle speed	Conflicts at crossing locations	Inadequate conspicuity/ visibility	Poor compliance with laws/ proper facility use	Insufficient separation from traffic
Engineering Treatments						
1. Access management	All	•	•	maybe		
2. Advance yield/stop lines	Pedestrians/Drivers		•	•	•	
3. Bicycle detection at signals	Bicyclists		•		•	
4. Bike lanes	Bicyclists				•	•

Treatment or Program	Primary Mode Affected	Relation to Safety Concerns				
		Excessive vehicle speed	Conflicts at crossing locations	Inadequate conspicuity/ visibility	Poor compliance with laws/ proper facility use	Insufficient separation from traffic
5. Bicycle pavement marking improvements	Bicyclists		•	•	•	•
6. Bicycle-tolerable rumble strips	Bicyclists/Drivers					•
7. Crossing islands and raised medians	Pedestrians/Drivers	•	•	•		•
8. Interchange design	Pedestrians/ Bicyclists	•	•			
9. Intersection geometric design	Pedestrians/Drivers	•	•	•		
10. Lighting and illumination	All		•	•		
11. Marked crosswalks	Pedestrians		•	•	•	
12. Pedestrian and bicycle overpasses/underpasses	Pedestrians/ Bicyclists		•			•
13. PHBs	Pedestrians/ Bicyclists)		•		•	
14. Pedestrian signals and push buttons	Pedestrians		•		•	
15. Rectangular rapid flashing beacons	Pedestrians		•	•	•	
16. Road diets	All	•	•	•		•
17. Roundabouts	All	•	•		•	
18. Separated bike lanes	Bicyclists	•			•	•
19. Sidewalk buffers and landscaping	Pedestrians	•	maybe	•	maybe	•
20. Sidewalks and curb ramps	Pedestrians				•	•
21. Traffic signal phasing	All		•	•	•	•
22. Traffic calming and management	All	•				
Education and Awareness Programs						
23. Child training and skills practice	Pedestrians/ Bicyclists		•	•	•	
24. General pedestrian/bicycle safety communication and outreach	All	maybe	maybe	•	•	
25. Safe routes to school	Pedestrians/ Bicyclists		maybe	maybe	•	
Enforcement Programs						

Treatment or Program	Primary Mode Affected	Relation to Safety Concerns				
		Excessive vehicle speed	Conflicts at crossing locations	Inadequate conspicuity/ visibility	Poor compliance with laws/ proper facility use	Insufficient separation from traffic
26. Adult school crossing guards	Pedestrians/ Bicyclists	•	•	•	•	
27. Automated enforcement	All	•	maybe		•	
28. Speed display devices	All	•			•	
29. Targeted law enforcement	All	•	•	•	•	

2.4.3 Pedestrian and Bicyclist Safety Variables

Analyzing crash data is a common method to evaluate the safety of pedestrians and bicyclists. These crashes typically involve motor vehicles, but it's important to acknowledge that not all crashes are reported, and some do not involve motor vehicles at all. Undocumented crashes can provide valuable insights into areas requiring safety enhancements. One strategy for identifying such areas is to pinpoint 'hot spots' where pedestrian or bicycle crashes are concentrated. These hot spots might encompass intersections or specific road segments that necessitate design improvements. Moreover, it's crucial to consider that crash likelihood is higher in locations with greater pedestrian and bicycle activity. To accurately assess the risk of pedestrian or bicycle crashes, exposure measures must be incorporated. For instance, the crash rate can be determined by dividing the total number of crashes during a specific period by a corresponding measure of exposure. Exposure measures encompass various factors such as pedestrian crossing volume, motor vehicle volume at crosswalks, total bicycle volume at intersections, or population density in census tracts. Additionally, variables categorized under 'Demand' can serve as proxies for pedestrian or bicycle exposure. Even in cases where specific data on pedestrian and bicycle volumes may not be available, exposure measures can still be estimated based on the time spent in areas where pedestrians and bicyclists are likely to encounter motor vehicles. A recent NCHRP Report outlines several variables related to pedestrian and bicycle safety, as presented in Table 16 (Lagerwey et al., 2015).

Table 16. Safety Variables, Reproduced from Lagerwey et al., (2015).

Example Variables	Relevance		Potential Location
	Ped	Bike	
Total number of pedestrian/bicycle crashes	●	●	S, Cr, Co, A
Fatal and severe injury pedestrian/bicycle crashes	●	●	S, Cr, Co, A
Pedestrian/bicycle crash rate	●	●	S, Cr, Co, A

Example Variables	Relevance		Potential Location
	Ped	Bike	
Proportion of pedestrians walking in the roadway	◐	◯	S
Proportion of pedestrians complying with "Don't Walk" signals	◐	◯	Cr
Proportion of bicyclists complying with red lights	◯	◐	Cr
Proportion of motorists complying with right turn on red restrictions	◐	◐	Cr
Proportion of motorists yielding to pedestrians in crosswalks	◐	◯	Cr
Number of "near misses" involving pedestrians/bicyclists	◐	◐	S, Cr, Co, A

Notes: ● = Very relevant, ◐ = Less relevant, ◯ = Not likely relevant, S = Segment, Cr = Crossing, Co = Corridor, A = Area

These variables, which are listed in Table 17, include traffic volume, pedestrian volume, transit presence, total through lanes, median type, crosswalks, on-street parking, and various types of markings and signs. Additionally, the table provides measurements for each variable, such as AADT for traffic volume and the presence or count of certain facilities like pedestrian hybrid beacon or rectangular rapid flashing beacon along a segment (Thomas et al., 2018).

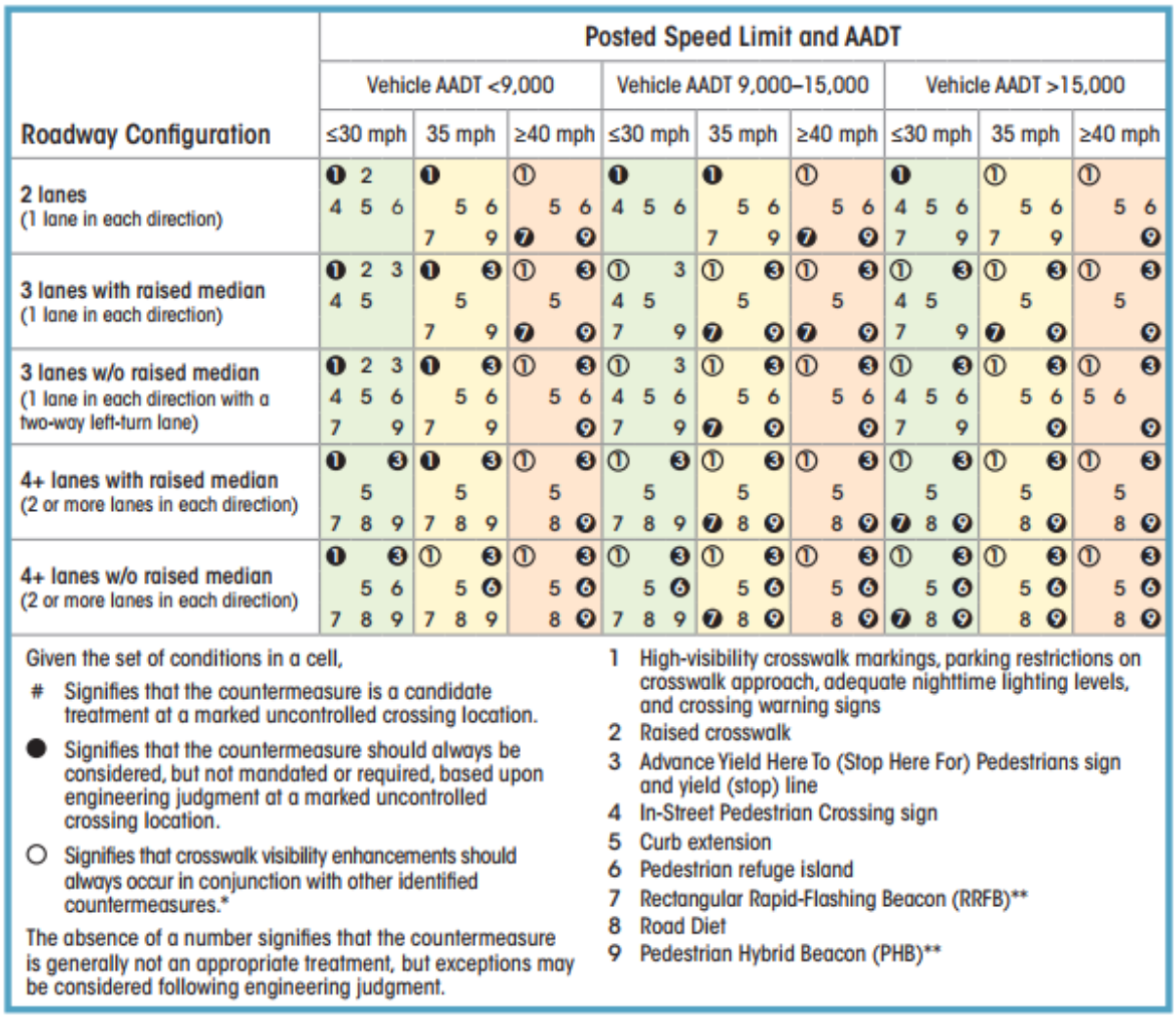
Table 17. Potential Pedestrian Crash Risk Variables for Segment Analysis, Reproduced from Thomas et al., (2018).

Segment-Related Roadway Variables	Measurements
Traffic volume	Typically, ADT or AADT are available for state road networks. Subtypes may include. • Major and minor road volumes (for intersections) • Volume assignment by functional class (surrogate measure) • Heavy vehicle percentages
Pedestrian volume	It is challenging to account for pedestrian volumes crossing a length or segment of roadway. Ideally, the number of pedestrians walking along the roadway and of pedestrians crossing anywhere along a segment could be included. It may be feasible to collect counts of pedestrians crossing at non-intersection-marked crosswalk locations.
Transit	• Presence of stops within X distance of segment midpoint or • endpoints • Number of stops along segment
Total through lanes	Number of through lanes (average, either end of segment; midpoint□number of through lanes; or number proportionally weighted)
Total through lanes	Number of through lanes (average, either end of segment; midpoint number of through lanes; or number proportionally weighted)
Median with/without crossing facilities	Presence of a continuous raised (not painted or TWLTL) median

Segment-Related Roadway Variables	Measurements
Median islands with pedestrian crossing	Count of raised median islands with pedestrian pass-through refuge along segments. Could consider median island presence at the intersection.
Two-way left-turn lane	Presence of TWLTL
Midblock crosswalks	Presence or count of marked crosswalks with unsignalized approaches along a segment
On-street parking	Presence (any, one, or both sides) or proportion of segment covered by striped parking
Pedestrian hybrid beacon or PHB	Presence or count of the facility type along a segment
Rectangular rapid flashing beacon	Presence or count of the facility type along a segment
High visibility crosswalk markings	Presence or count of the facility type along a segment
Advance stop/yield markings and signs	Presence or count of the facility type along a segment
Speed limit	Posted speed limit or weighted average speed limit along segment
Segment length	Length of segment; may be estimated from spatial data
Sidewalk coverage	Presence of sidewalks along zero, one, or both sides, or proportional coverage from front frontage data
Distance to nearest signalized crossing or activated beacon along same road	As described
Right- or left-turn lanes at adjacent intersections	Presence or counts of different lane types at adjacent intersection

2.4.4 Selection of Safety Countermeasures

As part of the Every Day Counts (EDC-5) program focused on pedestrian safety, the FHWA has recently revised its guide for improving safety at uncontrolled crossing locations (FHWA, 2021). The guide presents various measures to address pedestrian safety concerns, taking into account the characteristics of the roadway. Aligned with the objectives of the EDC-5 program, several recommended measures include Road Diets, Pedestrian hybrid beacons (PHBs), Pedestrian refuge islands, raised crosswalks, Crosswalk visibility enhancements, Rectangular Rapid Flashing Beacons (RRFBs), and Leading Pedestrian Intervals. A comprehensive matrix and list of these countermeasures, categorized according to roadway and traffic features, can be found in Figure 2 (Blackburn et al., 2018). The assignment of specific countermeasures to matrix cells is based on safety research, best practices, and established national guidelines. When implementing pedestrian crossings, FHWA advises agencies to carefully assess the available countermeasure options and select the most suitable combination of treatments, taking into consideration factors such as pedestrian volume, operational speeds, land use context, and site-specific features.



Roadway Configuration	Posted Speed Limit and AADT								
	Vehicle AADT <9,000			Vehicle AADT 9,000–15,000			Vehicle AADT >15,000		
	≤30 mph	35 mph	≥40 mph	≤30 mph	35 mph	≥40 mph	≤30 mph	35 mph	≥40 mph
2 lanes (1 lane in each direction)	● 2 4 5 6	● 5 6 7 9	○ 5 6 ● 9	● 4 5 6	● 5 6 7 9	○ 5 6 ● 9	● 4 5 6 7 9	○ 5 6 7 9	○ 5 6 ● 9
3 lanes with raised median (1 lane in each direction)	● 2 3 4 5	● ◉ 5 7 9	○ ◉ 5 ● ◉	○ 3 4 5 7 9	○ ◉ 5 ● ◉	○ ◉ 5 ● ◉	○ ◉ 4 5 7 9	○ ◉ 5 ● ◉	○ ◉ 5 ● ◉
3 lanes w/o raised median (1 lane in each direction with a two-way left-turn lane)	● 2 3 4 5 6 7 9	● ◉ 5 6 7 9	○ ◉ 5 6 ◉	○ 3 4 5 6 7 9	○ ◉ 5 6 ● ◉	○ ◉ 5 6 ◉	○ ◉ 4 5 6 7 9	○ ◉ 5 6 ◉	○ ◉ 5 6 ◉
4+ lanes with raised median (2 or more lanes in each direction)	● ◉ 5 7 8 9	● ◉ 5 7 8 9	○ ◉ 5 8 ◉	○ ◉ 5 7 8 9	○ ◉ 5 ● 8 ◉	○ ◉ 5 8 ◉	○ ◉ 5 ● 8 ◉	○ ◉ 5 8 ◉	○ ◉ 5 8 ◉
4+ lanes w/o raised median (2 or more lanes in each direction)	● ◉ 5 6 7 8 9	○ ◉ 5 ◉ 7 8 9	○ ◉ 5 ◉ 8 ◉	○ ◉ 5 ◉ 7 8 9	○ ◉ 5 ◉ ● 8 ◉	○ ◉ 5 ◉ 8 ◉	○ ◉ 5 ◉ ● 8 ◉	○ ◉ 5 ◉ 8 ◉	○ ◉ 5 ◉ 8 ◉

Given the set of conditions in a cell,

- # Signifies that the countermeasure is a candidate treatment at a marked uncontrolled crossing location.
- ● Signifies that the countermeasure should always be considered, but not mandated or required, based upon engineering judgment at a marked uncontrolled crossing location.
- ○ Signifies that crosswalk visibility enhancements should always occur in conjunction with other identified countermeasures.*

The absence of a number signifies that the countermeasure is generally not an appropriate treatment, but exceptions may be considered following engineering judgment.

1 High-visibility crosswalk markings, parking restrictions on crosswalk approach, adequate nighttime lighting levels, and crossing warning signs
2 Raised crosswalk
3 Advance Yield Here To (Stop Here For) Pedestrians sign and yield (stop) line
4 In-Street Pedestrian Crossing sign
5 Curb extension
6 Pedestrian refuge island
7 Rectangular Rapid-Flashing Beacon (RRFB)**
8 Road Diet
9 Pedestrian Hybrid Beacon (PHB)**

Figure 2. Pedestrian Crash Countermeasures by Roadway Feature (Blackburn et al., 2018).

2.4.5 Development of SPFs and CMFs for Pedestrians and Bicyclists

The NCHRP Project 17-56 undertook the task of creating several SPFs and CMFs specifically for various pedestrian treatments at unsignalized pedestrian crossings (Zegeer et al., 2017). After careful consideration of multiple options concerning the design of road features and traffic control devices, this particular study chose to assess the effectiveness of four distinct treatment types. These encompassed RRFBs, PHBs, pedestrian refuge islands, and advanced markings and signs indicating YIELD or STOP. The analysis of data involved the formulation of cross-sectional models and the utilization of EB analysis techniques, both before and after the implementation of treatments, to ascertain the impact on crashes, represented by CMFs. In the case of refuge islands, specific equations were fine-tuned to provide projections of the anticipated number of crashes. The parameter estimates for all models pertaining to refuge islands can be found in Table 18.

$$PEDCRASH = exp^{(a+City+b*Refuge\ Island\ Presence+c*AreaType)} AAADT^e PEDAADT^f \quad (25)$$
$$TOTCRASH = exp^{(a+City+b*Refuge\ Island\ Presence+d*Midblock\_int)} AAADT^e PEDAADT^f \quad (26)$$
$$INJCRASH = exp^{(a+City+b*Refuge\ Island\ Presence+d*Midblock\_int)} AAADT^e PEDAADT^f \quad (27)$$
$$TARGETCRASH = exp^{(a+City+b*Refuge\ Island\ Presence+d*Midblock\_int)} AAADT^e PEDAADT^f \quad (28)$$
$$INJTARGETCRASH = exp^{(a+City+b*Refuge\ Island\ Presence+d*Midblock\_int)} AAADT^e PEDAADT^f \quad (29)$$

Where,

$AADT =$ total AADT on the roadway being crossed

$AreaType =$ 1 if Suburban, 0 if Urban

$City =$ an intercept term specific for each city

$Midblock\_Int =$ 1 if intersection, 0 if midblock

$PEDAADT =$ total pedestrian AADT for midblock or intersection

$Refuge\ Island\ Presence =$ 1 if present, 0 if not present

Table 18. Parameter Estimates for Refuge Island Regression Models, Reproduced from Zegeer et al. (2017).

Parameter	Parameter Estimate (Standard Error)				
	Pedestrian	Total	Injury	RE+SS	Injury RE+SS
a	-10.4246 (-1.6409)	-5.5953 (-0.7761)	-6.4572 (-0.8886)	-8.3157 (-0.9754)	9.713 (-1.2009)
b	-0.3578 (-0.2153)	-0.2981 (-0.0956)	-0.3369 (-0.1148)	-0.2999 (-0.1258)	-0.3254 (-0.146)
c	-0.5715 (-0.3127)	n/a	n/a	n/a	n/a
d	n/a	0.473 (-0.1083)	0.4312 (-0.1183)	0.414 (-0.1224)	0.3728 (-0.1395)
e	0.6977 (-0.1694)	0.5192 (-0.0756)	0.5375 (-0.0846)	0.7235 (-0.0947)	0.778 (-0.1157)
f	0.3295 (-0.0486)	0.1224 (-0.0247)	0.1141 (-0.026)	0.1041 (-0.0237)	0.0986 (-0.0263)
overdispersion	n/a	0.7608 (-0.0305)	0.7075 (-0.0482)	0.9149 (-0.0521)	0.9837 (-0.1075)

Likewise, advanced YIELD or STOP markings and signs the following equations were generated to predict the expected number of crashes per year. The parameter estimates for all models related to advanced YIELD or STOP markings and signs are shown in Table 19.

$$PEDCRASH = exp^{(a+City+b*Advance\ StopYield\ Sign\ Presence+c*AreaType)} AAADT^e PEDAADT^f \quad (30)$$
$$TOTCRASH = exp^{(a+City+b*Advance\ StopYield\ Sign\ Presence+c*AreaType+d*Midblock\_int)} AAADT^e PEDAADT^f \quad (31)$$
$$INJCRASH = exp^{(a+City+b*Advance\ StopYield\ Sign\ Presence+d*Midblock\_int)} AAADT^e PEDAADT^f \quad (32)$$
$$TARGETCRASH = exp^{(a+City+b*Advance\ StopYield\ Sign\ Presence+d*Midblock\_int)} AAADT^e PEDAADT^f \quad (33)$$

$$INJTARGETCRASH = exp^{(a+City+b*Advance\ StopYield\ Sign\ Presence+d*Midblock\_int)} AAADT^{e} PEDAADT^{f} \quad (34)$$

Where,

$AADT$ = total AADT on the roadway being crossed

$AreaType$ = 1 if Suburban, 0 if Urban

$City$ = an intercept term specific for each city

$Midblock\_Int$ = 1 if intersection, 0 if midblock

$PEDAADT$ = total pedestrian AADT for midblock or intersection

$Advance\ Stop\ Yield\ Sign\ Presence$ = 1 if present, 0 if not present

Table 19. Parameter Estimates for Advanced YIELD or STOP Markings and Signs Models, Reproduced from Zegeer et al. (2017).

Parameter	Parameter Estimate (standard error)				
	Pedestrian	Total	Injury	RE+SS	Injury RE+SS
a	-6.5485 (1.6715)	-5.1484 (0.6466)	-5.5571 (0.7427)	-7.8277 (0.7854)	-8.7847 (1.0780)
b	-0.1470 (0.3295)	-0.0195 (0.2240)	-0.1367 (0.2714)	0.3520 (0.3284)	0.2801 (0.4228)
c	-0.9656 (0.4798)	-0.2668 (0.1462)	n/a	n/a	n/a
d	n/a	0.6124 (0.1172)	0.6039 (0.1265)	0.5200 (0.1283)	0.4786 (0.1550)
e	0.2501 (0.2041)	0.5021 (0.0634)	0.4384 (0.0705)	0.6752 (0.0767)	0.6761 (0.1044)
f	0.4003 (0.1011)	0.0949 (0.0257)	0.1006 (0.0270)	0.0880 (0.0248)	0.1026 (0.0287)
overdispersion	n/a	0.7151 (0.0307)	0.6908 (0.0488)	0.9038 (0.0530)	1.1215 (0.1149)

The study also generated SPFs for the remaining two categories of treatments. Upon analyzing the data in relation to untreated sites, it was observed that all four treatment options were associated with a reduced likelihood of pedestrian crashes. The CMFs for pedestrian collisions were as follows: PHBs (CMF of 0.453), PHBs with advanced YIELD or STOP signs and markings (CMF of 0.432), pedestrian refuge islands (CMF of 0.685), and advanced YIELD or STOP signs and markings (CMF of 0.75). These treatments exhibited the most significant impact in mitigating the risk of pedestrian collisions. Notably, the CMFs for PHBs and PHBs with advanced YIELD or STOP markings and signs differed significantly from 1.0. Furthermore, CMFs for some of the four pedestrian treatments were identified for various other types of collisions, including rear-end, sideswipe, and complete crashes. A comprehensive compilation of the recommended CMFs generated for this study can be found in Table 20.

Table 20. Potential Pedestrian Crash Risk Variables for Segment Analysis, Reproduced from Zegeer et al. (2017).

Treatment	Crash Type	Recommended CMF		Study Basis
		Estimate	Standard Error	
Refuge Island	Pedestrian	0.685	0.183	Median from two studies
	Total	0.742	0.071	Cross-section
	All Injury	0.714	0.082	Cross-section
	Rear-End/Sideswipe Total	0.741	0.093	Cross-section
	Rear-End/Sideswipe Injury	0.722	0.106	Cross-section
Advanced YIELD or STOP Markings and Signs	Pedestrian	0.750	0.230	Median from two studies
	Total	0.886	0.065	Before-after
	Rear-End/Sideswipe Total	0.800	0.076	Before-after
PHB	Pedestrian	0.453	0.167	Median from two studies
PHB + Advanced YIELD or STOP Markings and Signs	Pedestrian	0.432	0.134	Median from two studies
	Total	0.820	0.078	Before-after
	Rear-End/Sideswipe Total	0.876	0.111	Before-after
RRFB	Pedestrian	0.526	0.377	Cross-section

2.4.6 Studies on Pedestrian and Bicyclist Safety

The safety and behavior of cyclists and pedestrians crossing at greenway-road crossings on an urban greenway in New Orleans were studied by Anderson et al. (2019). The researchers gathered information on crossing practices, safety, and driving practices via direct observation and intercept questionnaires. They analyzed the link between motor vehicle behavior and the activation of crossing signals (rectangular fast flash beacons) for walkers and bicycles using logistic and negative binomial regression. According to the study, turning on the crossing signals made it less likely for walkers and bicycles to cross the street dangerously, but there was no connection between pedestrian use of the lights and motorists' stopping habits. However, cyclists had considerably increased probabilities of seeing moving cars fail to stop when the signal was triggered.

In order to assess pedestrian facilities and pinpoint necessary upgrades along roadways, Asadi-Shekari et al. (2015) created the pedestrian safety index (PSI). By comparing the current circumstances to a standard, a point system technique was put out to estimate this PSI. The technique was used to pinpoint issues already present and provide solutions. Additionally, the improvements to pedestrian safety outlined using this technique enhance the security of elderly and disabled pedestrians, who are most adversely affected by a lack of amenities. Dumbaugh and Li (2011) aimed to ascertain whether the occurrence of urban crashes is attributable to random error or features of the constructed environment. They utilized vehicle miles of travel as an approximation for random error and discovered a slight positive correlation with crashes

involving both motorists and pedestrians. Conversely, they identified stronger connections between crashes and characteristics of the built environment. Major risk factors for crashes were identified as arterial roadway miles, four-leg intersections, strip commercial uses, and big box stores. In contrast, pedestrian-friendly retail uses were associated with a lower frequency of crash incidents. The findings indicate that enhancing urban traffic safety necessitates a balance between mitigating safety risks and addressing conflicts in traffic, rather than relying exclusively on forgiving roadway design.

Using supervised association mining, Das et al. (2019) found trends in a database of vehicle-pedestrian crashes. The study used crash data from Louisiana from 2004 to 2011 for eight years in order to look for these tendencies. The use of association rules mining was made to achieve this. The findings suggested that nighttime road lighting might lessen the severity of pedestrian collisions. The study also identified a number of groups of interest, including male pedestrians who are more vulnerable to serious and fatal crashes, younger female drivers who are more likely to be in collisions, impaired pedestrians who are still at risk even when there is nighttime lighting on the road, middle-aged male pedestrians who are more vulnerable to collisions, and single vehicle crashes as the most frequent.

The study conducted by Ferenchak and Marshall (2017) aimed to identify areas in urban regions that have a high concentration of child pedestrian fatalities. The researchers employed spatial and statistical analysis to compare fatal collision rates in areas surrounding schools and other places that children frequent, such as parks, trails, and recreation spaces. The authors utilized 30 years of crash data for six American cities and concentrated on collisions that occurred in Denver, Colorado, in the first phase of their study. Their research indicated that areas around parks and schools have higher child pedestrian fatality rates than areas without a school or park, and fatality rates around parks are higher than those around schools. The authors concluded that actions aimed at ensuring child pedestrian safety should concentrate on parks as well as schools. A separate study also explored the incidence of child pedestrian injuries in urban crashes and the characteristics that influenced injury severity (Koopmans et al., 2015). The study discovered that children had a higher overall incidence of injuries compared to adult pedestrians, but the case fatality rate was lower. It was also found that most crashes for both children and adults occurred during favorable driving conditions. Younger age groups experienced injuries more frequently during warmer months than older groups. Midblock crashes increased as age decreased. The majority of crashes took place at locations with inadequate traffic controls but varied by age. For younger age groups, crashes were more likely to happen during daylight, on dry roads, and under clear weather conditions than for older groups.

A recent investigation conducted by Goswamy et al. (2023) explored the effectiveness of the RRFB in mitigating the severity of crashes occurring at pedestrian crossings. The study encompassed a dataset of 312 locations where pedestrians crossed, out of which 154 locations

were equipped with the RRFB, while the remaining 158 control locations lacked any countermeasures specifically designed for pedestrian crossings. The control locations exhibited similar characteristics in terms of traffic, roadway, and land use compared to the treatment locations. The objective of the study was to assess the impact of the RRFB and other variables on the severity of various types of crashes, including those involving pedestrians, nighttime incidents, total crashes, and rear-end collisions. The findings indicated that the RRFB had a positive influence on reducing nighttime crashes, particularly those classified as K and A crashes. However, the study did not observe a significant improvement in reducing rear-end collisions and overall crashes within the study area.

Another research study conducted by Guo et al. (2016) examined the effects of parallelogram-shaped pavement markings on vehicle speed and pedestrian safety at urban crosswalks. The study employed observational cross-sectional methods to evaluate the impact of these markings on vehicle speed and the occurrence of crashes near pedestrian crosswalks. The results demonstrated that the utilization of parallelogram-shaped pavement markings resulted in a notable decrease in vehicle speeds and violations of speed limits in the vicinity of pedestrian crosswalks. Furthermore, the implementation of these markings also led to a reduction in both the frequency and severity of crashes at pedestrian crosswalks. Kraidi and Evdorides (2020) conducted a study that aimed to create models for evaluating pedestrian safety, considering the impact of pedestrian and roadside activities. The researchers discovered that various factors significantly contributed to pedestrian crash risk, including the frequency of bus stops, parking, pedestrian crossings, traffic speed fluctuations, the number of intersecting side roads, and through and intersecting traffic volume. Additionally, the volume of violations committed by pedestrians and drivers was also a significant risk factor.

For different kinds of urban highway segments and junction facilities, a study was done to estimate the bicycle CMFs (Raihan et al., 2019). According to the research, bicycle collisions are decreased by lane width, speed limits, and grass in the median. However, the incidence of bicycle collisions rose when there were sidewalks and sidewalk barriers present. The study also discovered that an increase in bicycle activity increased the likelihood of collisions at junctions but decreased the likelihood of collisions on segments of roadways. Bus stops were found to increase the possibility of cycling crashes at junctions, but protected signal control improved bicycle security. The study by Davidse et al. (2019) looked at possible outcomes of collisions involving light mopeds in urban bike lanes. The research made many recommendations for improving the safety of both bikers and light-moped riders based on the data. These included clearing obstructions like poles from the bike path, adhering to recommendations for the minimum width of bike lanes given traffic volumes, enhancing visibility at intersections, putting traffic light control in place without interfering with traffic flows, and enacting a helmet law for riders of light mopeds and their passengers.

2.5 RELATIONSHIP BETWEEN SPEED AND SAFETY

2.5.1 Speed Measures in the HSM

The SPFs for the first edition of the HSM (AASHTO, 2010) and the future second edition of the HSM does not explicitly contain speed metrics despite the fact that they must be included when evaluating highway safety. The first version of HSM's Appendix 3E, titled 'Speed and Safety,' offers some background about speed restrictions and how they affect overall safety. The HSM also includes CMFs for the average operating speed change (before and after the crash occurrence), although these measurements need to be reexamined in light of the availability of more recent data sources like the NPMRDS.

2.5.2 Speed Limit

Corridor-level strategies for implementing measures to ensure compliance with speed restrictions rely on the establishment of speed limits. To safeguard the well-being of all individuals utilizing the roadway, including motor vehicles, bicycles, and pedestrians; it is important to impose a reasonable speed restriction that promotes a secure, consistent, and practical flow of traffic. Speed control measures, such as speed limits, play a vital role as they provide drivers with explicit instructions to guide their selection of suitable speeds while driving. This is particularly significant because drivers may not always possess the ability to independently determine appropriate speeds (Elvik, 2010). Figure 3 presents a visual representation of various types of speed limits.

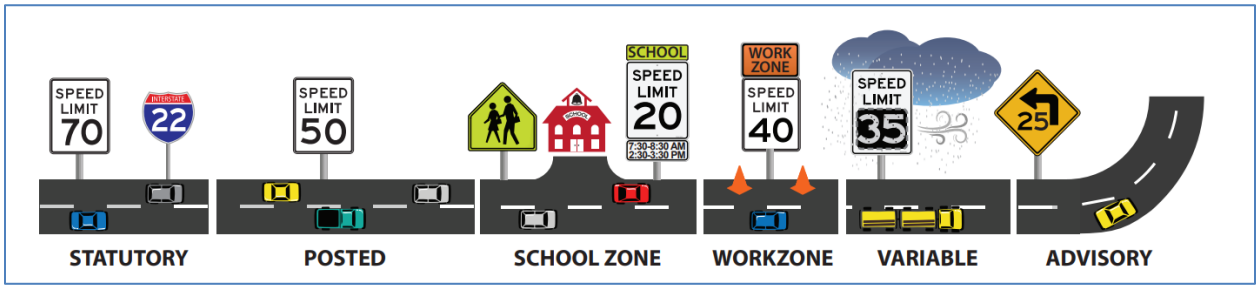


Figure 3. Types of Speed Limits (FHWA, 2016).

Statutory Speed Limit

State legislatures set statutory speed limits for each type of highway infrastructure, which might differ from state to state. A statutory speed restriction is legally binding and valid even if it is not posted, according to FHWA Speed restriction Basics. Examples include 55 mph on rural roadways, 70 mph on rural freeways, and 25 mph in residential or school areas (FHWA, 2016).

Posted Speed Limit

Posted speed limits, also known as regulatory speed limits, are those that are posted along the road and are made enforceable by law, according to FHWA Speed Limit Basics. A posted speed limit might be adjusted from the statutory speed limit by a municipal, county, or state transportation agency, or it could be the same as the legal speed determined by the State legislature (FHWA, 2016).

Advisory Speed Limit

Advisory speeds are intended to increase safety at road alignments like the sites of horizontal and vertical curves. These speed restrictions are typically established using an engineering speed study and in compliance with recommendations in the MUTCD (FHWA, 2016). The posted advisory speed is meant to notify motorists of any road conditions that call for a decrease in speed (Grabowski and Morrisey, 2007). According to Bonneson et al. (2007), the advised speed restriction needs to be established using the mean operating speed of truck drivers.

Variable Speed Limits

An essential component of an intelligent transportation system (ITS) is the VSL concept. This speed management system can modify the speed limit on various road segments based on the current flow of traffic or a predetermined speed control algorithm. Several American states, including Arizona, New Mexico, Oregon, and Washington, have already enacted VSL. According to studies, VSL practices have been found to lower average speeds (Garber and Srinivasan, 1998; Ullman and Rose, 2005). VSLs are seen to be an efficient countermeasure for avoiding speed-related collisions as well as aiding in the regulation of congestion, especially in work zones (Levin et al., 2019; Ullman and Rose, 2005). According to data from Pauw et al. (2018), the introduction of VSL resulted in a sharp decline in injury collisions (18% of overall crashes were reduced, and 6% of fatal and serious crashes). On rural, mountainous motorway routes for VSL, Saha et al. (2015) investigated how the road, weather, and collisions interacted. The research results indicated that the combination of meteorological conditions along with horizontal and vertical curves had a notable influence on the frequency of crashes.

Studies on Speed Limit and Safety

The role of reducing speed limits is crucial for ensuring road safety, as demonstrated by Kloeden's research conducted in Australia. In South Australia, the implementation of the Default Urban Speed Limit (DUSL) lowered the speed limit from 37 mph to 31 mph on all urban roads, unless specified otherwise, starting from March 1st, 2003. Kloeden et al. (2004) examined the effectiveness of this reduction by analyzing speed surveys and crash data. The results showed an average decrease of 1.4 mph in mean speeds on streets with reduced speed limits and 0.4 mph on arterial roads with 37 mph signs. Additionally, there was a significant reduction of 19.8% in casualty crashes on 37 mph roads and 4.6% on 37 mph arterial roads. In a subsequent evaluation

three years later, Kloeden et al. (2007) observed average speed reductions of 2.4 mph on reduced-speed streets and 1.3 mph on arterial roads. Moreover, casualty crashes decreased significantly by 23% on 31 mph roads and 16% on 37 mph arterial roads. Comparing data from 2016, Kloeden et al. (2017) noted that speeds on South Australian 31 mph roads remained mostly unchanged, although there was a potential increase in the number of vehicles exceeding 40 mph on local roads in Adelaide. On the other hand, Adelaide's 37 mph roads experienced a historical decline in vehicle speeds, while speeds on rural hill roads stabilized in 2016 after three years of decrease. Rural roads with speed limits of 62 mph and 68 mph remained stable since 2015, with fewer occurrences of high-speed instances observed over time.

Numerous studies have investigated the impact of speed limits on pedestrian safety. The introduction of metrication in 1974 resulted in an increase in the urban speed limit from 35 mph to 37 mph throughout Australia. Mclean and Anderson (2008) conducted a study to assess the effects of metrication and the subsequent reduction of the urban speed limit from 37 mph to 31 mph on pedestrian fatalities. They estimated the potential consequences of choosing a 37 mph speed limit instead of 31 mph since 1974. Heydari et al. (2014) proposed a methodology to analyze the influence of speed limit reduction on speeding behaviors, identifying various factors that either heightened or diminished speeding. While speed limit reductions proved effective for limits of 25 mph and 31 mph, they did not significantly reduce instances of excessive speeding, thereby posing risks to pedestrians and cyclists. Isaksson-Hellman and Töreki (2019) discovered that reducing the speed limit from 31-37 mph to 19-25 mph significantly decreased the likelihood of moderate-to-fatal injuries for cyclists involved in car collisions. Mitra et al. (2021) reported that lower speed limits in Korea reduced the likelihood of pedestrian fatalities, although they did not observe reductions in the overall number of pedestrian crashes and injuries.

Islam and El-Basyouny (2015) conducted a comprehensive assessment using Bayesian analysis to evaluate the safety impacts of reducing the posted speed limit in urban residential areas. They found that reducing the speed limit had a positive effect in reducing crashes of all severities. In Spain, a new law implemented in 2020 lowered the standard speed limit on two-lane roads from 31 mph to 19 mph to decrease crashes. Gonzalo-Orden et al. (2021) gathered evidence from other countries showing the positive effects of reducing vehicle speeds in urban areas. Son et al. (2022) conducted a study and found that lowering the speed limit effectively reduced the number of crashes, including serious injuries and fatalities. These findings suggest that lowering the speed limit reduces both the overall number and severity of crashes.

In their study, Kwayu et al. (2018) employed regression analysis to examine the consequences of increasing the speed limit on urban freeways in Michigan. They observed a rise in fatal, incapacitating, and total crashes, as well as incidents of vehicles veering off the road, following the speed limit increase. This impact was particularly notable on curved sections of the freeways. Cloutier and Lachapelle (2021) conducted an assessment of the effects of reducing speed limits

on collisions involving fatalities or severe injuries in Quebec, Canada. They identified a decreasing trend in incidents across various road segments, with a more substantial decline in segments characterized by higher initial speeds and larger speed reductions. It is crucial to thoroughly evaluate geometric characteristics when considering modifications to speed limits. Siddiqui et al. (2017) examined the impacts on the safety of a VSL system that provided advisory guidelines on the OR-217 freeway in Portland. Due to a lack of comprehensive crash data, they relied on surrogate safety measures to assess the system's effectiveness. The results indicated a decrease in the overall number of crashes, particularly rear-end collisions. The implementation of the VSL system led to reductions in both the average speed and the variability of speeds within and between lanes at specific locations. Additionally, it contributed to minimizing speed fluctuations along the corridor, resulting in smoother transitions in speed. However, the study emphasized that the recurrent activation of the advisory VSL system for short durations could potentially have adverse effects on speed consistency.

A study conducted by Tarko et al. (2019) examined various speed limit scenarios on interstate freeways in Indiana and determined that speed limits had a greater impact on mobility and safety when traffic conditions were uncongested, with limited effects during intermediate traffic conditions. They recommended a uniform speed limit of 70 mph on rural roads to enhance safety and mobility. However, they cautioned against raising speed limits on urban interstates due to safety considerations. In a separate investigation by Hu and Cicchino (2020), the effects of reducing the speed limit from 30 mph to 25 mph in Boston were analyzed, revealing a significant decrease in average speeds and a reduced likelihood of drivers exceeding the speed limit. These findings indicate that the reduction in speed contributed to improved safety by reducing instances of speeding. Silvano and Bang (2016) examined the consequences of changes in PSL on urban roads and identified a modest but statistically significant decrease in average free-flow speeds and speed variability with lower PSL. This decrease in speed and speed variance potentially contributed to a reduction in severe injury crashes. Conversely, raising the PSL led to higher average free-flow speeds without impacting speed variability. Social media data was utilized by Salazar-Miranda et al. (2022) to assess how Paris' slow zones affected the city's street life. According to the study, slow zones, which were put in place to make streets more pedestrian-friendly, increased human activity by 44% when compared to nearby regions without them, as shown by data from Twitter. More users and more tweets per user were responsible for this rise, proving that slow zones attracted more users and promoted higher social media involvement there. Another study by Alhomaidat et al. (2020) examined the consequences of increasing speed limits on freeways on the neighboring urban arterial roads. The results revealed a significant 13.9% rise in the frequency of crashes on adjacent arterials, even with a slight increase in freeway speed limits. This phenomenon, known as speed spillover, indicates that drivers are less likely to adhere to speed limits on arterial roads when the speed limits on freeways are raised. The impact of freeway speed on driver speeding behavior diminishes as the distance from the freeway increases. In a subsequent study, Alhomaidat et al. (2021) compared the speeds of

vehicles exiting the freeway with those already on the adjoining arterial road. They observed differences in average speeds between the two groups, with higher speeds observed on arterials adjacent to freeways with higher speed limits. For instance, on arterials near a 70 mph freeway, passenger cars exhibited higher mean speeds and 85th percentile speeds compared to arterials near a 55 mph freeway.

2.5.3 Design Speed

The concept of design speed pertains to the maximum safe speed at which vehicles can travel on a highway under favorable conditions, low traffic density, and in accordance with the highway's design characteristics (Berry and Belmont, 1951). The definition of design speed has undergone changes over time. Previous definitions focused on the speed adopted by the fastest group of drivers, whereas later definitions emphasized the safe speed that takes into account the highway's design features (AASHTO, 2004). Design speed serves as a basis for determining various geometric design elements of the roadway, including horizontal and vertical curvature, sight lines, super elevation, stopping sight distance, lane widths, and shoulder widths. However, relying solely on design speed does not guarantee alignment with posted speed limits and actual operating speeds, as designers often incorporate conservative design controls that surpass the minimum values specified in design guidelines. Different highway agencies may adopt different approaches for incorporating design speed into their design criteria, with some using it as the sole factor while others consider multiple geometric factors.

2.5.4 Operating Speed

Operating Speed Measures

Data regarding vehicle speeds and travel times are crucial for traffic engineers involved in street and highway design and operation. Operating speed refers to the speed chosen by drivers under prevailing conditions and is typically distributed normally. This parameter can be characterized by the mean speed and standard deviation (Donnell et al., 2018, 2009). Various measures of operating speed are significant in this context:

- **Spot speed:** This refers to the instantaneous speed of a vehicle passing at a specific location.
- **Time-mean speed:** Also known as mean speed or average speed, it represents the arithmetic average of all vehicle speeds over a specified period. It is associated with a specific point in time.
- **Space mean speed:** This is the average speed of all vehicles measured at a particular instant while traveling a given length of the roadway. It is a harmonic mean and accounts for spatial measures rather than temporal measures.
- **Standard deviation:** This statistical measure determines the dispersion of the data. The standard deviation of speed measures is the square root of the variance, which is the average of the squared differences from the mean speed.

- **Percentile speeds:** These are speeds at or equal to which a certain percentage of vehicle groups are traveling, such as the 15th percentile (representing slow-speed group), 50th percentile (mid-range speed group), and 85th percentile (high-speed group). The 85th percentile speed has been historically linked to speed limit setting, assuming that most drivers choose a rational speed to minimize risk.
- **Free-flow speed:** This refers to the speed when there are no constraints on the driver due to other vehicles or geometric and traffic control devices on the road, such as curves or traffic signals.
- **Ten-mph pace speed:** This speed range encompasses the largest percentage of vehicles in a distribution of spot speeds at a specific location.
- **Speed dispersion:** Research on speed dispersion characteristics is limited compared to the well-established measures of operating speed, such as percentile speed and space mean speed. Vehicle speed dispersion has been a significant focus in speed research, with the concept initially introduced by Solomon (1964). Solomon's research revealed a U-shaped curve, indicating that crash rates increased as speeds deviated from the average speed. Notably, the slope of the curve was steeper for slower speeds (see Figure 4). Vehicle speed dispersion is typically defined as the variation in vehicle speeds. Different researchers have employed various indicators to quantify speed dispersion, considering the specific objectives, methodologies, and data limitations of their studies. For instance, Wang et al. (2018) proposed two measures: the standard deviation of individual speeds and the average speed difference between neighboring vehicles.

Additionally, it is essential to evaluate the degree to which motor vehicles surpass the speed limits that are posted or specified by law. Traditional crash databases commonly use phrases such as ‘speeding’ or ‘exceeding the designated speed limit’ to categorize these incidents. However, it is important to offer additional clarification regarding these terms. The interpretation of speeding violations should consider the context, as the behavior of a motorcycle speeding in a residential neighborhood varies significantly from that of a large

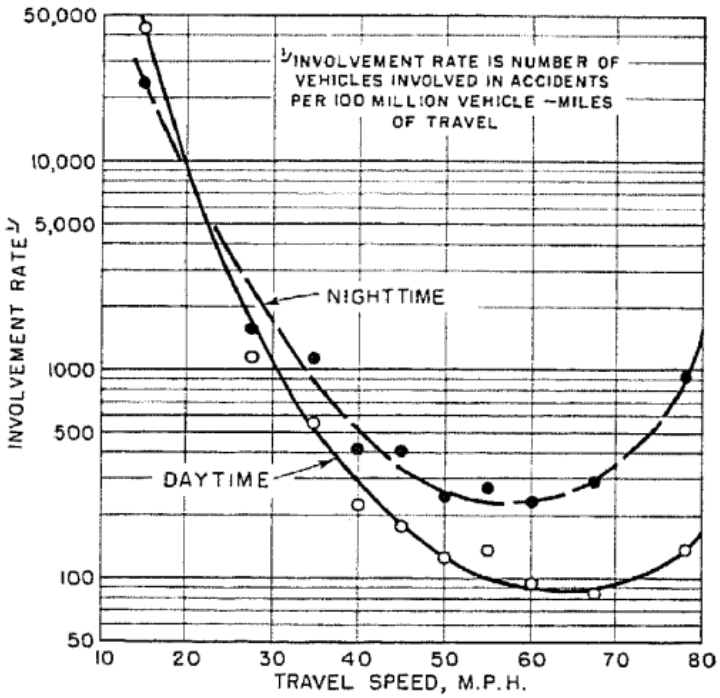


Figure 4. Solomon’s Curve (Solomon, 1964).

truck speeding on a freeway. It is important to recognize that each state has a fundamental speed law in place, which requires drivers to operate their vehicles safely by adhering to appropriate speed thresholds.

Studies on Operating Speed and Safety

The crash-speed relationship has been extensively studied, revealing varying viewpoints among researchers. Hauer (2009) conceptualized this relationship as a 'causal two-link chain' comprising human actions, speed evolution, and safety outcomes. Human activities, including the establishment and enforcement of speed limits, play a crucial role in shaping the evolution of vehicle speeds and drivers' speed choices. These factors, in turn, have a direct impact on roadway safety. The safety implications resulting from the speed evolution are subsequently considered in future decision-making processes related to human activities. The probability of crashes occurring is influenced by "pre-event" probabilities, which determine the number and frequency of crashes, while 'at the time of event' probabilities determine the severity of crashes (Haddon Jr, 1972). Hence, a driver's speed choice not only influences the potential severity of a crash based on the selected operating speed but also affects the likelihood of a crash occurring in the first place.

To gain a deeper understanding of the connection between speed preference and its impact on safety, researchers have utilized various measurements of operational speed. These measurements encompass individual velocities, average speeds along road segments, and variations in speed (Aarts and Schagen, 2006). Prior investigations (Fildes et al., 1991; Haglund and Åberg, 2000; Maycock et al., 1998) have employed data on vehicle speeds and conducted surveys or questionnaires with drivers to establish a clear relationship between higher individual driving speeds and an increased probability of collisions. Earlier studies have revealed a notable correlation between average operational speed and collision rates, particularly with regard to the severity of the crashes (Elvik et al., 2004, 2004; Hauer, 1971). Abdel-Aty and Radwan (2000) examined the extent of speeding in relation to the posted speed limits and found that male and young drivers exhibited a higher tendency to exceed the speed limits. In a study conducted in the United Kingdom, Taylor et al. (2000) uncovered a negative association between the average speed metric and the frequency of collisions overall. Furthermore, when analyzing different homogeneous groups based on road and traffic conditions, the results consistently demonstrated an increase in collision frequency with higher traffic speeds.

A study conducted by Pei et al. (2012) aimed to explore the correlation between speed and the likelihood of crashes. The researchers identified various factors that contribute to this connection, such as the design of the road, prevailing weather conditions, and the distribution of time. The findings of the study revealed a negative relationship between speed and crash risk, indicating that as speed increases, the probability of crashes decreases. Similarly, R. Yu et al., (2013) carried out a study focusing on crash data obtained from I-70 in Colorado over a one-year

duration. They utilized a Bayesian inference model that incorporated real-time variables like weather conditions, traffic, and road geometry. The results of the study demonstrated a significant association between weather conditions and the occurrence of crashes. Consistent with prior research, this study also found that crash segments with lower speeds and higher occupancy in the upstream segment 5-10 minutes before the crash exhibited an elevated risk of crash occurrence. It is important to note that the relationship between speed and crash risk may be influenced by factors such as traffic congestion, as well as other variables that can confound the analysis, such as severe weather conditions. Gargoum and El-Basyouny (2016) uncovered an inverse relationship between variations in speed and the frequency of crashes, indicating that lower speeds were associated with higher crash rates. Similarly, Imprialou et al. (2016) observed a negative correlation between the speed at which vehicles were operated and the frequency of crashes, regardless of their severity. Yu et al. (2018) scrutinized data from urban expressways and discovered that higher operating speeds during congested traffic conditions were linked to a decreased likelihood of crashes. Wang et al. (2018) examined segments of urban arterial roads and ascertained that a 1% rise in the average speed corresponded to a 0.70% increase in the total number of crashes. On rural roads, Dutta and Fontaine (2019) determined that lower average speeds were associated with higher crash frequencies, while an increase in the variability of speeds led to a higher occurrence of crashes. These investigations underscore the significance of managing speed to mitigate the risks of crashes on roads in both urban and rural settings.

In a study conducted by Xu et al. (2019), a semi-automatic filtering technique was utilized to differentiate GPS data points collected on elevated expressways from those obtained on surface roadways. The examination of speed variations involved the implementation of the cross-section speed standard deviation (MCSSD) and the standard deviation of the cross-sectional speed mean (SDCSM). Both hierarchical and non-hierarchical Poisson-gamma models indicated a positive correlation between SDCSM, MCSSD, and the occurrence of crashes. Hutton et al. (2020) utilized data from the Strategic Highway Research Program 2 (SHRP 2) Naturalistic Driving Study to investigate individual driving speeds and their connection to crash likelihood, considering various roadway parameters. Higher speed variations between trips were associated with an increased incidence of collisions. However, few other speed metrics demonstrated a significant relationship with collision frequency. In a path analysis conducted by Park et al. (2021) on city roadways, it was discovered that certain factors such as signalized junctions, traffic volume, and segment length had a favorable impact on collision rates. Conversely, specific roadway features like medians and curbs exhibited direct detrimental effects. The study also provided evidence supporting a causal relationship between crash frequency and speed variability.

In another study by Das et al. (2022) on roads in Dallas during the COVID-19 pandemic, several changes were observed. These included a decrease in traffic volume, an increase in average operating speed, an uptick in the frequency of fatal and serious collisions per mile traveled, and

an overall reduction in traffic volume. The study found that higher operating speeds in 2020 were associated with an increase in collisions, particularly on roads with speed limits set at 60, 65, and 70 mph, as depicted in Figure 5. However, when considering all speed limits and a three-year period, the relationship between average operating speed and crash frequency was found to be minor and negative. This suggests that driving at higher speeds on well-designed highways may not necessarily lead to more crashes. In a subsequent study by Das et al. (2023), a short-duration crash modeling technique was employed to examine the impact of operating speed on highway collisions during the COVID-19 period. The findings revealed variations in the effects of speed regulations on crash rates across different years and levels of collision severity. The influence of speed measurements was more pronounced in the 2020 models, indicating that the impact of speed measures depends on the severity of the crashes being studied. Additionally, the magnitude of the speed effects varied for different levels of collision severity.

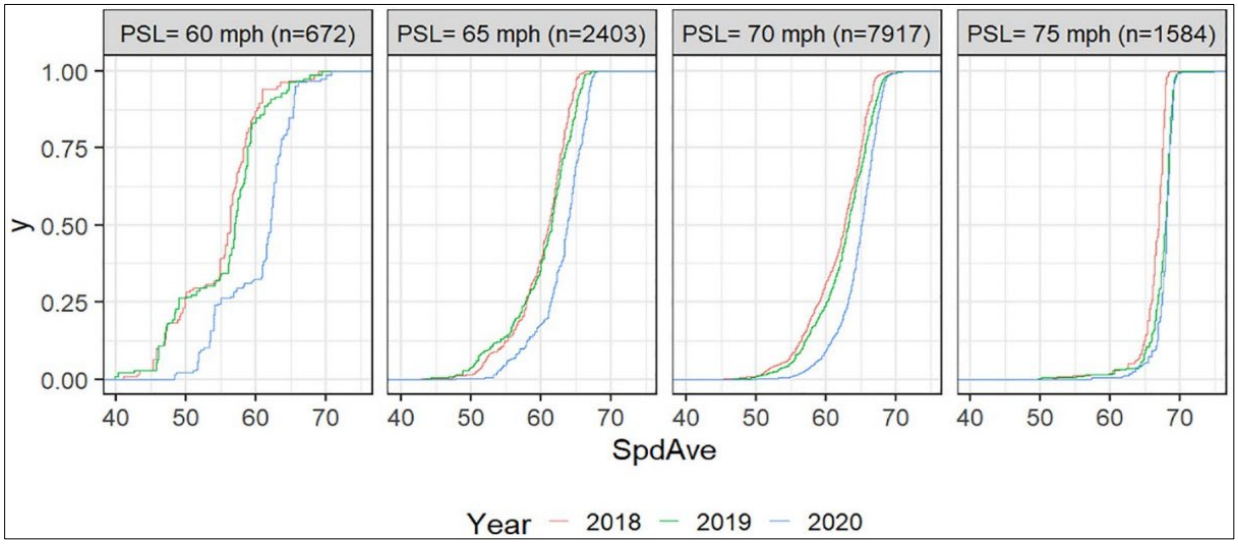


Figure 5. Cumulative Speed Plots by Posted Speed Limit (S. Das et al., 2022).

2.5.5 Speeding

Various terms have been used by researchers to describe speeding incidents. One such term is 'Free-Flow Episode,' which describes a scenario where a driver travels at or above a speed threshold set at 5 mph below the posted speed limit. For this to be included, the speed must be maintained for a minimum of 30 seconds. Another term is 'Speeding Episode,' which involves continuous driving at or above a speed threshold set at 10 mph over the posted speed limit, as shown in Figure 6. To qualify as an episode, this speed threshold must be sustained for at least 6 seconds (Richard et al., 2013b, 2016). According to another study by the same researcher, the definition of speeding varies depending on several factors, such as ad hoc or analytical criteria, risk and kinematics, psychological and subjective speeds, and behavior-based approaches (Richard et al., 2013a). Another study defines speeding as exceeding the posted speed limits or driving at speeds considered too high for the road or weather conditions at a specific location or time (Gargoum and El-Basyouny, 2016).

The detrimental impact of speeding on road safety should not be underestimated, given its profound influence on both the probability and consequences of vehicular collisions. Recent data released by the NHTSA in 2020 indicates that speeding was responsible for a distressing loss of 11,258 lives (NHTSA, 2020). NHTSA employs a comprehensive definition of a speeding-related crash, incorporating situations where any involved driver faces charges related to speeding or where law enforcement officials ascertain that factors such as racing, driving at excessive speeds given prevailing conditions, or surpassing the posted speed limit contributed to the incident. While there has been a modest decline in speeding-related crashes in recent years, it remains a predominant concern in the realm of traffic safety, as illustrated in Figure 7. Additionally, within the domain of urban and suburban arterials, where a significant number of pedestrian and bicycle fatalities occur, speeding has emerged as a pivotal causal factor in such regrettable events (Cai et al., 2021; Goel, 2021).

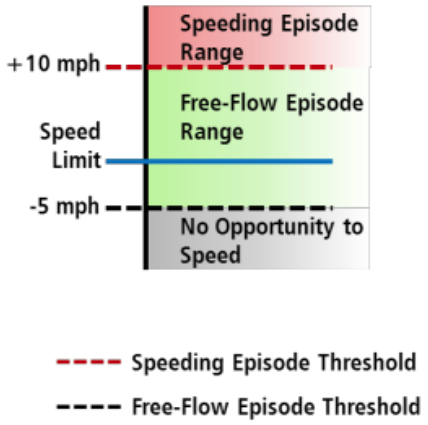


Figure 6. Definition of Speeding (Richard et al., 2016).

Studies on Speeding and Safety

A recent publication by the FHWA emphasized the significance of advancing speed management as a means to decrease injuries and fatalities associated with speeding on the nation's roadways (Xu et al., 2022). The COVID-19 pandemic showed an increase in excessive speeding nationwide, particularly on urban interstates, with drivers exceeding 100 mph. Despite a 13.2% reduction in VMT in 2020, the fatality rate rose to 1.37 fatalities per 100 million VMT, projecting a 7.2% rise in total fatalities and an 11% rise in speeding-related fatalities. Factors such as less congestion and high-risk drivers contributed to this trend. Beliefs, attitudes, and societal acceptance of speeding significantly influence drivers' behavior. This study underscores

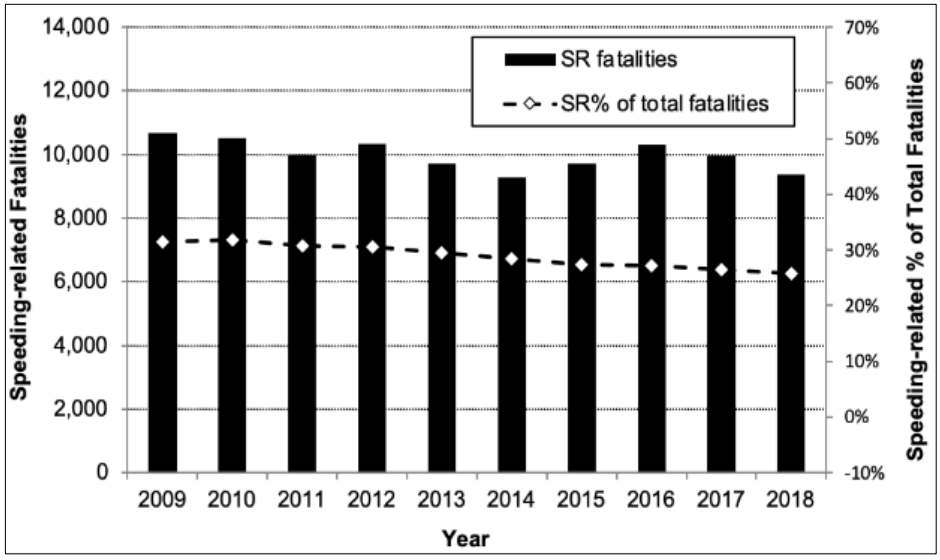


Figure 7. Speeding Related Fatalities (NHTSA, 2020).

the importance of speed management in addressing safety concerns related to speeding. Speed management involves establishing appropriate speed limits, reducing speeding, and mitigating speeding-related crashes. The USDOT Intermodal Speed Management Team, comprising the NHTSA, FHWA, and the Federal Motor Carrier Safety Administration (FMCSA), identified the following key approaches for implementing the USDOT speed management program (Xu et al., 2022).

- Establishing and executing comprehensive speed management programs and strategies at the jurisdictional level.
- Developing strategies for setting speed limits that prioritize the safety and needs of all road users, taking into account contextual factors and not solely relying on drivers' observed speeds.
- Implementing evidence-based safety measures to promote safe speeds and protect the well-being of all individuals using the road network.
- Enhancing the reporting of crash data by incorporating specific information on speeding-related incidents, ensuring consistency and enabling the identification of contributing factors.
- Implementing visible and transparent enforcement measures, alongside educational initiatives and awareness campaigns, rather than solely focusing on enforcement actions.
- Incorporating considerations of equity into the decision-making process for speed management.

In a study conducted by Wang and Cicchino (2023) on excessive speed during the COVID-19 pandemic, it was observed that a significant proportion of vehicles were traveling at speeds 5 to 10 mph above the posted limit. The research focused on urban expressways and major roads, with the highest occurrences of speeding observed during weekday rush hours and weekend afternoons. After accounting for factors such as road type, time of day, day of the week, and traffic volume, the analysis revealed a 22% increase in the likelihood of exceeding the speed limit by 5 mph in 2020 compared to 2019, along with a 51% increase in the risk of surpassing the speed limit by 10 mph. These findings indicate a concerning surge in speeding incidents during the COVID-19 pandemic. Stiles et al. (2023) conducted a study to examine the influence of COVID-19 stay-at-home measures on alterations in collision characteristics, timing, and severity on urban streets, considering the impact of reduced traffic volumes and increased velocities. The results revealed a robust correlation between diminished traffic levels and heightened crash severity. The findings indicated that higher speeds were associated with more severe collisions, and there was a reduction in crash occurrences during the morning peak hours. Furthermore, there was a significant decrease in the frequency of collisions, which are typically more prevalent in congested areas. Additionally, the data demonstrated an increase in the proportion of collisions attributed to drunk driving and speeding, underscoring the significance of these factors in comprehending the shifts in crash patterns throughout the course of the pandemic.

In the research conducted by Yadav and Velaga (2020), the focus was placed on investigating the influence of alcohol consumption on speeding behavior, while also analyzing the probabilities of crashes occurring. The study employed driving simulator experiments, wherein participants were subjected to various Blood Alcohol Concentration (BAC) levels, namely 0%, 0.03%, 0.05%, and 0.08%. The findings of the study indicated a noteworthy escalation in driving speed as the BAC levels increased. Moreover, the investigation revealed that crash probabilities were significantly greater in urban settings compared to rural environments across all BAC levels.

Using survey data collected between the years 2000 and 2018, the study by Sultana (2018) examined the trend in driver speed behavior on Perth's metropolitan road network in Australia. The results showed that, in 2000, 53% of cars in the metropolitan network adhered to or went below the prescribed speed limits. The compliance rate significantly increased on average between 2003 and 2015, reaching 64.1% in 2015 compared to the compliance rate seen in 2000. Additionally, the survey that was done in 2018 showed a notable jump in compliance rates, with a climb of 5.3% to a record high of 69.5%, as depicted in Figure 8. The findings indicate the need for speed enforcement strategies to account for variations in driver speed behaviors related to factors such as road type, speed limit, and temporal variables including the day of the week and time of day.

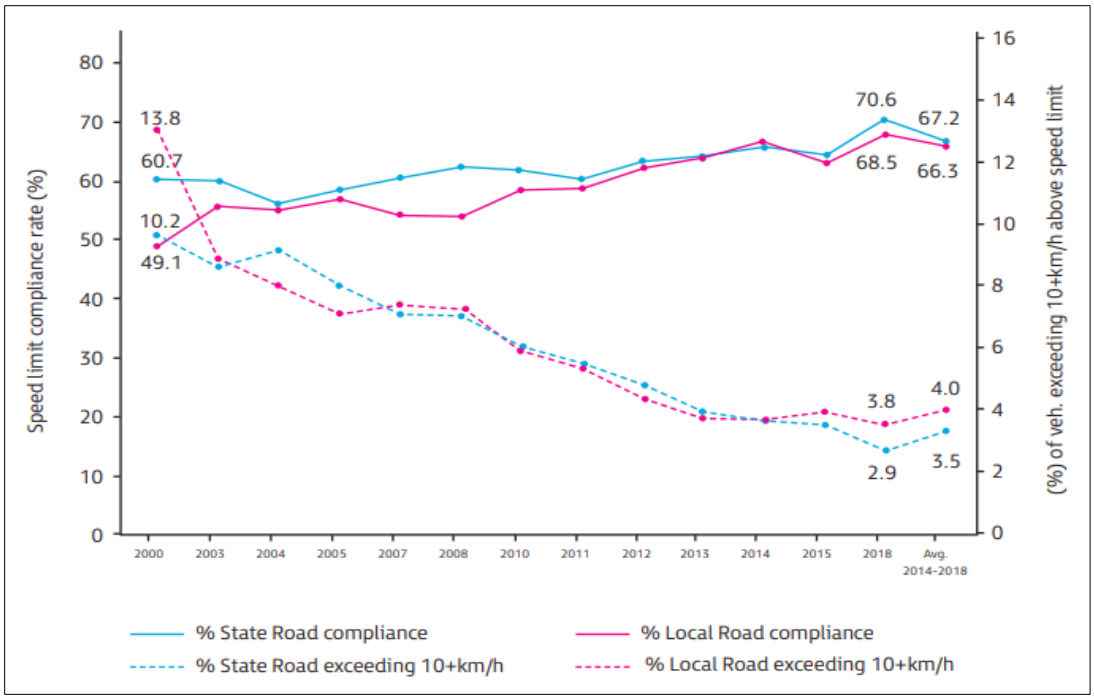


Figure 8. Speeding Trend in Perth's Metropolitan Road Network in Australia (Sultana, 2018).

Richard et al. (2013b) conducted a study on the speeding behavior of drivers in urban and rural settings over a three to four-week period of naturalistic driving. The research aimed to analyze speeding causes, factors, prediction, classification, and proposed interventions for reducing speeding incidents. The study identified four distinct categories of speeding behaviors exhibited by individual drivers, as illustrated in Figure 9. These included infrequent or unintentional instances of exceeding speed limits, trip-specific situational speeding, casual speeding characterized by frequent but minor speed violations per trip, and habitual or chronic speeding. Regression models, employing both logistic regressions to predict the likelihood of speeding and linear regressions to determine the extent of speeding, were developed. The examination identified noteworthy indicators of excessive speed, encompassing factors associated with the driver, such as sex and age, contextual factors like the hour of the day and day of the week, and behavioral factors such as attitudes towards irresponsible driving. Furthermore, an extension of the research conducted in this study identified various types of speeding behaviors and driver categories, as illustrated in Figure 10. The research produced several significant discoveries. Firstly, the presence of more hazardous elements in instances of speeding confirmed the widely held belief that different forms of speeding are linked to varying levels of risk. Secondly, observations from personal accounts indicated that location-specific factors played a role in the occurrence or absence of speeding events. Thirdly, indirect measures suggested that certain aspects of the

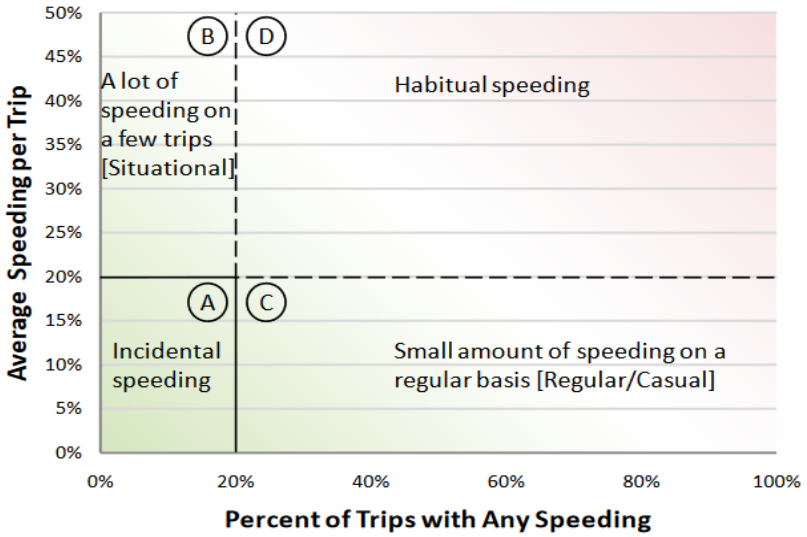


Figure 9. Types of Speeding (Richard et al., 2013b).

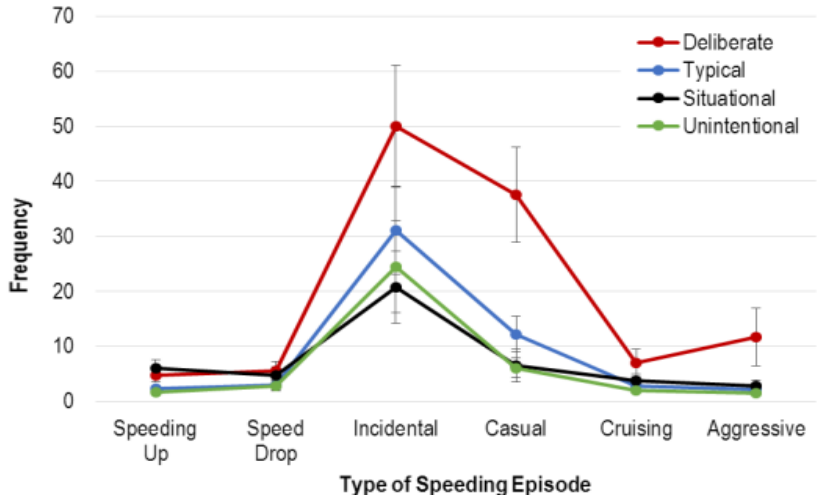


Figure 10. Average frequency of different types of Speeding Episodes by Driver Types sites (Richard et al., 2016).

driving environment had an impact on speeding behavior. Lastly, similarities in types of speeding were observed in both rural and urban areas where data was collected (Richard et al., 2016).

A report issued by the World Health Organization (WHO) highlighted several key factors that contribute to drivers' speeding behavior and emphasized the importance of speed management in enhancing roadway safety (WHO, 2017). Apart from the designated speed limits, drivers' speeds are influenced by various factors, including the driver's age and gender. Male drivers and young drivers tend to exhibit a higher likelihood of speeding, which consequently leads to an overrepresentation of these groups in speed-related crashes. Additionally, the driver's blood alcohol concentration, as well as road layout, surface quality, and the vehicle's power and maximum speed, can also impact speed-related behaviors (see Figure 11).

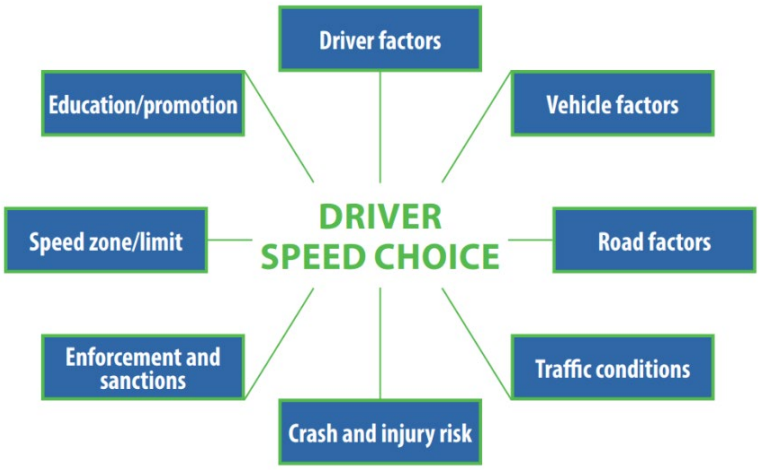


Figure 11. Factors Affecting Speed Choice (WHO, 2017).

In a study conducted by Cai et al. (2021), the effects of speed management strategies on the proportions of speeding incidents were examined in urban and suburban arterials. Probe speed data was utilized to calculate speeding proportions using a Beta regression model and grouped random parameter modeling. The results demonstrated that the grouped random parameter model outperformed alternative approaches, allowing for a better understanding of the diverse effects of road features and other contributing factors on speeding across different road types. Findings indicated that road segments with more intersections tended to exhibit lower speeds, resulting in adjusted proportions of different speed ranges. Moreover, the presence of asphalt pavement was associated with increased speeding proportions across all arterial types, while strategies such as lane narrowing and short blocks showed potential for decreasing the proportion of speeding incidents on suburban commercial roads.

A study conducted by Tankasem et al. (2022) investigated the impact of automatic speed control (ASC) on speeding behavior and intentions on urban arterial highways with mixed traffic. The research findings revealed cognitive shifts resulting from ASC in relation to speeding behavior. Specifically, the perception of speeding decreased, with reduced acceptance and increased

resistance to regulation. Drivers reported a decrease in both actual speeding behavior and the inclination to speed. Furthermore, ASC altered the relative significance of various factors, amplifying the influence of close friends and family members on drivers' propensity to speed, while emphasizing the role of intention and perceived control in shaping speeding behavior.

2.5.6 Safe Speeds

The Safe System Approach (SSA) has been widely adopted by the U.S. Department of Transportation (USDOT) as the primary framework for addressing road safety concerns. This comprehensive strategy is highly regarded within the transportation community as an effective approach to mitigate risks in our complex and extensive transportation network. The SSA aims to prevent crashes and minimize the harm caused when crashes do occur. It takes a holistic view, recognizing both human errors and vulnerabilities and incorporates multiple measures to protect all individuals. In line with the SSA, the DOT's National Roadway Safety Strategy and ongoing safety initiatives are committed to achieving a future where there are no fatalities or severe injuries on the roads. In accordance with the core principles of the SSA, the FHWA has identified five pillars: safe road users, safe vehicles, safe speeds, safe roads, and post-crash care (see Figure 12). Within the SSA, safe speeds play a critical role in reducing the risk of crashes and the severity of injuries. Recognizing the importance of implementing safe speeds across the entire road network, the SSA emphasizes a combination of measures. These include setting appropriate speed limits, designing roads that encourage safe speeds, and utilizing technologies like speed cameras to regulate vehicle velocities. Establishing suitable speed limits is a vital aspect of speed management and should take into account factors such as road design, traffic volume, and the presence of pedestrians and cyclists in the road environment. By incorporating these measures, the SSA aims to enhance roadway safety and create environments that prioritize the well-being of all road users (Finkel et al., 2020; USDOT, 2022a).

Figure 12. Principles and Pillars of SSA.

Energy Transfer

Traveling at high speeds increases the likelihood of crashes and exacerbates the resulting injuries. This is primarily due to the limited time available for drivers to respond to unexpected events and the greater kinetic energy generated by faster-moving vehicles, which intensifies the impact of collisions. Kinetic energy refers to the energy that is related to the motion of an object and is dependent on both its mass and velocity, as demonstrated by Equation (35). The mathematical relationship between kinetic energy and velocity is quadratic, meaning that any increase in velocity results in a disproportionately greater increase in kinetic energy. For example, doubling the velocity of an object will result in a four-fold increase in its kinetic energy, while tripling the velocity will lead to a nine-fold increase in kinetic energy. As such, even minor changes in velocity can have a significant impact on the energy generated during a collision.

$$E_k = \frac{1}{2}mv^2 \tag{35}$$

Where,
E_k = Kinetic energy (Joules)
m = Mass (kg)
v = Velocity (m/s)

Effect of Speed on Crash and Injury Severity

Numerous studies have confirmed the significant role of speed in crash likelihood. Elvik (2013) conducted research that revealed a decrease in the likelihood of casualty crashes when mean traffic speeds were reduced in response to speed limit reductions. Similarly, Kloeden et al. (2002) established a relationship between driver speed exceeding the speed limit and an increased likelihood of involvement in a casualty crash. These findings underscore the fact that even minor reductions in speed can yield substantial decreases in road crashes. Furthermore, extensive research has explored the relationship between speed and crash severity. Elvik (2013) demonstrated that fatal crashes exhibited a more substantial decline compared to all injury crashes when mean speed was reduced. In other words, a reduction in mean speed corresponded to a decrease in crash severity. A model presented by Wramborg (2005) provides insights into how speed influences the severity of specific types of crashes. This study introduced three relationships between impact speed and the likelihood of fatalities, as illustrated in Figure 13. These relationships assume that the involved vehicles have equal mass and velocity. According to these probability curves, pedestrian/cyclist incidents at speeds of 19 mph, side impact collisions at speeds of 31 mph, and head-on collisions at speeds of 44 mph have a 10% probability of resulting in a fatality (Hall et al., 2021; Jurewicz et al., 2015) (see Figure 13).

Tingvall and Haworth, (1999) examined the speed thresholds in relation to the type of infrastructure and traffic. Their findings suggest that if a pedestrian is struck by a well-designed car traveling at speeds exceeding approximately 30 km/h (18 mph), the human tolerance threshold is likely to be surpassed. In urban areas where there is a desire for higher speeds, one effective approach is to separate pedestrian crossings from the flow of traffic. Alternatively, pedestrian crossings, zones, or vehicles must be carefully designed to ensure that speeds are limited to a maximum of 30 km/h (18 mph). This principle can also be extended to infrastructures where only car-to-car collisions are possible. While it is generally expected that well-designed cars will possess a maximum safety threshold of 70 km/h (43 mph) for frontal impacts and 50 km/h (31 mph) for side impacts, it is possible to tolerate higher speeds if the interface between the vehicle and the infrastructure is thoughtfully engineered. In fact, speeds exceeding 100 km/h (62 mph) may be considered permissible under certain circumstances.

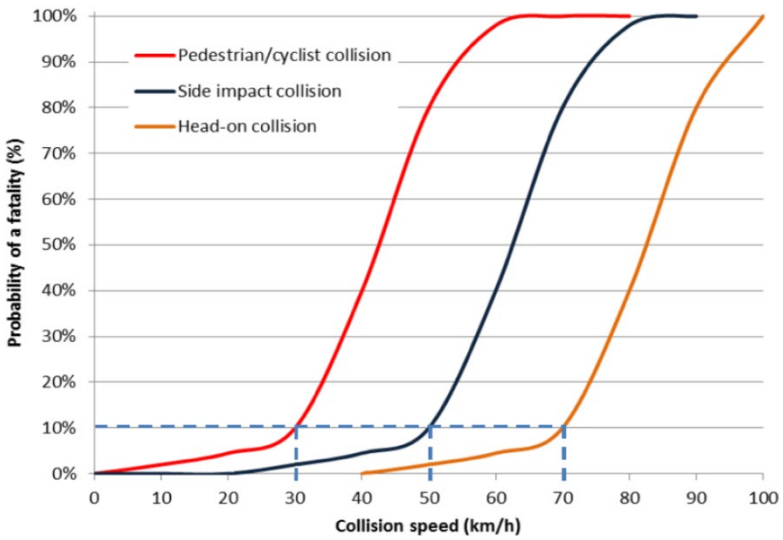


Figure 13. Relationships Between a Motorised Vehicle Collision Speed and Probability of a Fatality (Jurewicz et al., 2015; Wramborg, 2005).

Table 21 outlines the potential long-term maximum travel speeds that can be associated with infrastructure, assuming the adoption of best practices in vehicle design and the utilization of 100% restraint systems.

Table 21. Possible Long-Term Maximum Travel Speeds (Tingvall and Haworth, 1999).

Type of infrastructure and traffic	Possible travel speed	
	kmh	mph
Areas where pedestrians and vehicles may come into conflict	30	18
Intersections with potential side impacts between vehicles	50	31
Roads where vehicles may collide head-on	70	43
Roadways where there is no chance of a side or frontal collision, only collision with the infrastructure	100+	62+

Power models illustrating the connections between average speed and the occurrence of injury or fatal crashes, as well as the number of injuries or fatalities, were presented by Nilsson (2004).

Figure 14 visually represents these relationships, clearly showing that as speed increases, there is a corresponding rise in the percentage change in casualties. Furthermore, Elvik et al. (2004) conducted an extensive meta-analysis of studies examining the correlation between travel speeds and casualty rates. This analysis encompassed 98 distinct studies, which collectively offered 460 estimates regarding the association between changes in mean traffic speed on a road and changes in the fatality rate. The included studies were conducted between 1966 and 2004, with around half of the estimates derived from studies conducted after 1990. Both rural and urban roads were considered, covering a speed range of approximately 25 km/h (15 mph) to 120 km/h (75 mph). Nilsson (2004) presented power models that illustrate the relationships between average speed and the incidence of injury or fatal crashes, as well as the number of injuries or fatalities. These relationships are visually represented in Figure 14, where it is evident that as speed increases, there is a corresponding increase in the percentage change in casualties. In addition to Nilsson's models, Elvik et al. (2004) conducted a comprehensive meta-analysis of research investigating the association between travel speeds and casualty rates. This analysis encompassed 98 distinct studies, which collectively provided 460 estimates regarding the relationship between changes in mean traffic speed on a road and changes in the casualty rate. The data included studies conducted between 1966 and 2004, with approximately half of the estimates derived from studies conducted after 1990. The study encompassed both rural and urban roads and considered a speed range spanning from approximately 25 km/h (15 mph) to about 120 km/h (75 mph). Data from 20 countries were included in the analysis. The findings of this meta-analysis are summarized in Table 22, providing a comprehensive overview of the results.

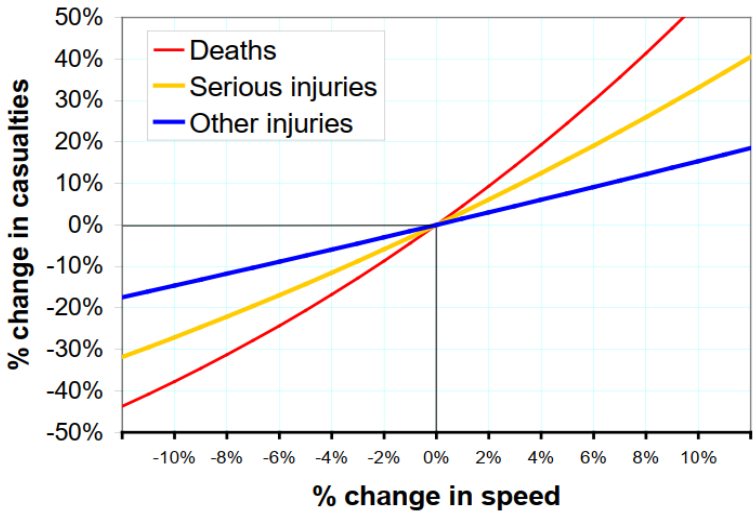


Figure 14. Relationship Speed Changes and Changes in Casualty Rates (Elvik et al., 2004; Nilsson, 2004).

Table 22. Relationship Speed Changes and Changes in Casualty Rates (Elvik et al., 2004; Hall et al., 2021).

Change in:	Change in Mean Speed					
	Speed reduction			Speed increase		
	-10%	-5%	-1%	+1%	+5%	+10%
Deaths	-38%	-21%	-4%	+5%	+25%	+54%
Serious Injuries	-27%	-14%	-3%	+3%	+16%	+33%

Other Injuries	-15%	-7%	-1%	+2%	+8%	+15%
Property Damage Crashes	-10%	-5%	-1%	+1%	+5%	+10%

In a study conducted by Tefft (2013), the risk of severe injury or fatality for pedestrians involved in collisions with forward-moving vehicles was estimated. The analysis utilized crash data from the US and focused on pedestrian impacts by cars, light trucks, vans, and sport utility vehicles. The findings revealed that the average risk of a pedestrian sustaining an injury classified as Abbreviated Injury Scale 4 or greater severity increased with impact speed. At an impact speed of 17.1 mph, the average risk reached 10%. This risk escalated to 25% at 24.9 mph, 50% at 33.0 mph, 75% at 40.8 mph, and 90% at 48.1 mph, as demonstrated in Figure 15 and Figure 16. Similarly, the average risk of death showed a similar trend. At an impact speed of 24.1 mph, the risk of death reached 10%, followed by 25% at 32.5 mph, 50% at 40.6 mph, 75% at 48.0 mph, and 90% at 54.6 mph.

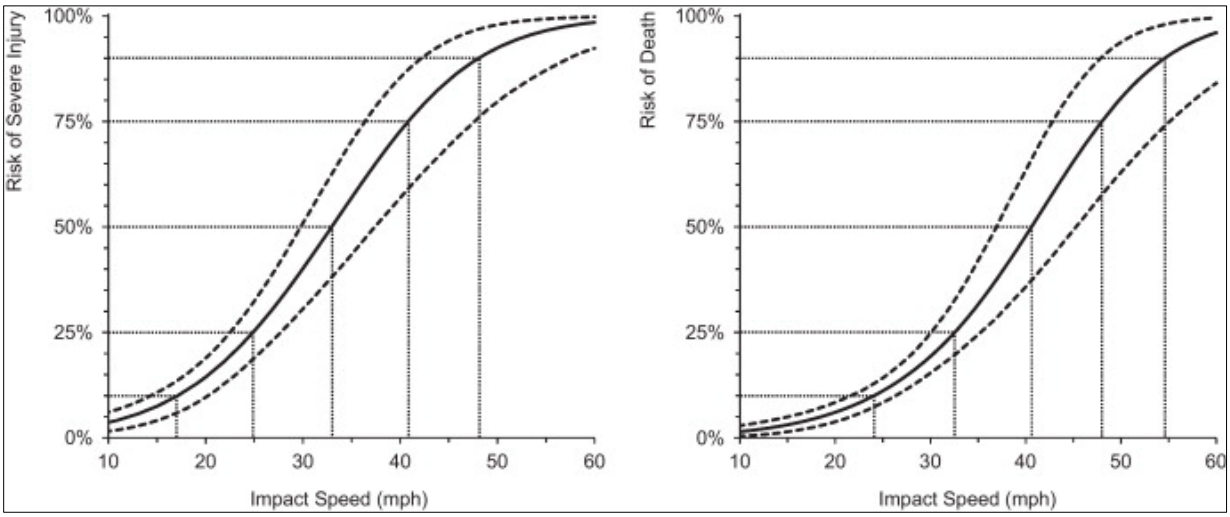


Figure 15. Risk of Severe Injury (Left) and Death (Right) of Pedestrians in Relation to Impact Speed (Tefft, 2013).

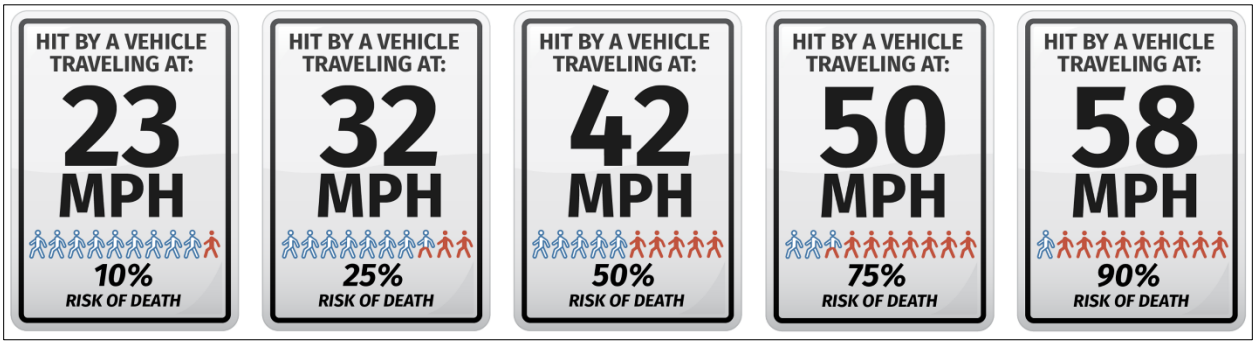


Figure 16. Impact Speed and a Pedestrian's Risk of Death (USDOT, 2022b).

Studies on Safe Speed

In the US, cities and states have been increasingly adopting the SSA to tackle transportation inequalities, safeguard vulnerable road users, and establish safer speed limits. In a recent study, Ngo et al. (2022) presented examples of how two states, Washington and California, and two cities, Philadelphia and Portland, Oregon, have implemented this approach to achieve the ambitious target of zero fatalities. The study revealed that these states and cities have taken substantial steps to institutionalize the SSA and have utilized it as a cornerstone for their policies that shape their operations at both program and project levels. By adopting this approach, they aim to transform the transportation system into a more equitable, efficient, and safer one. The study highlights the significance of implementing the SSA to achieve road safety goals and the importance of institutionalizing it into policies for effective and long-lasting results.

Steinmetz et al. (2018) conducted a study to assess the effectiveness of the SSA in enhancing road safety in Mildura, Australia. The study included two speed management scenarios, Treatment scenario 1 and Treatment scenario 2, both of which included two broad speed limit regions. Additional changes were made to certain roads in Treatment scenario 2 to further supplement the speed management plan. Australian Road Assessment Program (AusRAP) star rating assessments and Australian National Risk Assessment Model (ANRAM) fatal and serious injury crash estimates were conducted to compare the treatment scenarios with the baseline scenario. The study found that Treatment scenario 1 reduced the proportion of the network with 1 or 2-star ratings and increased the proportion with 3+ stars by 32%. It also reduced Fatal and Serious Injury crashes by 45%. Treatment scenario 2 had a similar impact with the added road modifications. These findings highlight the effectiveness of the SSA and demonstrate the significance of implementing speed management plans to enhance road safety.

In response to the persistent occurrence of serious injuries resulting from crashes at intersections, Candappa et al. (2013) conducted an in-depth investigation of these crashes in Victoria. The study identified a key principle that emphasized the need to limit side impact crash speeds to under 31 mph, as exceeding this speed threshold exceeds the biomechanical tolerance of the human body. Additionally, minimizing angles and conflict points was recognized as an essential principle for improving intersection safety. The investigation produced several existing and new designs that incorporate these principles, which have the potential to enhance intersection safety. Another study by the same authors implemented an innovative trial at a signalized intersection in Victoria, Australia with the aim of aligning approach speeds to SS speeds. The trial included a combination of signals, reduced speed limits, and a Safety Platform. Speeds were measured using pneumatic tubes, and video footage was taken over a two-week period. The study found that the reduced speeds resulted in crash kinetic energy (KE) levels that were more aligned with SS principles, with KE levels estimated to be just above the recommended level of 96.5 kJ, compared to Control 1 and 2 where KE levels were closer to double the tolerable levels at 189

kJ. The trial suggests that aligning approach speeds with SS principles can lead to safer outcomes for road users (Candappa et al., 2016).

The implementation of the Safe System (SS) policy requires that speed limits for the road and traffic system be designed based on human biomechanical and competency parameters, taking crash injury severity factors into consideration. While most Australian States have yet to fully apply SSA principles to speed zoning, the NSW Speed Zoning Guidelines aim to guide the setting of speed limits that balance mobility, road safety, and community concerns. However, vocal opposition to lowering speed limits often hinders governments from implementing safer limits. In this context, Mooren et al. (2014) analyzed the NSW practices and identifies specific departures from the SSA in setting speed limits, as well as ways to shift community attitudes towards safer speeds. The study suggests that setting speed limits based on the 85th percentile of free travel speeds is irresponsible and dangerous.

Jurewicz and Turner (2011) conducted a study with the aim of exploring the emerging idea of setting speed limits based on severe crash risks, which result in fatal and serious injury outcomes. The research was based on Austroads studies and proposed several alternative approaches to determine speed limits. These approaches included evaluating crash history, assessing inherent crash risk due to the road environment, or a combination of both. Additionally, the intended road function was considered an indicator of the mobility level expected by the public. To apply a risk-based approach to speed limits, the study explored the use of established network-level risk assessment tools, such as AusRAP. The authors stressed the significance of considering not only the road environment but also the mobility needs of the public in setting speed limits. The study provided valuable insights into alternative approaches to speed limit setting, which could potentially reduce the risk of severe crashes on roads.

To achieve a SS, it is crucial to align the speed limit with the infrastructure, or vice versa. The idea of reducing speed limits in areas where the road environment changes is not novel. For several decades, Victoria has been lowering speed limits in areas with rough surfaces, road events, or roadwork to mitigate safety risks. The study by Beer (2011) discussed the challenges associated with implementing a speed limit policy that would genuinely achieve a SS. The study made several key findings, including the roads farthest from the SS being 62 mph and 68 mph roads with possible collisions with fixed poles or trees. The study also suggested that investing in wire-rope safety barriers to prevent run-off-road crashes could improve the road environment towards a SS without changing the speed limit. From a SS perspective, these areas should have speed limits below 25 mph. If infrastructure is not readily available to mitigate the risk to vulnerable road users, the study suggests lowering speed limits in these areas to 25 mph, assuming a reasonable level of exposure.

Harm minimization is a promising strategy for setting speed limits and achieving safer transportation systems. Recognizing this potential, Austroads commissioned a study with the aim of striking a balance between harm minimization and mobility when establishing speed limits. In pursuit of this objective, Fildes et al. (2006) reported on developments to date as well as future initiatives in the field. Building on these efforts, Jurewicz and Hall (2009) presented an approach for setting speed limits based on harm minimization. The approach was developed with the aid of a recent Austroads project conducted by the Australian Road Research Board (ARRB). The project aimed to review the principles for setting speed limits in the context of the SS, which extended beyond the current guidelines. The project extensively reviewed literature and analyzed data to explore the relationship between road infrastructure, driver speeds, speed limits, and crash outcomes. A panel of speed management policy stakeholders from Austroads assessed the role of road infrastructure features in speed limit setting, resulting in recommendations for appropriate speed limit setting policy and principles that align with the SS framework. The study also identified possible intermediate speed management policy options to assist the transition toward the SS. Jurewicz (2010) introduced a new set of principles for setting speed limits based on harm minimization, which is a fundamental aspect of the SSA. The SS aims to eliminate serious injuries and fatalities in the road transportation system by focusing on safer roads and reduced speeds. The ARRB and road jurisdiction stakeholders have developed revised speed limit-setting guidelines for Austroads. These guidelines provide a framework for assessing the readiness of a route to align with the SS principles and determining appropriate speed limits and necessary road improvements to ensure passenger safety. It is acknowledged that implementing all SS road characteristics may not be economically feasible in the short and medium term, and not all speed limit changes may be instantly welcomed by the public. Therefore, the study proposes various harm-reduction strategies as an initial step towards implementing the SS.

Road safety remains a significant issue worldwide, and effective speed management plays a crucial role in achieving safer roads. The importance of implementing the SSA to enhance road safety by aligning speeds with human biomechanical and competency parameters. This approach serves as a foundation for establishing consistent and safe speed limits. The findings suggest that the implementation of speed management plans can lead to a significant reduction in low-rated roads and fatal and severe injury crashes. However, opposition from communities may pose challenges to implementing safer speed limits. Therefore, educating the public and involving them in the decision-making process is crucial to achieving long-lasting results. Table 23 provides a comprehensive summary of the studies reviewed concerning safe speeds in the context of the SSA.

Table 23. Summary of Studies on Safe Speeds.

Study	Aim/Objective	Key Findings/ Summary
Ngo et al. (2022)	Present examples of states and cities implementing the SSA to achieve zero fatalities	▪ Washington, California, Philadelphia, and Portland have institutionalized the SSA to shape their safety policies. ▪ The SSA can transform the transportation system into a safer, more equitable, and more efficient one.
Steinmetz et al. (2018)	Assess the effectiveness of SSA in enhancing road safety in Mildura, Australia	▪ The study included two speed management scenarios with different speed limits. ▪ Treatment scenario 1 reduced FSI crashes by 45%. ▪ Treatment scenario 2, with additional road modifications, had a similar impact, demonstrating the effectiveness of the SSA.
Candappa et al. (2013)	Investigate crashes at intersections and identify key principles for improving intersection safety	▪ Limiting side impact crash speeds to under 31 mph is crucial for intersection safety. ▪ Minimizing angles and conflict points are essential principles for improving intersection safety. ▪ Existing and new intersection designs can be modified to incorporate these principles to enhance safety.
Candappa et al. (2016)	Implement an innovative trial to align approach speeds to SS speeds	▪ An innovative trial implemented at a signalized intersection in Victoria, Australia, aimed at aligning approach speeds to SS speeds. ▪ Reduced speeds resulted in crash kinetic energy (KE) levels more aligned with SS principles. ▪ Aligning approach speeds with SS principles can lead to safer outcomes for road users.
Mooren et al. (2014)	Analyze NSW practices and identify specific departures from the SSA in setting speed limits	▪ Opposition to lower speed limits can hinder governments from implementing safer limits. ▪ Setting speed limits based on the 85th percentile of free travel speeds is irresponsible and dangerous, and there are ways to shift community attitudes toward safer speeds.
Jurewicz and Turner (2011)	Explore the emerging idea of setting speed limits based on severe crash risks	▪ Emphasized the significance of considering the intended road function and mobility needs of the public in setting speed limits. ▪ Explored the use of network-level risk assessment tools.
Beer (2011)	Discuss challenges associated with implementing a speed limit policy that achieves a SS	▪ 62 mph and 68 mph roads have the highest risk of collisions with fixed poles or trees, making them farthest from the SS. ▪ Speed limits around schools and areas with vulnerable road users have the highest community outrage and should be lowered to 25 mph from a SS perspective.

Study	Aim/Objective	Key Findings/ Summary
Fildes et al. (2006)	Evaluate the potential of harm minimization in speed limits	▪ Australian speed limits are higher than those internationally, prompting the need for a SSA. ▪ The SSA aims to balance harm minimization and mobility when setting speed limits.
Jurewicz and Hall (2009)	Develop an approach to setting speed limits based on harm minimization	▪ Harm minimization approach can guide speed limit setting consistent with SS principles. ▪ Road infrastructure features should be considered in speed limit setting. ▪ Intermediate speed management policy options can guide the transition toward SS
Jurewicz (2010)	Introduce a new set of principles for setting speed limits based on harm minimization	▪ The newly introduced speed limit setting principles provide a framework for assessing a road's SS readiness. ▪ SS road features may be economically viable in the short and medium term and not all speed limits would be immediately acceptable to the public.

2.6 IMPACT OF WEATHER ON SAFETY

Crashes that take place in unfavorable weather conditions such as rain, sleet, snow, fog, icy or wet pavement, and other weather-related factors are known as adverse or inclement weather-related crashes. Weather conditions encompass various factors such as reduced visibility, precipitation, strong winds, and extreme temperatures, all of which can impact the friction of the road surface, driver performance, and vehicle condition. These factors have the potential to elevate the likelihood and seriousness of crashes. Numerous research studies have examined driver behavior and the frequency of crashes during adverse weather conditions. A concise overview of these pertinent studies is provided below.

2.6.1 Pedestrian and E-scooter Safety

Asli (2022) examined pedestrian incidents in urban circular junctions and identified lighting, road surface characteristics, and atmospheric conditions as significant factors. Qiu and Fan (2022) determined that urban environments and wet road surfaces reduce the probability of fatal pedestrian injuries at intersections. Ferreira et al. (2022) identified factors that adversely affect the safety perception of pedestrians and cyclists, including pollution, lack of vegetation, inclement weather, inclines, and long commuting distances. Pobudzei et al. (2023) investigated the severity of injuries in pedestrian-vehicle crashes and revealed that e-scooter incidents increase during favorable weather conditions. Cerný et al. (2023) explored the limitations of automatic emergency braking (AEB) systems in detecting pedestrians and found that the efficiency of AEB is influenced by vehicle speed and weather conditions. In sunny weather, vehicles are 20% more inclined to come to a halt in front of pedestrians compared to light rain.

2.6.2 Connected and Automated Vehicle Safety

Environmental factors can have an impact on the effectiveness of advanced driver assistance systems (ADAS). Research has indicated that the overall risk of crashes during snowfall remains consistent over time (Andrey, 2010). ADAS has shown effectiveness in reducing crashes, especially in urban areas with clear daylight conditions and rural areas with clear daylight conditions Masello et al. (2022). The ability of (AEB) systems to detect pedestrians is influenced by vehicle speed and environmental conditions, resulting in a higher probability of stopping in sunny weather compared to light rain (Cerný et al., 2023). Autonomous vehicles and connected autonomous vehicles have the potential to contribute to the reduction of pedestrian crashes, with weather conditions, lighting, and road classifications playing significant roles (Susilawati et al., 2023). Risk factors associated with crashes involving electric vehicles (EVs) include intersections, daytime, dry road conditions, clear weather, urban roads, traffic signals, and angular collisions (Weibull et al., 2023). Considering and understanding environmental conditions is crucial for optimizing the performance and safety benefits of ADAS, AEB, and autonomous vehicles.

2.6.3 Traffic Safety in General

Weather Conditions as a Contributing Factor to Traffic Crashes

The research conducted by Xu et al. (2018) revealed that unfavorable weather conditions contribute to increased crash risks on freeways in California, particularly when there are interactions between upstream occupancy and light rain. Several studies also examined the utilization of weather and significant variables for real-time crash prediction (Rongjie Yu et al., 2013; Yu et al., 2015; Yu and Abdel-Aty, 2014b; W. Yu et al., 2019). Wen et al., (2019) developed a Bayesian spatiotemporal model that identified associations between crash frequency and risk factors such as curves, slopes, traffic composition, and weather conditions. Strong et al. (2010) discussed the existing gaps in weather and transportation research and emphasized the importance of proactive safety management with up-to-date traffic data. By analyzing Pennsylvania crash data, Kelarestaghi et al. (2017) found that adverse weather conditions and young drivers contribute to a reduction in crash severity, while factors such as unbelted passengers, motorcycles, heavy trucks, and pedestrians increase crash severity. Theofilatos (2019) conducted a study in Athens, Greece, utilizing real-time traffic and weather data and discovered that the intensity of rainfall strongly influences the occurrence of crashes. Buddhavarapu et al. (2013) investigated crash severities in Texas, focusing on pavement surface conditions and horizontal curves, and established a significant impact of the Distress Index and International Roughness Index (IRI) on crash injury severity.

The study conducted by A. Das et al., (2017) aimed to gain insights into the impact of visibility on safety from a visibility perspective. The research findings identified several factors that were

significantly associated with safety, including curved roads, drivers of different age groups, roads with higher speed limits, traffic signalization, roads with low friction, undivided roads, and the absence of nighttime lighting. Weather conditions such as heavy precipitation, wet pavement, strong winds, frozen precipitation, reduced visibility, flooding, extreme temperatures, and other related factors affect the operations of commercial motor vehicles (CMV) and driver safety. This important subject, which had received limited attention, was explored by Rossetti and Johnsen, (2011).

When examining crashes involving multiple vehicles on high-speed roadways during rainfall, Jung et al. (2012) discovered noteworthy factors such as inadequate car following, wind speed, and actions of the driver at fault. Olowosegun et al. (2022) emphasized the impact of attributes related to the road, weather, and time on the severity of crashes, particularly on slippery road surfaces. Sun et al. (2022) identified elements that influence the severity of bicycle-motor vehicle crashes in urban and suburban areas of Beijing, including the type of vehicle, signal control, and lighting conditions. These investigations highlight the importance of diverse factors, such as driver conduct, road conditions, and environmental elements, in comprehending and addressing the seriousness of crashes during adverse weather conditions.

Cafiso et al. (2021) established indicators to assess the quality of pavement surfaces and geometric design using SPFs. By applying SPFs to analyze the impact of specific factors on the occurrence of crashes, they derived CMFs using the coefficients obtained from the models. The study focused on developing CMFs for various criteria, including Grip Number, International Roughness Index, curvature change ratio, curvature coefficient of variation, maximum superelevation deficit, and minimum lane width. Generalized linear modeling techniques were employed, assuming a negative binomial distribution error structure, to fit the models. Equation (36) in the study provides detailed information on the selected model formulation.

$$\hat{E}(Y) = L \times e^{a_0 + a_1 \ln(AADT)} \times e^{\sum_{i=1}^{n} \beta_i x_i} \tag{36}$$

Where: $\hat{E}(Y)$ is the projected annual crash frequency, L is the length of the segment (m), $AADT$ is the segment average annual daily traffic (veh/day), a_0, a_1 and β_i are the model parameters, and x_i are the explanatory variables. The distribution of the crash frequency around $\hat{E}(Y) = u$ is a negative binomial with variance described using Equation (37).

$$Var(Y) = u + k \times u^2 \tag{37}$$

Where k is the dispersion parameter of the negative binomial distribution.
The model can be rewritten similarly to the HSM form:

$$\hat{E}(Y) = e^{a_{0(base)}} \times L \times AADT^{a_1} \times (CMF_1 \times \ldots \times CMF_k) \tag{38}$$

Where CMF_i are the CMFs for the base conditions, derived from the SPF coefficients:

$$CMF_i = e^{\beta_i(x_i - x_{i(base)})} \quad (39)$$

And $a_{0(base)}$ is the constant coefficient of estimated SPF, adjusted for the base conditions:

$$a_{0(base)} = e^{a_0 + \sum \beta_i x_{i(base)}} \quad (40)$$

The baseline conditions were established using reference values for the covariates. The model parameters, as well as the dispersion parameter for the negative binomial distributions, were estimated using the maximum likelihood technique.
Table 24 displays the CMFs for the indicators of pavement condition.

Table 24. CMFs for Pavement Condition Indicators (Cafiso et al., 2021).

Crash type	B	X_{base}	X_{min}	X_{max}	CMF X_{min}	CMF X_{max}
Grip Number – GN						
Total crashes	-2.148 (-2.840, -1.456)	O.45	0.25	0.65	1.54	0.65
Run-off-the road	-3.054 (-4.534, -1.574)	O.45	0.25	0.65	1.84	0.54
Other	-1.971 (-2.680, -1.263)	O.45	0.25	0.65	1.48	0.67
Dry	-2.048 (-2.780, -1.315)	O.45	0.25	0.65	1.51	0.66
Wet	-3.508 (-5.035, -1.981)	O.45	0.25	0.65	2.02	0.50
Daytime	-2.025 (-2.780, -1.270)	O.45	0.25	0.65	1.50	0.67
Nighttime	-2.628 (3.796, -1.460)	O.45	0.25	0.65	1.69	0.59
International Roughness Index – IRI						
Total crashes	0.073 (0.021, 0.124)	4.00	2.52	7.96	0.90	1.33
Run-off-the road	0.109 (0.013, 0.205)	4.00	2.52	7.96	0.85	1.54
Other	0.059 (0.008, 0.111)	4.00	2.52	7.96	0.92	1.26
Dry	0.088 (0.035, 0.140)	4.00	2.52	7.96	0.88	1.41
Daytime	0.058 (0.002, 0.115)	4.00	2.52	7.96	0.92	1.26
Nighttime	0.085 (0.011, 0.159)	4.00	2.52	7.96	0.88	1.40

Weather Conditions as the Sole Study Factor to Traffic Crashes

Norrman et al. (2000) explored the relationships between road slipperiness, crash probability, and winter road maintenance frequency. The study's results indicated that the risk of a traffic collision varied depending on the type of road slipperiness. The most common cause of crashes was slippery conditions caused by rain or sleet on a frozen road surface. Najafi et al. (2015) developed regression models using data on collisions and pavement conditions in New Jersey to examine how friction impacts the frequency of vehicle crashes under wet and dry conditions in different urban settings. The findings revealed that friction affects the frequency of vehicle collisions in both dry and wet conditions. Abohassan et al. (2022) investigated the impact of

altering pavement friction levels on traffic safety during snowstorms in urban areas, focusing on Edmonton, the capital of Alberta. The results showed a highly significant association between pavement friction and traffic safety. When the pavement friction exceeded 0.6, a significant reduction in collisions was expected, while pavement friction below 0.35 was predicted to result in a significant increase in collisions. Additionally, the study found that arterial roads had a substantially higher crash rate compared to collector roads.

El-Basyouny et al. (2014a) established 12 weather conditions based on temperature, snow, rain, and wind speed and developed multivariate safety models using 11 years of daily weather and crash data from Edmonton, Alberta, Canada. The results revealed that severe collisions (resulting in injury or death) had a greater impact on PDO crashes compared to adverse weather conditions, resulting in a 4.5% to 45% increase in PDO crashes. All types of crashes showed statistical significance and a strong correlation with sudden weather changes that led to heavy rain or snow. El-Basyouny et al., (2014b) investigated the influence of meteorological conditions, particularly abrupt and intense snowfall or rainfall, on the type of crashes. The findings demonstrated that snowfall and temperature had statistically significant associations with all collision categories, indicating that as snowfall intensity increased, crashes increased, and as temperature increased, crashes decreased. Rainfall, on the other hand, had minimal impact. Gim (2022) analyzed demographic data from nationwide road traffic crashes that occurred between 2011 and 2015 to examine the severity of injuries in collisions involving senior drivers (aged 65 or older). The findings showed that the peak picnic season in October had the greatest impact on an increase in severity, but this magnitude was observed primarily in minor or less serious collisions. Additionally, in January, when road conditions were poor, minor or less serious collisions escalated into serious ones. Notably, between April and September, when the weather was favorable, the severity of fatal injuries decreased.

Examining the correlation between weather conditions and crash risk, Bergel-Hayat et al. (2013) observed significant associations on a monthly basis, varying depending on the type of road. Poor road weather conditions, such as icy rain and slippery roads, were identified as contributing to higher crash risks, with motorways exhibiting a particularly elevated risk in adverse weather and road conditions (Malin et al., 2019). Qin et al. (2006) conducted a study on the impact of snowstorms on road safety, revealing that both the severity of crashes and winter maintenance efforts played a role in crash rates and casualties. Implementing proactive winter maintenance measures was found to have a significant positive effect on traffic safety. These studies underscore the importance of considering weather conditions and implementing effective maintenance strategies to mitigate crash risks across different types of roads.

In analyzing weather-related crashes, Khan et al. (2008) identified spatial patterns and significant clusters based on different weather conditions. Jackson and Sharif (2016) studied rain-related fatal crashes in Texas, highlighting counties with potential contributing factors. Brijs et al.

(2008) used an autoregressive model to examine the impact of weather conditions on crash counts in the Netherlands. Assumptions related to weather effects were found to be significant. Lee et al. (2018) utilized Structural Equation Modeling (SEM) to analyze the relationship between water depth, rainfall, and traffic crashes in Seoul, Korea. These studies employed statistical techniques to understand the influence of weather on crash patterns and provided insights into specific regions and factors to focus on for further analysis and research.

Jin et al. (2014) proposed a method for spatial optimization aimed at identifying optimal locations for weather station deployment within an extensive regional transportation network. By utilizing crash data associated with weather conditions, a safety concern index was developed. This information was then used to analyze routes that offer comprehensive spatial coverage of the area, enabling the identification of the most suitable locations for weather station placement through a maximizing technique. In another approach, Lee et al. (2022) explored the use of Global Navigation Satellite Systems signals to assess the safety of urban roadways for navigation purposes. The aim was to enable automated vehicles to plan their trajectory while avoiding hazardous road segments affected by adverse weather conditions, thereby ensuring safe operations. This decision-making process relied on the analysis of physical route characteristics, vehicle capabilities, and weather conditions.

Table 25. Studies on Weather and Safety.

Study	Modeling Approach	Impact of Weather on Safety
Qin et al. (2006)	Macroscopic analysis	▪ The severity of snowstorms, considering factors such as duration, intensity, and wind speed, leads to an increase in traffic crashes and casualties.
Brijs et al. (2008)	Integer autoregressive model	▪ Results suggest that serial temporal correlation can account bias reduction
Khan et al. (2008)	Spatial analysis	▪ Applied spatial statistical techniques ▪ Identified notable patterns of weather-related crashes
Bijleveld et al. (2009)	Aggregate level analysis	▪ Performed an analysis of the aggregate effect of weather conditions on crashes in the Netherlands
Andrey (2010)	Matched-pair framework	▪ Over time, there is no significant alteration in the relative risk of casualties during snowfall.
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 commercial motor vehicles (CMVs) on our Nation’s highways

Study	Modeling Approach	Impact of Weather on Safety
Jung et al. (2012)	Sequential logistic regression	▪ Examined effect of contributing factors on the severities of multivehicle-involved crashes (during inclement weather) on high-speed roadways of Wisconsin ▪ Used a sequential logistic regression approach to perform analysis
Khan et al. (2012)	Spatial analysis	▪ Conducted spatial statistical techniques to identify substantial patterns of weather-related crashes.
Buddhavarapu et al. (2013)	Ordered probit (OP) response model	▪ On two-lane horizontal bends, skid number was not significantly connected with collision injury severity. ▪ IRI and the Distress Index were shown to have a statistically significant impact on injury severity.
Bergel-Hayat et al. (2013)	Time series model	▪ Examined the link between weather conditions and crash risk at an aggregate level (on a temporal basis)
El-Basyouny et al. (2014a)	Full Bayesian (FB) context via a Markov chain Monte Carlo simulation	▪ Examined the aggregated impact of adverse weather on crash.
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 roadway weather information systems (RWIS)
Theofilatos and Yannis (2014)	Linear regression	▪ Provided a review of the effect of traffic and weather characteristics on road safety ▪ Identified the gaps and discuss the need for further research
Najafi et al. (2015)	Regression analysis	▪ Friction affects not only the frequency of vehicle crashes in wet conditions but also in dry conditions.
Das et al. (2017)	Parametric model, MCA, topic model	▪ Examined the implications of inclement weather on safety from the perspective of visibility and other key issues.
Kelarestaghi et al. (2017)	Spearman correlation test	▪ adverse weather conditions and the presence of young drivers tend to reduce the severity of crashes.
Lee et al. (2018)	SEM	▪ Conducted a systematic approach to analyze weather crash relations using data from Seoul, Korea.
Xu et al. (2018)	Logistic regression models	▪ Environmental information improved the crash risk prediction model's fit and prediction performance.
Malin et al. (2019)	Concept of random point process	▪ Investigated the relative crash risk of different road weather conditions and combinations of conditions.

Study	Modeling Approach	Impact of Weather on Safety
Theofilatos et al. (2019)	Cusp catastrophe theory and NB model	▪ Results show that rainfall is linearly associated with crashes ▪ Average flow shows a non-linear relationship with crashes
Wen et al. (2019)	Bayesian spatiotemporal model	▪ Measured the association 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 effect of real-time traffic and weather on crashes.
Gim (2022)	Conventional ordered logit model	▪ Poor road conditions in January increase the severity of minor crashes. In contrast, favorable weather conditions from April to September reduce the severity of fatal injuries.
Abohassan et al. (2022)	Negative binominal techniques	▪ Collisions decrease with pavement friction above 0.6, but increase with pavement friction below 0.35.
Asli (2022)	Friedman test	▪ Weather conditions were found to be the first factor in pedestrian crashes in urban roundabouts.
Ferreira et al. (2022)	Literature Review	▪ Bad weather conditions negatively affected the users' safety perception.
Masello et al. (2022)	Experiments	▪ The efficiency of ADAS is significantly dropped during bad weather conditions
Norrman et al. (2022)	Data statistics	▪ The greatest risk of crashes occurred when the road surface was slippery due to rain or sleet freezing.
Olowosegun et al. (2022)	Ordered probit models	▪ Weather was found to affect the severity of crash injuries.
Qiu and Fan (2022)	Logistic regression	▪ Severe weather only has impacts at non-intersections.
Sun et al. (2022)	A two-stage approach integrating random parameters logit model and two-step clustering algorithm	▪ Weather had statistically significant random effects on the injury severity in urban areas
Susilawati et al. (2022)	A two-level Bayesian Poisson lognormal model	▪ Weather had significant effects on vehicle–pedestrian crashes in all road classifications.
Weibull et al. (2023)	Literature review	▪ Weather is one of the most frequently reported risk factors
Cerný et al. (2023)	Experiments	▪ The efficiency of AEB was significantly influenced by both vehicle speed and weather conditions.

2.7 SUMMARY

This chapter summarizes the findings of Task 2:

- Examined the SPFs in the HSM and other innovative statistical and machine learning models.
- Addressed the existing limitation in traditional safety calculations by incorporating the safety of non-motorists, such as pedestrians and cyclists, into the overall safety calculations. By considering the specific vulnerabilities and risks faced by non-motorists, the findings ensure a more comprehensive and inclusive approach to safety planning and decision-making processes. This inclusion facilitates the identification of areas of concern and enable the implementation of targeted measures to enhance the safety of all road users.
- Provided a comprehensive understanding of safe speed within the context of the SSA. It combines theoretical foundations with empirical evidence, identifying key factors and strategies that contribute to maintaining appropriate speeds and reducing crash risks.
- Investigated the complex relationship between various factors, including operating speed, geometric variables, rainfall, posted speed limits, and crash outcomes. By conducting a thorough analysis, the memorandum provides valuable insights into the interplay between these variables and their impact on urban roadway crash risk.

Through the synthesis of gathered information, knowledge, and practices, this chapter provides valuable insights. The findings contribute to the development of improved SPFs, the inclusion of non-motorist safety considerations, and a better understanding of the relationship between key factors and crash outcomes on urban roadways.

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

The Project Team identified four major data sources to perform the analysis. These data sources are:

- Short duration (10-minute interval) operating speed data from National Performance Management Research Data Set (NPMRDS) and INRIX XD.
- Traffic crash data from CRIS.
- Roadway inventory data from the Road-Highway Inventory Network Offload (RHiNO).
- Weather data from Copernicus, the European Union's Earth Observation Program.

3.2.1 Speed Data: NPMRDS and INRIX XD

Three readily accessible options exist for capturing speed information on Texas roadways: NPMRDS, the recently released Performance Network from the FHWA, and the INRIX XD network (Federal Highway Administration, 2019). The Project Team currently has a contract with INRIX to obtain travel time data on its XD network, which is conflated onto the RHiNO network. NPMRDS, procured by the FHWA, is free to state departments of transportation and metropolitan planning organizations for research.

3.2.2 Crash Data: CRIS (2018-2022)

The Project Team collected 5 years (2018-2022) of crash data from TxDOT's CRIS. CRIS data elements are divided into three major groups: (a) crash event characteristics, (b) primary person characteristics, and (c) vehicle (unit) characteristics.

3.2.3 Roadway Inventory Data: RHiNO

The Project Team acquired roadway inventory data from two different sources: (a) 2021 RHiNO, and (b) 2021 TxDOT Roadway Inventory (Texas Department of Transportation, 2024, 2025a). An examination of these data sets showed that they are the same. RHiNO provides a detailed data dictionary and additional supporting GIS files, the Project Team used the 2021 RHiNO as the main layer on which the other data layers were conflated.

3.2.4 Weather Data: Copernicus

Copernicus, the Earth Observation Program of the European Union, serves the collective interests of citizens by observing and analyzing the Earth's environment. Its primary goal is to

offer unrestricted access to satellite Earth observations, in situ data, and modeling information. Coordinated by the European Commission, Copernicus operates through agreements with international organizations. The European Centre for Medium Range Weather Forecasts (ECMWF) oversees the implementation of Copernicus Climate Change Service (C3S) and Copernicus Atmospheric Monitoring Service (CAMS) on behalf of the European Union. The data collection was conducted for 4 years (2019-2022).

3.3 DATA PREPARATION FRAMEWORK

The Project Team has followed an extensive and replicable data preparation and conflation framework to develop the data product (P1) for RTI Project 0-7144. This document provides detailed information on the implementation of this replicable methodological procedure. These sources encompassed road inventory data, which was obtained through the RHiNO system, providing detailed information about the urban freeway infrastructure. For an in-depth understanding of traffic patterns, granular operating speed data was derived from the NPMRDS and INRIX XD, offering 10-minute interval operating speed data. The Project Team further gathered the weather data. Finally, crash data was collected from CRIS, enabling an analysis of safety factors and their impact on urban roadways. The databases were organized into two primary temporal clusters: 1) Annual level, and 2) Short-duration level. Figure 17 provides a flowchart of the data conflation procedure.

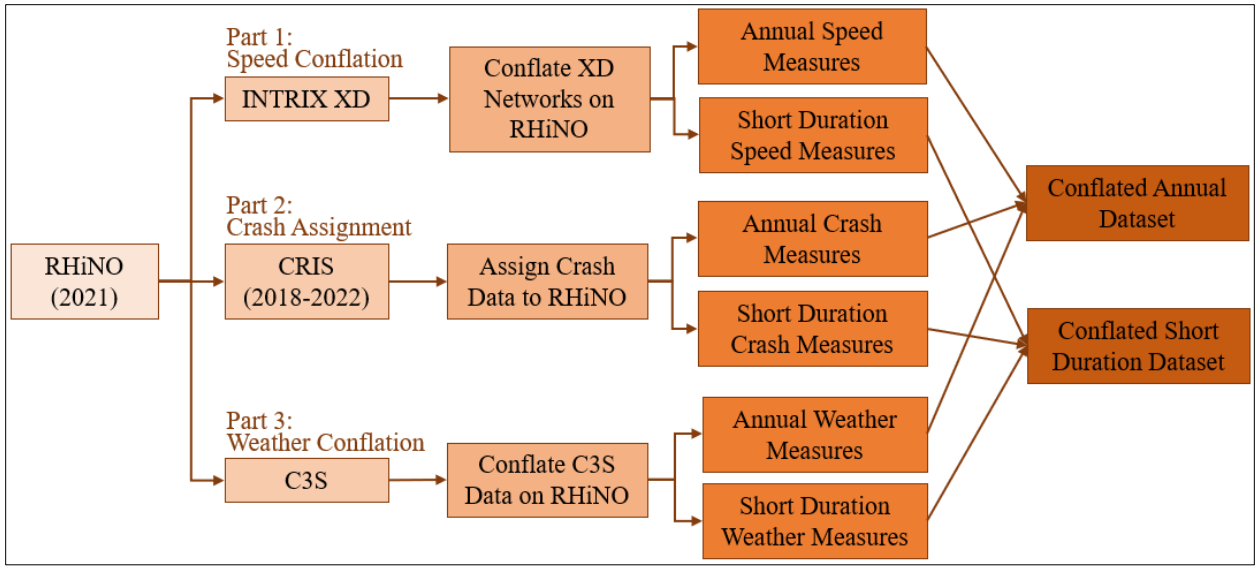


Figure 17. Data Conflation on Texas Urban Roadways.

3.3.1 RHiNO Segments

The Project Team first collected the RHiNO dataset for the year 2021. This dataset encompasses a total of 883,837 roadway segments, divided into 341,843 rural and 541, 994 urban segments. The focus of 0-7144 Project is on the urban network, which the Project Team further categorized

based on functional classification into two groups: major urban roads with 173,913 segments, and minor urban roads with 368,081 segments (see Table 26).

Table 26. Statistics of TxDOT Urban Facilities.

Facility Code	Facility Name	Number of RHiNO Segments	Total Lenth (mi.)
U1	Urban Interstate	23,832	5,186.3
U2	Urban Other Freeway and Expressway	19,340	4,741.7
U3	Urban Other Principal Arterial	46,582	9,403.8
U4	Urban Minor Arterial	32,683	9,144.1
U5	Urban Major Collector	51,476	16,231.7
U6	Urban Minor Collector	2,828	1,132.8
U7	Urban Local	365,253	80,122.2
	Total	541,994	125,962.6
U1-U5	Major Urban	173,913	44,707.5
U6-U7	Minor Urban	368,081	81,255.1

The data preparation was conducted by using R and Python programming languages and ArcGIS Pro software. Figure 18 illustrates the R code chunk used to extract the major and minor urban roadway segments from the overall dataset. In the code, "Urban1_5" indicates major urban roads, and "Urban6_7" indicates minor urban roads.

```r
# Required packages for script.
require(data.table)
require(dplyr)
require(sf)
require(tidyverse)
require(tigris)

# Sets the working directory to the same folder as the R script is located in.
setwd("C:/Users/j_l848/OneDrive - Texas State University/RTI_07144/3 Task 3/Data/RHiNO/")

rh <- st_read("2021/TxDOT_Roadway_Linework_wAssets.shp")

#Urban Subset
subset_df <- rh %>%
  subset(RU %in% c(2, 3, 4))

#Urban fun_sys subset
Urban1_5 <- subset_df %>%
  subset(F_SYSTEM %in% c(1,2, 3, 4,5))

#Urban fun_sys subset
Urban6_7 <- subset_df %>%
  subset(F_SYSTEM %in% c(6,7))
```

Figure 18. R Code for Selecting Urban Roads.

Upon a thorough examination of all data attributes within the accessible RHiNO dataset, the Project Team pinpointed six specific attributes capable of aiding in the identification of various urban facilities according to the definitions outlined in the HSM . Table 27 provides more details of these attributes.

Table 27. Attributes Relevant to Urban Roadway Selection.

Item Name	Column Name	Definition
RURAL-URBAN-CODE	RU	1=Rural (Population < 5,000) 2=Small Urban (Population 5,000 – 49,999) 3=Urbanized (Population 50,000 – 199,999) 4=Large Urbanized (Population 200,000+)
FUNCTIONAL-CLASSIFICATION *(Updated codes for YE2014)*	F_SYSTEM	1=Interstate 2=Other Freeway and Expressway 3=Other Principal Arterial 4=Minor Arterial 5=Major Collector 6=Minor Collector 7=Local
RECORD-TYPE (Pre YE2004: 8=HPMS Tolls) (Updated for YE2018 to remove 6=Functionally Classified City Street)	REC	0=Grade Separated Connector (YE2014 new code) On-System: 1=On-System Mainlanes 2=On-System Right Frontage Road 3=On-System Left Frontage Road Off-System: 5=County Road 7=City Street 8=Non-TxDOT Toll Authority Road (YE2014 new code) 9=Federal Road (YE2014 new code)
NUMBER-OF-THROUGH-LANES	NUM_LANES	Does not include turning, climbing, or auxiliary lanes, but does include Super 2 and exclusive HOV / HOT lanes
MEDIAN-TYPE	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
CLIMBING-PASSING-CENTERTURNING-LANE	CLMB_PS_LANE	1=Continuous Two-way Left Turn Lane 2=Super 2 Lane 3=Climbing / Passing Lane

In addition, the Project Team further divided road segments based on the facility types and supplied the facility code for each segment. “UI”: urban interstate; “UIF”: urban interstate freeway; “UFE”: urban others minor freeway; “UFEF”: urban other main freeways; “1U”: urban 1 lane undivided road; “2U”: urban 2 lanes undivided road; “2D”: urban 2 lanes divided road; “3T”: urban 3 lanes left turn road; “3U”: urban 3 lanes undivided road; “3D”: urban 3 lanes divided road; “4U”: urban 4 lanes undivided road; “4D”: urban 4 lanes divided road; “5T”: urban 5 lanes left turn road; “5U”: urban 5 lanes undivided road; “5D”: urban 5 lanes divided road; “6U”: urban 6 lanes undivided road; “6D”: urban 6 lanes divided road; “7T”: urban 7 lanes left turn road; “7U”: urban 7 lanes undivided road; “7D”: urban 7 lanes divided road; “8T”:

urban 8 lanes left turn road; “8U”: urban 8 lanes undivided road; “8D”: urban 8 lanes divided road; The R code, Figure 19, is used developed to cluster the data by these defined facility types.

```
# Create a new label column based on the Facility_Code and condition columns
subset_df_label <- Urban1_5 %>%
  mutate(Facility_Code = case_when(
    F_SYSTEM == 1 & (REC == 1|REC == 0| REC == 8) ~ "UI",
    F_SYSTEM == 1 & (REC == 2 | REC == 3|REC == 7) ~ "UIF",
    F_SYSTEM == 2 & (REC == 1|REC == 0|REC == 8) ~ "UFE",
    F_SYSTEM == 2 & (REC == 2 | REC == 3|REC == 7) ~ "UFEF",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 1  ~ "1U",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 2 & MED_TYPE == 0 & (CLMB_PS_LA == 0|CLMB_PS_LA == 3) ~ "2U",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 2 & (MED_TYPE == 2|MED_TYPE == 3|MED_TYPE == 4|MED_TYPE == 5|MED_TYPE == 6|MED_TYPE == 7) ~ "2D",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & (NUM_LANES == 2| NUM_LANES == 3) & CLMB_PS_LA == 1 ~ "3T",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 3 & MED_TYPE == 0 & (CLMB_PS_LA == 0|CLMB_PS_LA == 3) ~ "3U",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 3 & (MED_TYPE == 2|MED_TYPE == 3|MED_TYPE == 4|MED_TYPE == 5|MED_TYPE == 6|MED_TYPE == 7) ~ "3D",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 4 & MED_TYPE == 0 & (CLMB_PS_LA == 0|CLMB_PS_LA == 3) ~ "4U",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 4 & (MED_TYPE == 2|MED_TYPE == 3|MED_TYPE == 4|MED_TYPE == 5|MED_TYPE == 6|MED_TYPE == 7) ~ "4D",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & (NUM_LANES == 4|NUM_LANES == 5) & CLMB_PS_LA == 1 ~ "5T",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 5 & MED_TYPE == 0 & (CLMB_PS_LA == 0|CLMB_PS_LA == 3) ~ "5U",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 5 & (MED_TYPE == 2|MED_TYPE == 3|MED_TYPE == 4|MED_TYPE == 5|MED_TYPE == 6|MED_TYPE == 7) ~ "5D",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 6 & MED_TYPE == 0 & (CLMB_PS_LA == 0|CLMB_PS_LA == 3) ~ "6U",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 6 & (MED_TYPE == 2|MED_TYPE == 3|MED_TYPE == 4|MED_TYPE == 5|MED_TYPE == 6|MED_TYPE == 7) ~ "6D",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & (NUM_LANES == 6|NUM_LANES == 7)  & CLMB_PS_LA == 1 ~ "7T",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 7 & MED_TYPE == 0 & (CLMB_PS_LA == 0|CLMB_PS_LA == 3) ~ "7U",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 7 & (MED_TYPE == 2|MED_TYPE == 3|MED_TYPE == 4|MED_TYPE == 5|MED_TYPE == 6|MED_TYPE == 7) ~ "7D",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & (NUM_LANES == 7|NUM_LANES == 8)  & CLMB_PS_LA == 1 ~ "8T",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 8 & MED_TYPE == 0 & (CLMB_PS_LA == 0|CLMB_PS_LA == 3) ~ "8U",
    (F_SYSTEM == 3 | F_SYSTEM == 4 | F_SYSTEM == 5) & NUM_LANES == 8 & (MED_TYPE == 2|MED_TYPE == 3|MED_TYPE == 4|MED_TYPE == 5|MED_TYPE == 6|MED_TYPE == 7) ~ "8D",
    TRUE ~ "Other" ))
```

Figure 19. R Code for Dividing Roads by Facility Type.

As RHiNO is only limited to roadway data, the Project Team used other data sources to acquire information on operational speed, crash, and precipitation information. This document offers concise instructions on both annual and short-duration speed, crash, and precipitation summaries for RHiNO segments. Due to the data size and processing issue, annual and short duration data preparations were conducted for major (U1 to U5) and minor (U6 and U7) urban roadways separately.

3.4 ANNUAL DATA PREPARATION

3.4.1 Speed Measure Assignment

The original XD network of Texas provides operational speed information every ten seconds. To acquire the speed measures of RHiNO segments (conflating information from XD), the primary challenge of this conflation process lies in handling different segmentation and misalignment between the two networks. The core objective of conflation is to create an appropriate integration by establishing one-to-one or one-to-many relationships between XD segments and urban RHiNO roadways. This conflation process aims to harmonize and align data from the INRIX XD network with urban roadways, overcoming the challenges of segmentation and misalignment to provide a cohesive and accurate dataset for analysis. The main steps are outlined as follows:

Step 1: Preprocessing of RHiNO

In this initial step, a new field called ‘unique_id’ is introduced as a unique identifier for road segments within the existing layer. Additionally, the length of each segment is calculated and added to the ‘rhino_len’ field. The detailed code is shown in Figure 20.

```python
import arcpy
# Create a spatial reference object for the output coordinate system
out_coordinate_system = arcpy.SpatialReference('NAD 1983 UTM Zone 15N')

# Specify the full path to the input layer
rhino = r"C:\Users\j_l848\OneDrive - Texas State University\RTI_07144\3 Task 3\Data\2U_4D_4U_3T\Urban6_7\Urban6_7_Ori\Urban6_7.shp"

# Specify the full path to the output layer
rhino_proj = r"C:\Users\j_l848\OneDrive - Texas State University\RTI_07144\3 Task 3\Data\2U_4D_4U_3T\Urban6_7\Urban6_7_Project\Urban6_7_Project.shp"

# Run the tool
arcpy.Project_management(rhino, rhino_proj, out_coordinate_system)

################################Add unique_id to RHINO
# Add a new field for the unique ID
unique_id_field = "unique_id"
arcpy.AddField_management(rhino_proj, unique_id_field, "LONG")

# Populate the new field with the value "Urban67"
arcpy.CalculateField_management(rhino_proj, "Fclty_C", "'Urban67'", "PYTHON3")

# Calculate the unique ID field based on row numbers
with arcpy.da.UpdateCursor(rhino_proj, [unique_id_field]) as cursor:
    for i, row in enumerate(cursor):
        row[0] = i
        cursor.updateRow(row)

################################Length
# Add a new field for shape length
shape_length_field = "rhino_len"
arcpy.AddField_management(rhino_proj, shape_length_field, "DOUBLE")

# Calculate the shape length for each feature
arcpy.CalculateField_management(rhino_proj, shape_length_field, "!shape.length!", "PYTHON")
```

Figure 20. Python Code for Preprocessing of RHiNO.

Step 2: Creating Buffers Around RHiNO

This step involves the creation of a buffer layer around urban roadways. The buffer distance is set to “25 m or 82 ft”. The side type of the buffer is ‘FULL,’ and the end type of the buffer is “FLAT”. The buffer is generated to provide spatial context and facilitate subsequent conflation. Figure 21 presents the Python code of this step.

```python
# Specify the full path to the output layer
poly = r"C:\Users\j_l848\OneDrive - Texas State University\RTI_07144\3 Task 3\Data\2U_4D_4U_3T\Urban6_7\Urban6_7_buffer\Urban6_7_buffer.shp"

# Set the buffer distance (50 meters)
buffer_distance = "25 Meters"

# Set the buffer side type to "FULL" and end type to "FLAT"
side_type = "FULL"
end_type = "FLAT"

# Perform the buffer analysis
arcpy.Buffer_analysis(rhino_proj, poly, buffer_distance, side_type, end_type)
```

Figure 21. Python Code for Creating Buffers Around RHiNO.

Step 3: Establish the relationship between RHiNO and XD

For each buffered urban segment, the code, Figure 22, identifies any XD segments that intersect with it. During this process, the ‘Intersect_Length’ is calculated, representing the length of intersections between the polygon and XD segments. Furthermore, the ‘Percentage’ is computed as follows: Percentage = (Intersect_Length / Line_Length (XD)) * 100.

```python
import arcpy
# Set the input and output layer names
line_layer = r"C:\Users\j_l848\OneDrive - Texas State University\RTI_07144\3 Task 3\Data\2U_4D_4U_3T\XD_Project\XD_Project.shp"
polygon_layer = "C:\Users\j_l848\OneDrive - Texas State University\RTI_07144\3 Task 3\Data\2U_4D_4U_3T\Urban6_7\Urban6_7_buffer\Urban6_7_buffer.shp"
output_layer = "xd_Urban6_7_1"

# Create an empty feature class for the output layer
arcpy.management.CreateFeatureclass("in_memory", output_layer, "POLYLINE", spatial_reference=line_layer)

# Add fields for line length, intersect length, and percentage
arcpy.management.AddField(output_layer, "Line_Length", "DOUBLE")
arcpy.management.AddField(output_layer, "Intersect_Length", "DOUBLE")
arcpy.management.AddField(output_layer, "XDSegID", "text") # XDSegID
arcpy.management.AddField(output_layer, "Percentage", "DOUBLE")
arcpy.management.AddField(output_layer, "Polygon_ID", "LONG")
arcpy.management.AddField(output_layer, "RHINO_len", "DOUBLE")

# Open an insert cursor for the output layer
insert_cursor = arcpy.da.InsertCursor(output_layer, ["SHAPE@", "Line_Length", "Intersect_Length", "XDSegID", "Percentage", "Polygon_ID","RHINO_len"]) # XDSegID

# Create an insert cursor for the new layer
insert_cursor = arcpy.da.InsertCursor(output_layer, ["SHAPE@", "Line_Length", "Intersect_Length", "XDSegID", "Percentage", "Polygon_ID","RHINO_len"]) # XDSegID

# Iterate over each line feature
with arcpy.da.SearchCursor(line_layer, ["SHAPE@", "OID@", "SHAPE@LENGTH","XDSegID"]) as line_cursor: # XDSegID
    for line_row in line_cursor:
        line_geometry = line_row[0]
        line_length = line_row[2]
        XDSegID = line_row[3]

        # Flag to check if any intersections found
        intersect_found = False

        # Iterate over each polygon feature
        with arcpy.da.SearchCursor(polygon_layer, ["SHAPE@", "OID@","unique_id","rhino_len"]) as polygon_cursor:
            for polygon_row in polygon_cursor:
                polygon_geometry = polygon_row[0]

                # Check if line intersects with polygon
                if line_geometry.crosses(polygon_geometry):
                    intersect_found = True
                    intersect_geometry = line_geometry.intersect(polygon_geometry, 2)
                    intersect_length = intersect_geometry.length
                    percentage = (intersect_length / line_length) * 100

                    # Insert a new row with line and intersect information
                    insert_row = (
                        intersect_geometry,
                        line_length,
                        intersect_length,
                        XDSegID,
                        percentage,
                        polygon_row[2],
                        polygon_row[3]
                    )
                    insert_cursor.insertRow(insert_row)

        # If no intersections found, insert the line with null polygon information
        if not intersect_found:
            continue

# Clean up
del insert_cursor
```

Figure 22. Python Code for Establish the Relationship Between RHiNO and XD.

Step 4: Data Cleaning

The output generated in the previous step requires further refinement. It begins with the calculation of the sum of urban segments based on the XD line percentage. If the sum exceeds 100, corrective action is taken. Specifically, the percentages are sorted, and the one with the lowest percentage is eliminated until the sum is reduced to less than 100. Subsequently, the remaining percentages are adjusted proportionally to ensure that the cumulative percentage of each urban segment totals exactly 100. The detailed code is presented in Figure 23.

```python
import geopandas as gpd
import pandas as pd

# Read the shapefile and convert it to a DataFrame
data = gpd.read_file('C:/Users/j_l848/OneDrive - Texas State University/RTI_07144/3 Task 3/Data/2U_4D_4U_3T/5T/5T_xd/5T_conflated.shp')

df = pd.DataFrame(data)

df["Pct_poly"] = df["Intersect_"] *100/ df["RHINO_len"]

def process_data(df):
    # Group the data by "polygonID" and calculate the sum of "percentage"
    grouped_data = df.groupby("Polygon_ID")["Pct_poly"].sum()

    # Keep track of whether any changes have been made
    changes_made = False

    for polygon_id, total_percentage in grouped_data.items():
        if total_percentage > 100:
            polygon_data = df[df["Polygon_ID"] == polygon_id]
            if len(polygon_data) > 1:
                sorted_percentages = polygon_data.sort_values("Pct_poly", ascending=False)
                df = df.drop(sorted_percentages.tail(1).index)
                changes_made = True
            else:
                df.loc[polygon_data.index, "Pct_poly"] = 100
                changes_made = True

    return df, changes_made

# Repeat the process until no changes are made
changes = True
while changes:
    df, changes = process_data(df)

# Step 1: Group the data by "polygonID" and calculate the sum of "percentage"
grouped_data = df.groupby("Polygon_ID")["Pct_poly"].sum()

# Step 2-4: Check if the sum of "percentage" for a "polygonID" is less than 100 and apply the necessary operations
for polygon_id, total_percentage in grouped_data.items():
    if total_percentage < 100:
        polygon_data = df[df["Polygon_ID"] == polygon_id]
        if len(polygon_data) > 1:
            # Step 4b: Sort the "percentage" values in descending order
            sorted_percentages = polygon_data.sort_values("Pct_poly", ascending=False)
            # Step 4c: Increase the percentages proportionally to make the sum equal to 100
            proportions = sorted_percentages["Pct_poly"] / sorted_percentages["Pct_poly"].sum()
            adjusted_percentages = proportions * 100
            df.loc[sorted_percentages.index, "Pct_poly"] = adjusted_percentages
        else:
            # Step 4d: Set the single "percentage" value to 100
            df.loc[polygon_data.index, "Pct_poly"] = 100

# The resulting DataFrame "df" will have the desired modifications

df.to_csv("5T_xd.csv", index=False)
```

Figure 23. Python Code for Data Cleaning.

Step 5: XD Level Speed Summary

The result of Step 5 is a CSV file detailing the correlation between XD and RHiNO segments. This relationship is quantified by the overlap percentage of each RHiNO segment with corresponding XDs. The unique identifier for XD segments is labeled as ‘xd_id’, while for RHiNO segments, it's ‘unique_id’. Utilizing the ‘xd_id’ obtained from this step, raw XD speed data can be downloaded. Subsequently, a summary of XD-level speeds can be compiled using the Python code illustrated in Figure 24.

```
Sys.setlocale("LC_TIME", "English")
library(dplyr)
library(lubridate)
library(data.table)

setwd('C:/Users/dma147/Texas State University/Das, Subasish - Data/2U_4D_4U_3T/3T/Spd_Data_xd/3T_xd_18/')

data_ori1 = fread('3T_xd_18.csv')
#replace line 8.9
#########################################
tmc_list = fread('XD_Identification.csv',fill = TRUE)
tmc_list = unique(tmc_list$xd)
tmc_list = split(tmc_list, ceiling(seq_along(tmc_list)/30)) #set it according to your computer performance.
table = data.frame()
for (ls in tmc_list)
{
        data = data_ori1[data_ori1$xd_id %in% ls]
        data1 = data %>% group_by(xd_id) %>% summarize (SpdAve= mean(speed, na.rm=TRUE), SpdStd = sd(speed, na.rm=TRUE), Spd85 = quantile(speed, c(.85), na.rm=TRUE), PSL = first(reference_speed))

        data$timestamp = as.POSIXct(data$measurement_tstamp)
        data$year = year(data$timestamp)
        data$month = months(data$timestamp)
        data$weekday = weekdays(data$timestamp)
        data$hour = hour(data$timestamp)
        data_day = data[data$hour >5 & data$hour <18]
        data_day = data_day %>% group_by(xd_id) %>% summarize(SpdAveDay= mean(speed, na.rm=TRUE), SpdStdDay = sd(speed, na.rm=TRUE))
        data1 = merge(data1,data_day, by = "xd_id", all.x = TRUE, all.y = FALSE)
        data_night = data[data$hour<6 | data$hour >17]
        data_night = data_night %>% group_by(xd_id) %>% summarize(SpdAveNight= mean(speed, na.rm=TRUE), SpdStdNight = sd(speed, na.rm=TRUE))
        data1 = merge(data1,data_night, by = "xd_id", all.x = TRUE, all.y = FALSE)

        data_MTWT = data[data$weekday %in% c('Monday', 'Tuesday', 'Wednesday', 'Thursday')]
        data_MTWT = data_MTWT %>% group_by(xd_id) %>% summarize(SpdAveMTWT= mean(speed, na.rm=TRUE), SpdStdMTWT = sd(speed, na.rm=TRUE))
        data1 = merge(data1,data_MTWT, by = "xd_id", all.x = TRUE, all.y = FALSE)

        data_FSS = data[data$weekday %in% c('Friday', 'Saturday', 'Sunday')]
        data_FSS = data_FSS %>% group_by(xd_id) %>% summarize(SpdAveFSS= mean(speed, na.rm=TRUE), SpdStdFSS = sd(speed, na.rm=TRUE))
        data1 = merge(data1,data_FSS, by = "xd_id", all.x = TRUE, all.y = FALSE)

        data_lpsl = data[data$speed > data$reference_speed & !is.na(data$speed)]
        data_lpsl = data_lpsl %>% group_by(xd_id) %>% summarize(SpdFFAve= mean(speed, na.rm=TRUE), SpdFF85 = quantile(speed, c(.85), na.rm=TRUE))
        data1 = merge(data1,data_lpsl, by = "xd_id", all.x = TRUE, all.y = FALSE)

        table = rbind(table,data1)
}
###################################
#replace below
setwd('C:/Users/dma147/Texas State University/Das, Subasish - Data/2U_4D_4U_3T/3T/SpeedMeasures/')
write.csv(table,"data_summary_3T_XD_18.csv", row.names = FALSE) #please rename the output file accordingly.
```

Figure 24. R Code for XD Level Speed Summary.

Step 6: RHiNO Level Speed Summary

XD-level speed summary was aggregated to the RHiNO level speed summary based on the CSV file acquired from Step 4 (see Figure 25). The speed measures used in this study are listed Table 28.

```
Sys.setlocale("LC_TIME", "English")
require(foreign)
require(data.table)
require(dplyr)

TMC_N <- fread("C:/Users/dma147/Texas State University/Das, Subasish - Data/2U_4D_4U_3T/3T/xd_3T/xd_3T.csv")
TMC_N <- TMC_N[,c("Polygon_ID", "XDSegID", "Pct_poly")]

TMC_N$unique_id <- TMC_N$Polygon_ID

Spd_N <- fread("C:/Users/dma147/Texas State University/Das, Subasish - Data/2U_4D_4U_3T/3T/SpeedMeasures/data_summary_3T_XD_18.csv")

TMC_Spd_N <- left_join(TMC_N,Spd_N,by = c('XDSegID' = 'xd_id'))

TMC_Spd_N <- TMC_Spd_N[!is.na(TMC_Spd_N$SpdAve),]

TMC_Spd_N$tmc_perc <- TMC_Spd_N$Pct_poly/100

TMC_Spd_N$TMC_Ratio <- TMC_Spd_N$Pct_poly

TMC_Spd_N_Ave <- TMC_Spd_N %>% group_by(unique_id) %>% transmute(Spd18=weighted.mean(SpdAve,TMC_Ratio,na.rm = TRUE),
                                                                 SS18=sqrt(sum(SpdStd^2*tmc_perc^2)),
                                                                 SEF18=weighted.mean(Spd85,TMC_Ratio,na.rm = TRUE),
                                                                 PSL18=weighted.mean(PSL,TMC_Ratio,na.rm = TRUE),
                                                                 SpdD18=weighted.mean(SpdAveDay,TMC_Ratio,na.rm = TRUE),
                                                                 SSD18=sqrt(sum(SpdStdDay^2*tmc_perc^2)),
                                                                 SpdN18=weighted.mean(SpdAveNight,TMC_Ratio,na.rm = TRUE),
                                                                 SSN18=sqrt(sum(SpdStdNight^2*tmc_perc^2)),
                                                                 SWD18=weighted.mean(SpdAveMTWT,TMC_Ratio,na.rm = TRUE),
                                                                 SSWD18=sqrt(sum(SpdStdMTWT^2*tmc_perc^2)),
                                                                 SWE18=weighted.mean(SpdAveFSS,TMC_Ratio,na.rm = TRUE),
                                                                 SSWE18=sqrt(sum(SpdStdFSS^2*tmc_perc^2)),
                                                                 SFF18=weighted.mean(SpdFFAve,TMC_Ratio,na.rm = TRUE),
                                                                 SFEF18=weighted.mean(SpdFF85,TMC_Ratio,na.rm = TRUE)
                                                                 )
TMC_Spd_N_Ave <- unique(TMC_Spd_N_Ave)

#write.csv(TMC_Spd_N_Ave,file="C:/Users/dma147/Texas State University/Das, Subasish - Data/2U_4D_4U_3T/3T/Final_Spd/data_summary_3T_XD_18.csv",row.names = FALSE)
```

Figure 25. R Code for RHiNO Level Speed Summary.

Table 28. Speed Measure Variables and Definitions.

Attribute Name	Definition
SpdAve	Average speed determined for the year using all data
SpdStd	Standard deviation of speed determined for the year using all data
Spd85	85th percentile speed determined for the year using all data
PSL	Post Speed Limit (not real posted speed, a measure derived from INRIX average speed values)
SpdAveDay	Average speed determined for the year (hour > 5 and hour < 18) using all data
SpdStdDay	Standard deviation of speed determined for the year (hour > 5 and hour < 18) using all data
SpdAveNight	Average speed determined for year (hour > 17 and hour < 24 and hour > -1 and hour < 6) using all data
SpdStdNight	Standard deviation of speed determined for year (hour > 17 and hour < 24 and hour > -1 and hour < 6) using all data
SpdAveMTWT	Average speed determined for the year (Mon, Tue, Wed, Thu) using all data
SpdStdMTWT	Standard deviation of speed determined for the year (Mon, Tue, Wed, Thu) using all data
SpdAveFSS	Average speed determined for the year (Fri, Sat, Sun) using all data
SpdStdFSS	Standard deviation of speed determined for the year (Fri, Sat, Sun) using all data
SpdFFAve	Average speed determined for the year using speed data where 5-min speed is > PSL (or PSL+5 or +10)
SpdFF85	85th percentile speed determined for the year using speed data where 5-min speed is > PSL (or PSL+5 or +10)

3.4.2 Crash Data Assignment

This section provides an explanation of the RHiNO level annual crash summary. The Project Team downloaded Texas crash data (CRIS) from 2018 to 2022. The crash event files encompass various details, including the date and time of the crash, the severity of the crash, and the latitude and longitude coordinates corresponding to the crash events. The Project Team first displayed the crash points on the map using the "Display XY" tool. Then, the Project Team applied "Near" function tool in ArcGIS Pro (see Figure 26) to assign crash events to the nearest roadway segments. An attribute named 'Near_FID' was assigned to each crash, wherein 'Near_FID' signifies the FID number (the roadway segment's corresponding row number, commencing from 0 in the crash characteristics file) where the crash occurred. The Project Team chose 25 meters as the threshold to assign a crash event on the nearest roadway segment. This means only if a crash event is within 25 meters of a roadway segment, this crash event is it assigned to this specific roadway segment. If a crash event cannot be assigned to any roadway segment, its "Near_FID" attribute is be equal to -1. The Project Team filtered out the crash events whose "Near_FID" equals -1 because they cannot relate to any roadway segments. After crash events are assigned to roadway segments, the Project Team summarized the total number of crashes that happened on each roadway segment by crash severity. Five crash severity levels (KABCO) were used to

assign the crashes by severity levels. KABCO is an acronym used in the field of traffic safety and accident investigation to classify the severity of injuries resulting from motor vehicle crashes. Each letter in the acronym corresponds to a specific level of injury severity: K (Killed or Fatal), A (Apparent or Incapacitating Injury), B (Non-Incapacitating Injury), C (Possible Injury), and O (No Injury).

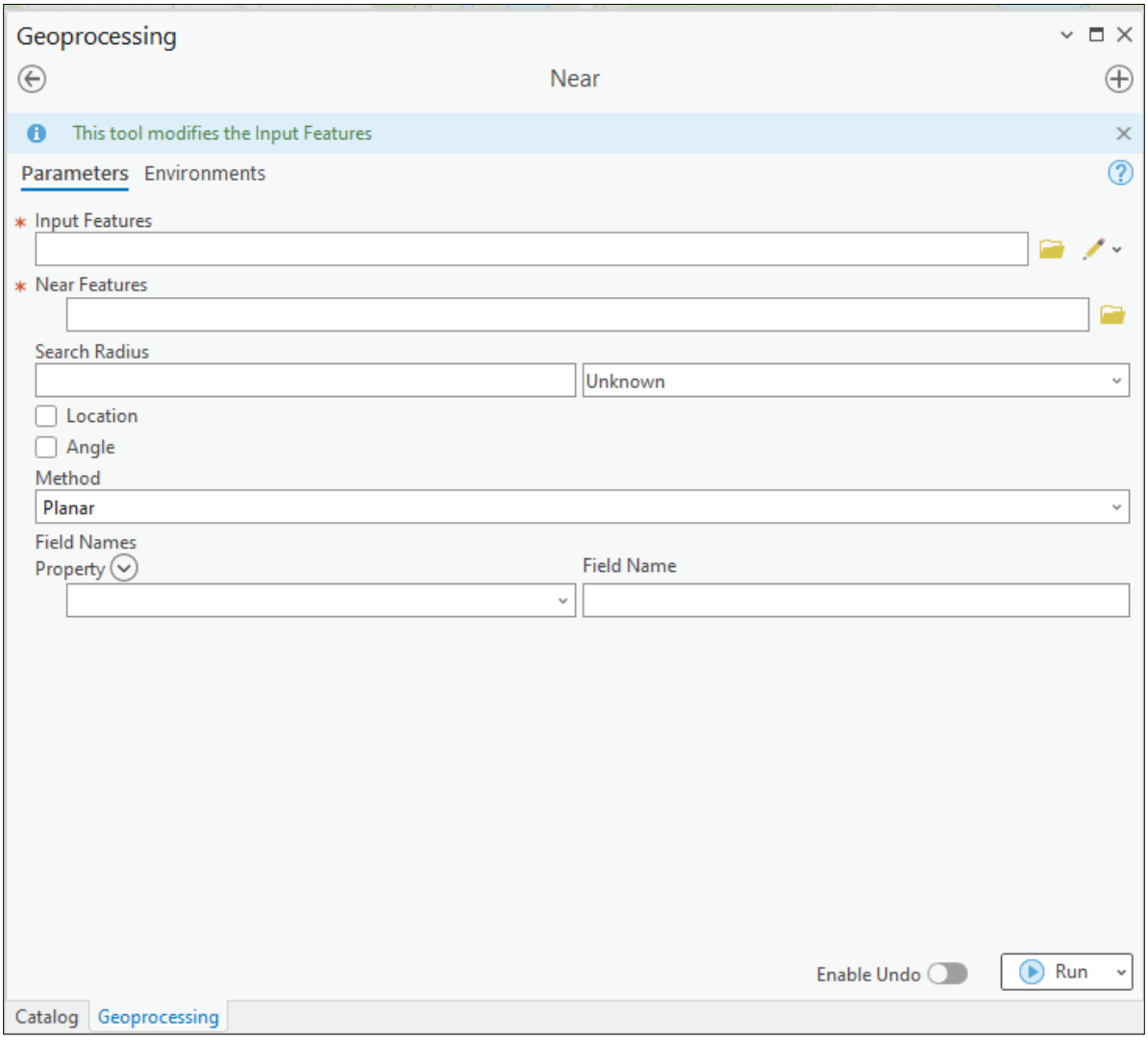


Figure 26. Near (Analysis) Tool in ArcGIS Pro.

3.4.3 Precipitation Measure Assignment

For short-duration precipitation data, the first step was to download hourly precipitation data, one month at a time, for an entire year, for 2019, 2020, 2021, and 2022 from the Copernicus Climate Data Store (CDS). To access the data, two prerequisites need to be followed: 1) A valid account on *cds.climate.copernicus.eu* with login credentials; 2) Necessary permissions to access the ERA5-Land hourly data product (see Figure 27). The steps are described below.

3.4.4 Steps to Download Data

Step 1: Log In and Navigate to the CDS: Access the CDS website and log in with your credentials. Once logged in, navigate to the 'Your Requests' section, as shown in Figure 28.
Step 2: Queue the Download Requests: Submit 12 separate requests, each corresponding to a different month of the year for each year. This should be done by selecting the ERA5-Land hourly data product and specifying the time range for each month. Ensure that all days for a month are selected, and all hours for each day. There is also a need to set the bounding box to the bounding box of the state of Texas as seen in Figure 28.
Step 3: Monitor Request Status: After submission, requests appear in the 'Queued' status. The system will automatically process these requests. The page will auto-refresh the status of the requests by default.

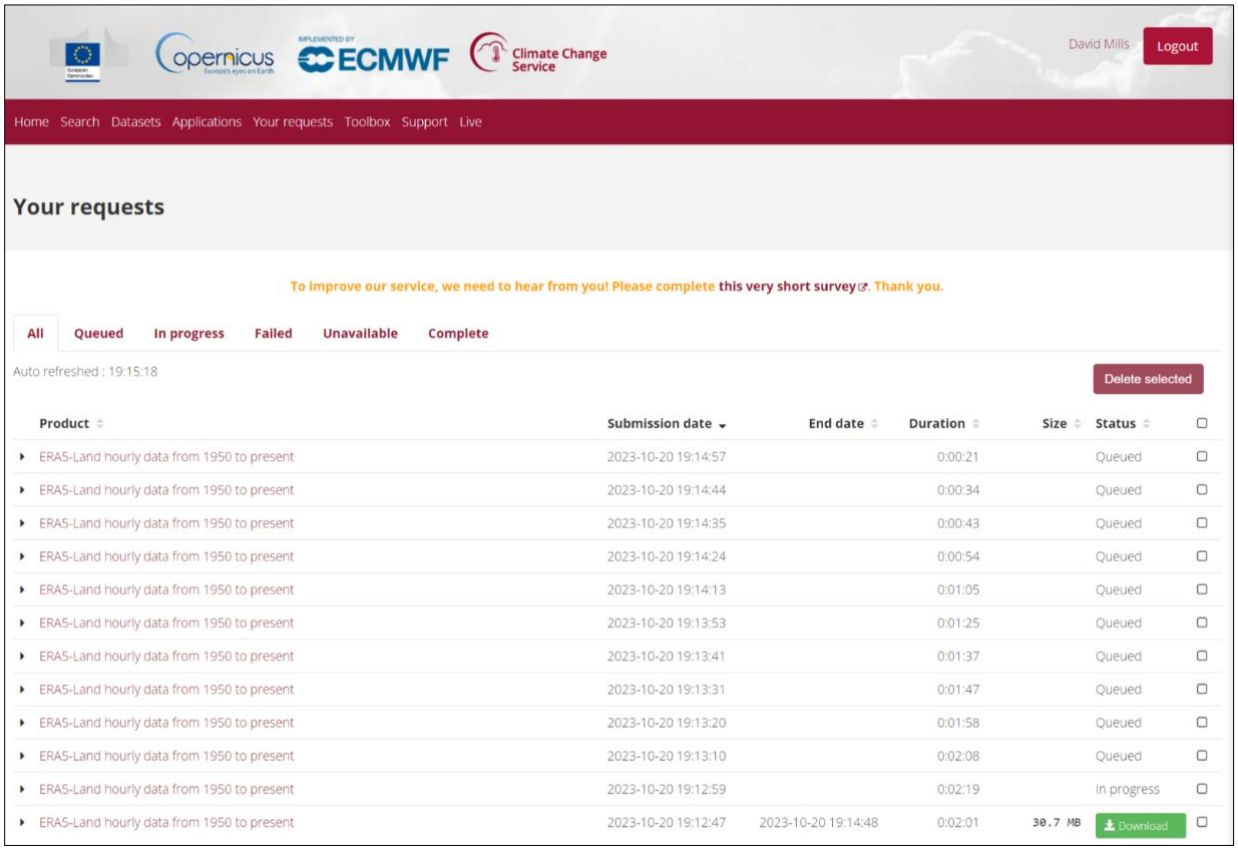


Figure 27. Precipitation Gridded Binary Data Downloaded from Copernicus.

Step 4: Downloading the Data: Once a request status changes to 'Complete,' a 'Download' button will appear. Click this button to download the dataset for the respective month. The file size will be indicated next to the button.
Step 5: File Storage and Management: As the data is downloaded, files are stored in an organized manner in OneDrive storage. A separate folder is created for each year.

Step 6: Data Verification: After the downloading is completed, the integrity of the data files was verified by loading data using the R package 'sp' to load an object of class 'SpatialGridDataFrame' for analysis. Each file corresponded to the monthly data requested, and files were labeled after verification according to the month and year of the data.
Step 7: Repeat for All Months: Repeat the download process for each month/year until all 12 monthly data files are successfully downloaded for each year.

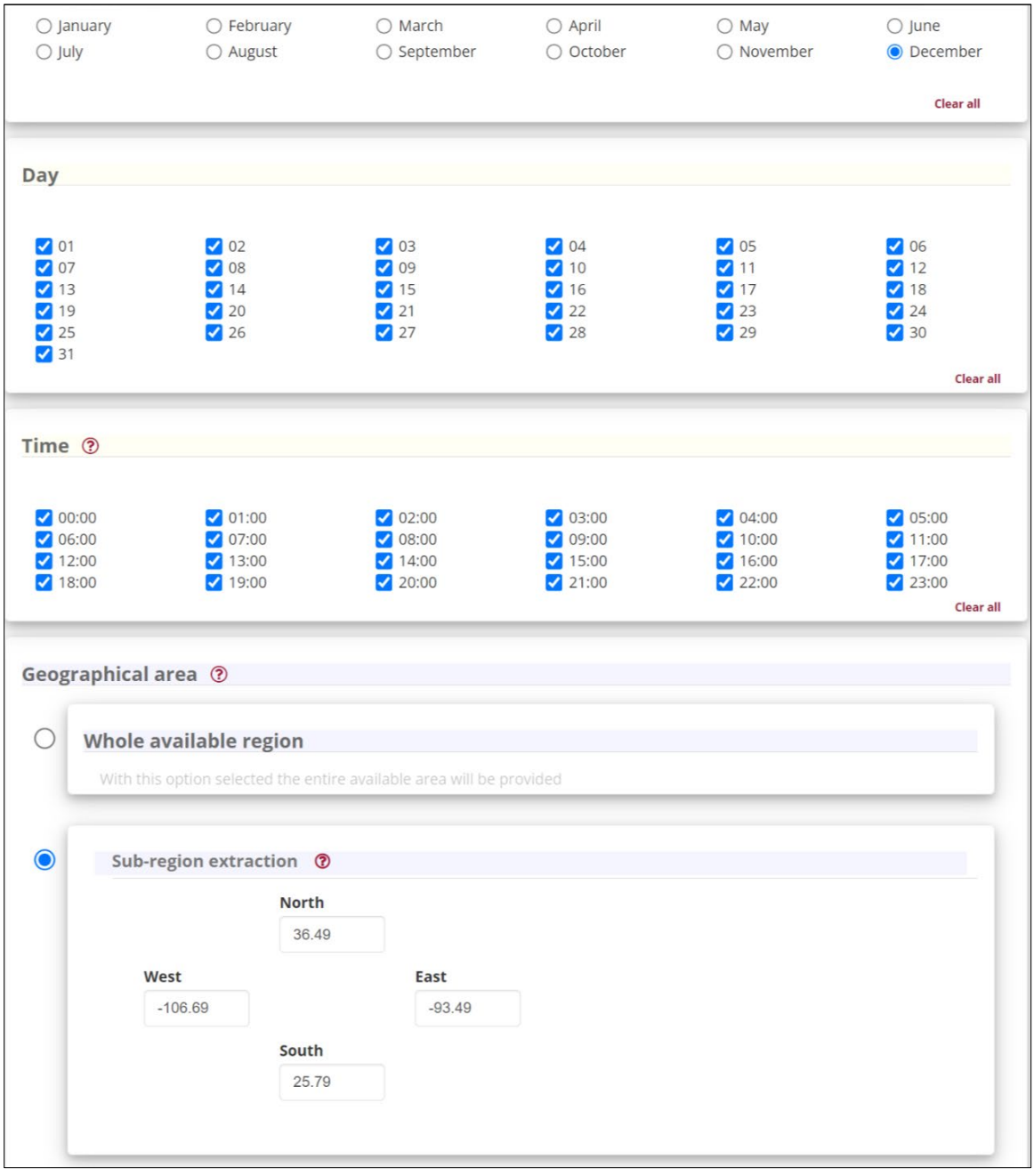


Figure 28. Data Requests Details.

3.4.5 Contents of the Precipitation GRIB Files

After the data download completion, there are 48 Gridded Binary (GRIB) files for the years 2019-2022, one for each month (see Figure 29). Inside the files are bands that total the number of hours in each of the downloaded months. For example, January of 2019 has 744 bands (31 Days * 24 hours = 744 bands). Each band is stored as a floating-point vector of length 14,124. Each band has a spatial attribute for 'xmin', 'ymin', 'xmax', and 'ymax' representing the grid area covered and spatial resolution. In Figure 30, 14,124 bands are seen as .1 x .1 degree resolution grid squares, with NA grids being represented as gray grid squares.

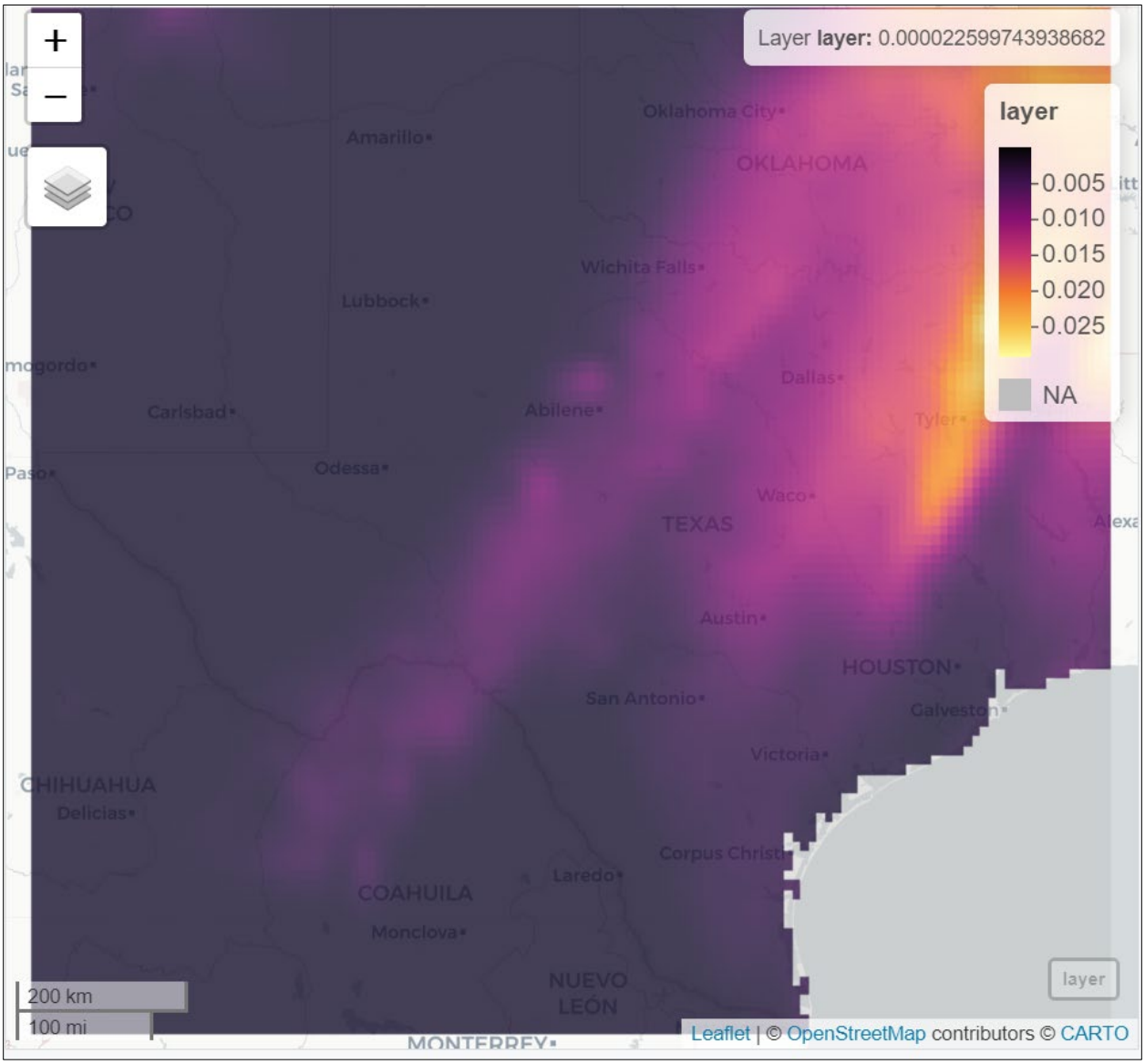


Figure 29. GRIB File for January 1st, 2019 at 0000-0059 Hours.

Converting GRIB files to CSV Files for Improved Processing Time

Converting GRIB files containing hourly weather data for one month into more manageable hourly CSV files is useful for improved processing time. GRIB files are commonly used in meteorology but can be cumbersome for certain types of data processing due to their size, file type, and spatial nature.

Step 1: Assessing GRIB File Size: The original GRIB file for January 2019 is approximately 29 megabytes, which is manageable but not optimal for rapid data processing.

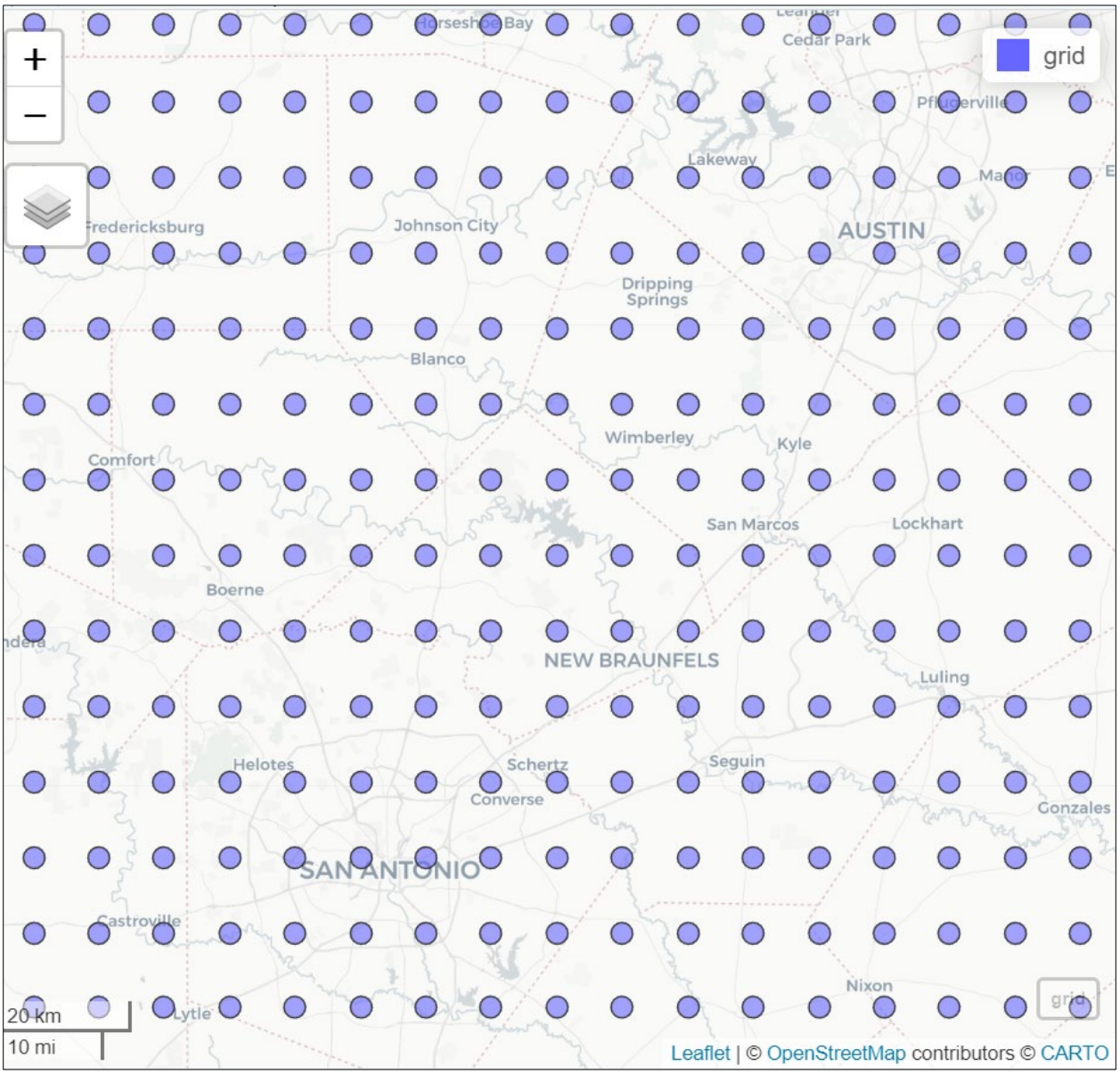


Figure 30. Point Shapefile Corresponding to GRIB Data.

Step 2: Generating Point Shapefile from GRIB. A point shapefile is created from the GRIB file, where each point represents the center of a grid cell in the spatial data. The shapefile contains 14,124 points, each corresponding to a grid cell in the GRIB file, a zoomed-up image of the point shapefile can be seen in Figure 30.

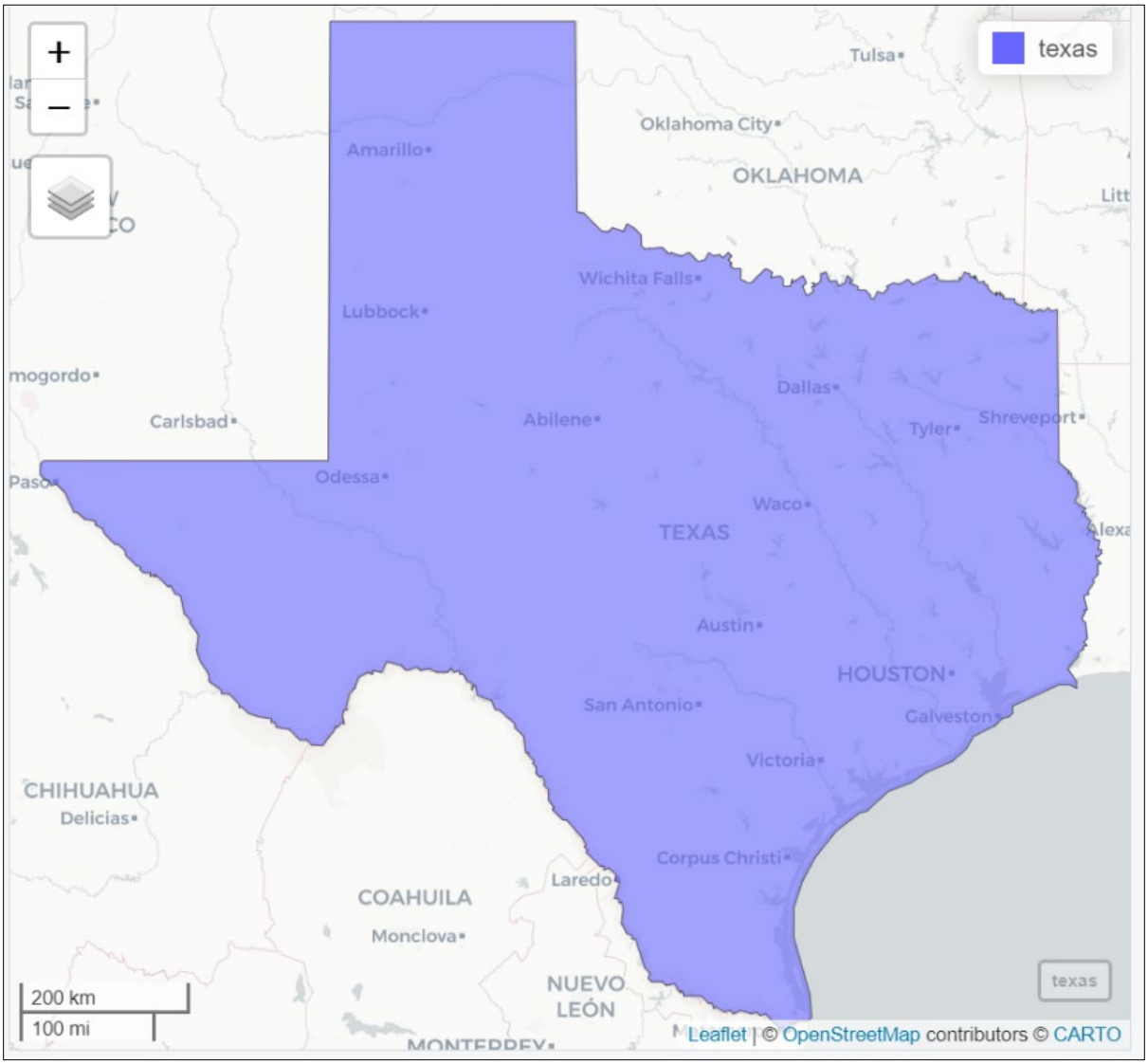


Figure 31. TIGRIS Shapefile of Texas.

Step 3: Clipping the Shapefile to the Area of Interest: To focus on a specific region, such as Texas, the shapefile is clipped using a Texas boundary shapefile obtained from the TIGRIS database, as seen in Figure 31. The clipping is performed in a common Coordinate Reference System (CRS) to maintain geographical accuracy.

Step 4: Reducing Data Points: After clipping, the number of data points is reduced to fit the Texas boundary, resulting in 6,607 points that fall within the state as seen in Figure 32. The 'NEAR_FID' column is maintained to match the original GRIB file index. A nearest neighbor

index is created from this shapefile to reference the GRIB data, which is stored in the CSV file for faster processing in the next step.

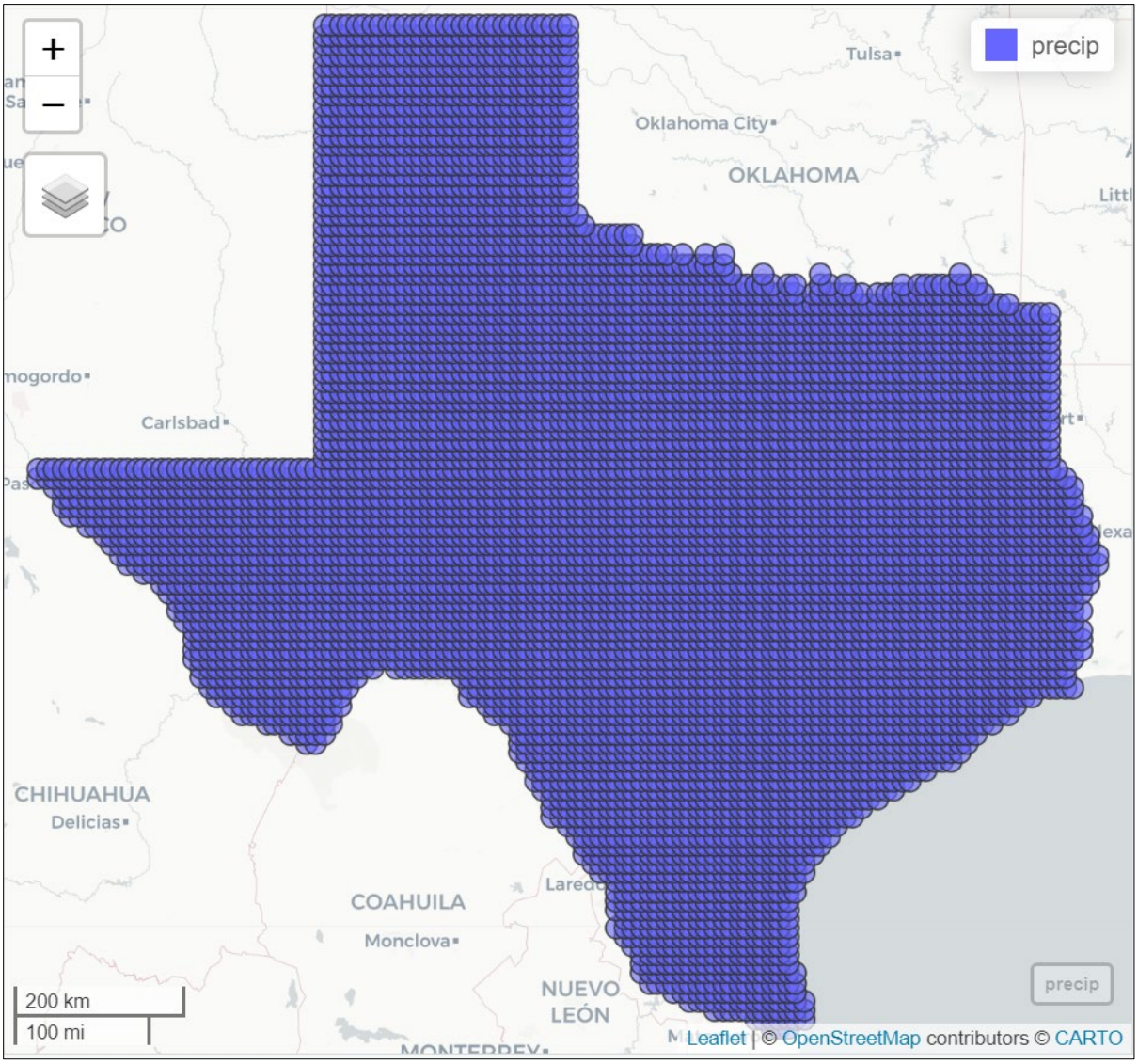


Figure 32 Precipitation Points Shapefile.

Step 5: Converting to CSV: The clipped data is then exported to CSV format, creating one file per hour, thus converting the monthly GRIB data into 744 individual hourly CSV files for the month of January 2019 and the respective number of hourly CSV files for the number of hours in each month afterward.

Step 6: Verifying CSV File Size: The size of each CSV file is approximately 82 KB, which is significantly smaller and more manageable than the original GRIB file. The advantages of conversion are described below:

- **Improved Data Processing Speed:** Smaller file sizes in CSV format allow for quicker data manipulation and analysis, especially when working with hourly data.
- **Targeted Dataset:** Trimming the shapefile to the specific region of interest eliminates irrelevant data, resulting in a more refined and focused dataset suitable for analysis.
- **Compatibility and Accessibility:** CSV files are widely supported by various software and programming languages, making it easier to share and work with the data across different platforms.
- **Ease of Use:** CSV files are easier to inspect and modify, allowing for a more straightforward data validation and quality control process.

3.4.6 Precipitation Data Processing

The Project Team has developed an R script to process and analyze extensive precipitation data, specifically focusing on hourly precipitation measures. The script enables us to transform granular precipitation data into structured insights.

- **Initialization of Analytical Tools:** The first step involves loading a suite of R packages. Several R packages were used in this process. The package 'lubridate' facilitates sophisticated date-time manipulation, 'dplyr' and 'tidyr 'are for data manipulation and tidying, and 'data.table' offers the capability to handle large datasets with speed and efficiency.
- **Workspace Configuration:** To maintain consistency in file paths and operations, the Project Team configured the working directory dynamically to the script's location. This step is bypassed on Shiny servers during application development due to their unique operational environment.
- **Numerical Representation:** The Project Team disabled R's default scientific notation to ensure that numerical data is readily interpretable without the necessity for transformation.
- **Data Aggregation Workflow:** The Project Team utilized a nested loop to iterate through years and months, systematically transforming hourly data into a structured statistical matrix. This approach enables efficient categorization and computation of aggregated statistics.
- **Spatial Data Integration:** In Step 4 of the Data Conversion section, the Project Team integrates each precipitation data point with its corresponding geographic identifier by utilizing the *id_match.csv* file, thereby incorporating the relevant spatial dimension into the analysis.
- **Data Exportation:** In the final stage of this process, the Project Team exported the amalgamated data into CSV files. These files serve as the endpoint of the analytical process, offering a comprehensive and easily accessible record for subsequent analysis, reporting, and peer review.

Utilizing the precipitation data processing script, the Project Team employs the capabilities of data science to undertake comprehensive and replicable investigations.

3.5 SHORT DURATION DATA PREPARATION

The previous section outlines the procedure for conflating annual speed, crash, and precipitation data. The current section details the short-duration aspects of these measures; the overall steps are very similar to the annual data preparation. The Project Team's goal is to obtain these measures by year, season (across all four seasons), day of the week (distinguishing between weekdays and weekends), and time of day (daytime and nighttime).

3.5.1 Speed Measure Assignment

In alignment with the annual speed summary procedure, the Project Team adapted and redeveloped the code for step 5 (Figure 33) and 6 (Figure 34) to obtain the short-duration speed summary. It's important to note that this process demands significant computational resources, often necessitating the use of high-performance computing to manage the extensive data processing involved.

```
import multiprocessing
import pandas as pd
import numpy as np
from datetime import datetime,timezone
import os

def process_tmc_chunk(data):
    # Dynamically determine the unique years for conditions
    unique_years = data['Year'].unique()

    # Update conditions to include all unique years
    conditions = [
        {"Year": year, "Season": ["Winter", "Spring", "Summer", "Fall", "All"],
        "DOW": ["Weekday", "Weekend", "All"], "TOD": ["Daytime", "Nighttime", "All"]}
        for year in unique_years
    ]

    # Placeholder for results
    results = []

    # Iterate over unique xd_id values
    for xd in data['xd_id'].unique():
        for condition in conditions:
            for season in condition["Season"]:
                for dow in condition["DOW"]:
                    for tod in condition["TOD"]:
                        # Filter data based on conditions and current xd_id
                        mask = (data['Year'] == condition["Year"]) & (data['xd_id'] == xd)
                        if season != "All":
                            mask &= (data['Season'] == season)
                        if dow != "All":
                            mask &= (data['Week'] == dow)
                        if tod != "All":
                            mask &= (data['TOD'] == tod)

                        filtered_data = data[mask]

                        # Compute SpdAvg and SpdStd
                        spd_avg = filtered_data['speed'].mean()
                        spd_std = filtered_data['speed'].std()
                        Spd_85 = filtered_data['speed'].quantile(0.85)

                        Spd_FF85 = filtered_data['SpdFF85'].mean(skipna=True)

                        # Calculate PSL rounded to the nearest 5
                        raw_PSL = filtered_data['reference_speed'].mean(skipna=True)
                        if pd.isna(raw_PSL):
                            PSL = pd.NA
                        else:
                            PSL = 5 * round(raw_PSL / 5)

                        TT_Avg = filtered_data['travel_time_seconds'].mean(skipna=True)
                        TT_Std = filtered_data['travel_time_seconds'].std(skipna=True)
                        PVTT = (filtered_data['travel_time_seconds'].std(skipna=True))*100/(filtered_data['travel_time_seconds'].mean(skipna=True))

                        results.append({"XD_id": xd, "Year": condition["Year"], "Season": season,
                                        "DOW": dow, "TOD": tod, "SpdAvg": spd_avg, "SpdStd": spd_std,'Spd85': Spd_85, 'SpdFF85': Spd_FF85, 'PSL': PSL, 'TTAvg':TT_Avg, 'TTStd': TT_Std,'PVTT': PVTT})
```

(a)

```python
if __name__ == '__main__':

        # get the path of the directory where the script is located
        dir_path = os.path.dirname(os.path.realpath(__file__))

        # define the relative file path to the data file
        file_path = os.path.join(dir_path, "2U", "2U_XD", "2U_xd_18", "2U_xd_18_4.csv")

        # read the CSV file using pandas
        data = pd.read_csv(file_path)_

        # Convert the measurement_tstamp to datetime
        data['datetime'] = pd.to_datetime(data['measurement_tstamp'])

        # Extract year, month, day of the week and hour
        data['Year'] = data['datetime'].dt.year
        data['Month'] = data['datetime'].dt.month
        data['DOW'] = data['datetime'].dt.dayofweek
        data['Hour'] = data['datetime'].dt.hour
        data['days'] = data['datetime'].dt.day

        # Assign seasons based on month
        data['Season'] = data['Month'].apply(lambda x: 'Winter' if x in [12, 1, 2] else
                                             ('Spring' if x in [3, 4, 5] else
                                             ('Summer' if x in [6, 7, 8] else 'Fall')))

        # Assign Daytime/Nighttime based on hour
        data['TOD'] = data['Hour'].apply(lambda x: 'Daytime' if 6 <= x <= 18 else 'Nighttime')

        # Assign Weekday/Weekend based on DOW
        data['Week'] = data['DOW'].apply(lambda x: 'Weekday' if x < 5 else 'Weekend')

        data['SpdFF85'] = data.groupby(['Month', 'days'])['speed'].transform(pd.Series.quantile, q=0.85)

        tmc_name = data['xd_id'].unique()

        chunk_size = 100  # adjust this to control the chunk size
        tmc_chunks = [tmc_name[i:i+chunk_size] for i in range(0, len(tmc_name), chunk_size)]

        # Determine the number of cores, subtracting 1 to reserve one core
        num_cores_to_use = multiprocessing.cpu_count() - 2

        # Determine the number of cores, subtracting 1 to reserve one core
        num_cores_to_use = multiprocessing.cpu_count() - 2

        with multiprocessing.Pool(processes=num_cores_to_use) as pool:

            list_of_filtered_a = [data[data['xd_id'].isin(tmc1)] for tmc1 in tmc_chunks]

            results = pool.map(process_tmc_chunk, list_of_filtered_a)
```

(b)

Figure 33. Python Code for XD Level Short Duration Speed Summary. (a) Python Code Part 1; (b) Python Code Continued.

```python
import pandas as pd
import numpy as np
from datetime import datetime,timezone
from tqdm import tqdm
import os
import multiprocessing

def process_tmc_chunk(unique_id_list):

    # get the path of the directory where the script is located
    dir_path = os.path.dirname(os.path.realpath(__file__))
    # define the relative file path to the data file
    file_path = os.path.join(dir_path, "xd_2U.csv")
    # read the CSV file using pandas
    TMC = pd.read_csv(file_path)
    TMC = TMC[["Polygon_ID", "XDSegID", "Pct_poly"]]
    # remove rows with missing values in 'SpdAve' column for TMC_Spd_N
    TMC = TMC.dropna(subset=['XDSegID'])
    # compute tmc_perc for TMC_Spd_N
    TMC['tmc_perc'] = TMC['Pct_poly']
    TMC['unique_id'] = TMC['Polygon_ID']
    TMC['tmc'] = TMC['XDSegID']

    # define the relative file path to the data file
    file_path2 = os.path.join(dir_path, "XD_level_spd_merged", "2U", "18.csv")

    # read the CSV file using pandas
    tb = pd.read_csv(file_path2)
    tb.rename(columns={'XD_id': 'tmc'}, inplace=True)

    season_name = np.unique(tb["Season"])
    time_name = np.unique(tb["TOD"])
    dow_name = np.unique(tb["DOW"])
    year_name = np.unique(tb['Year'])
    o=0
    tb_1 = pd.DataFrame (np.zeros(500000 * 13).reshape(500000, 13), columns = ['unique_id','Year','Season','DOW', 'TOD', 'SpdAvg','SpdStd','Spd85','SpdFF85','PSL','TTAvg','TTStd','PVTT'])

    for unique_id in unique_id_list:
        b = TMC.loc[(TMC['unique_id'] == unique_id)].reset_index(drop=True)
        TMCs = np.unique(b['tmc'])
        #print(unique_id)
        ## add year later
        for year in year_name:
            for season in season_name:
                for dow in dow_name:
                    for time in time_name:

                        a_1 = tb.loc[(tb['tmc'].isin(TMCs)) & (tb['Season'] == str(season)) & (tb['DOW'] == str(dow)) & (tb['TOD'] == str(time))].reset_index(drop=True)

                        if not a_1.empty:

                            TMC_Spd_N = b.merge(a_1, on='tmc')

                            tb_1['unique_id'][o] = str(unique_id)
                            tb_1['Year'][o] = str(year)
                            tb_1['Season'][o] = str(season)
                            tb_1['DOW'][o] = str(dow)
                            tb_1['TOD'][o] = str(time)
                            tb_1['SpdAvg'][o] = np.average(TMC_Spd_N['SpdAvg'], weights=TMC_Spd_N['tmc_perc'])
                            tb_1['SpdStd'][o] = np.average(TMC_Spd_N['SpdStd'], weights=TMC_Spd_N['tmc_perc'])
                            tb_1['Spd85'][o] = np.average(TMC_Spd_N['Spd85'], weights=TMC_Spd_N['tmc_perc'])
                            tb_1['SpdFF85'][o] = np.average(TMC_Spd_N['SpdFF85'], weights=TMC_Spd_N['tmc_perc'])
                            tb_1['PSL'][o] = myround(np.average(TMC_Spd_N['PSL'], weights=TMC_Spd_N['tmc_perc']))
                            tb_1['TTAvg'][o] = np.average(TMC_Spd_N['TTAvg'], weights=TMC_Spd_N['tmc_perc'])
                            tb_1['TTStd'][o] = np.average(TMC_Spd_N['TTStd'], weights=TMC_Spd_N['tmc_perc'])
                            tb_1['PVTT'][o] = np.average(TMC_Spd_N['PVTT'], weights=TMC_Spd_N['tmc_perc'])
                            o = o+ 1
```

(a)

```python
if __name__ == '__main__':

    with multiprocessing.Pool() as pool:

        # get the path of the directory where the script is located
        dir_path = os.path.dirname(os.path.realpath(__file__))

        # define the relative file path to the data file
        file_path = os.path.join(dir_path, "xd_2U.csv")

        # read the CSV file using pandas
        TMC = pd.read_csv(file_path)

        #TMC = pd.read_csv('C:/Users/j_l848/OneDrive - Texas State University/0_7131_LJL/Data/3_Task3/TMC_SpeedData/TMC_SpeedData/Roads_with_TMC_N_TX.csv')
        TMC = TMC[["Polygon_ID", "XDSegID", "Pct_poly"]]
        # remove rows with missing values in 'SpdAve' column for TMC_Spd_N
        TMC = TMC.dropna(subset=['XDSegID'])
        # compute tmc_perc for TMC_Spd_N
        TMC['tmc_perc'] = TMC['Pct_poly']
        TMC['unique_id'] = TMC['Polygon_ID']
        TMC['tmc'] = TMC['XDSegID']

        unique_id_list = np.unique(TMC["unique_id"])

        chunk_size = 200  # adjust this to control the chunk size
        unique_id_chunks = [list(unique_id_list[i:i+chunk_size]) for i in range(0, len(unique_id_list), chunk_size)]

        results = pool.map(process_tmc_chunk, unique_id_chunks)

        print(results)
```

(b)

Figure 34. Python Code for RHiNO Level Short Duration Speed Summary. (a) Python Code Part 1; (b) Python Code Continued.

3.5.2 Crash Data Assignment

To generate a comprehensive analysis akin to an annual crash summary, the Project Team initiated the process by employing the 'Near' feature tool within ArcGIS Pro. This tool was instrumental in effectively assigning crash events to specific roadway segments. The rationale behind this method lies in the accessibility of pertinent crash data and crash timestamps within the designated crash dataset. The Project Team executed a Python code designed to extract relevant insights from the available crash data. This two-tiered methodology, integrating ArcGIS Pro for initial segment assignment and Python coding for subsequent short-duration analysis, aligns with a systematic process to assign crashes to segments (see Figure 35).

```
import multiprocessing
import pandas as pd
import numpy as np
from datetime import datetime,timezone
import os
def process_tmc_chunk(data):
    # Dynamically determine the unique years for conditions
    unique_years = data['Year'].unique()
    # Update conditions to include all unique years
    conditions = [
        {"Year": year, "Season": ["Winter", "Spring", "Summer", "Fall", "All"],
        "DOW": ["Weekday", "Weekend", "All"], "TOD": ["Daytime", "Nighttime", "All"]}
        for year in unique_years]
    # Placeholder for results
    results = []
    # Iterate over unique xd_id values
    for xd in data['uni_id_new'].unique():
        for condition in conditions:
            for season in condition["Season"]:
                for dow in condition["DOW"]:
                    for tod in condition["TOD"]:
                        # Filter data based on conditions and current xd_id
                        mask = (data['Year'] == condition["Year"]) & (data['uni_id_new'] == xd)
                        if season != "All":
                            mask &= (data['Season'] == season)
                        if dow != "All":
                            mask &= (data['Week'] == dow)
                        if tod != "All":
                            mask &= (data['TOD'] == tod)
                        filtered_data = data[mask]
                        # Compute crash severity counts
                        Crash_5 = (filtered_data['Crash_Sev_'] == 5).sum()
                        Crash_4 = (filtered_data['Crash_Sev_'] == 4).sum()
                        Crash_3 = (filtered_data['Crash_Sev_'] == 3).sum()
                        Crash_2 = (filtered_data['Crash_Sev_'] == 2).sum()
                        Crash_1 = (filtered_data['Crash_Sev_'] == 1).sum()

                        results.append({"uni_id_new": xd, "Year": condition["Year"], "Season": season,
                                        "DOW": dow, "TOD": tod, "Crash_5": Crash_5, "Crash_4": Crash_4,"Crash_3": Crash_3,"Crash_2": Crash_2,"Crash_1": Crash_1})
    # Convert results to dataframe
    results_df = pd.DataFrame(results)
    # get the path of the directory where the script is located-+
    dir_path = os.path.dirname(os.path.realpath(__file__))
    folder = os.path.join(dir_path, "Crash_Summary", "18")
    if not os.path.exists(folder):
        os.makedirs(folder)  # create the directory if it doesn't exist

    filename = str(xd) + '.csv'
    filepath = os.path.join(folder, filename)
    results_df.to_csv(filepath, index=False)
    return results_df
```

Figure 35. Python Code for Short Duration Crash Summary on RHINO Segments.

3.5.3 Weather Measure Assignment

The Project Team has developed an R script to process and analyze extensive precipitation data, specifically focusing on hourly precipitation measures. The script enables the user to transform granular precipitation data into structured insights.

1. **Initialization of Analytical Tools:** The first step involves loading a suite of R packages. R package 'lubridate' facilitates sophisticated date-time manipulation, 'dplyr' and 'tidyr' are for

data manipulation and tidying, and ‘data.table’ offers us the capability to handle large datasets with speed and efficiency.

2. **Workspace Configuration:** To maintain consistency in file paths and operations, the Project Team configures the working directory dynamically to the script's location. This step is bypassed on Shiny servers during application development due to their unique operational environment (see Figure 36).

```
temporalCategories <- function(data) {
  # Extract the day from column names
  # Assuming column names are in the format XMMdd, e.g., X0101
  day_info <- substr(names(data), 2, 3)
  day_info <- as.integer(day_info)

  # Extract the hour from column names
  # Assuming column names are in the format XDDHH, e.g., X0100
  hour_info <- substr(names(data), 4, 5)
  hour_info <- as.integer(hour_info)

  # Determine the year
  year <- years[1]

  # Apply this function to each day in your dataset
  weekend <- sapply(day_info, is_weekend)

  # Apply this function to each hour in your dataset
  daytime <- sapply(hour_info, is_daytime)

  # Create an empty category variable.
  category <- rep(NA, length(weekend))

  # Create 4 categories
  # 1. Weekend Daytime
  # 2. Weekend Nighttime
  # 3. Weekday Daytime
  # 4. Weekday Nightime
  for (i in 1:length(weekend))
  {
    if (weekend[i] == TRUE & daytime[i] == TRUE)
    {
      category[i] <- 1
    } else if (weekend[i] == TRUE & daytime[i] == FALSE)
    {
      category[i] <- 2
    } else if (weekend[i] == FALSE & daytime[i] == TRUE)
    {
      category[i] <- 3
    } else
    {
      category[i] <- 4
    }
  }

  # Return the result.
  return(category)
}
```

Figure 36. Temporal Categories Function.

3. **Numerical Representation:** The Project Team deactivates R's default scientific notation to ensure that the numerical data being worked with is immediately interpretable without the need for transformation.
4. **Defining Temporal Variables:** The analysis spans multiple years, segmented into monthly intervals.
5. **Development of Classification Functions:** Functions have been written by the Project Team to ascertain whether a date falls on a weekend or a weekday and to classify hours as daytime or nighttime based on predefined thresholds. These classifications are instrumental for the layered analysis that follows.
6. **Temporal Categorization Function:** The custom function 'temporalCategories' tags each data column with time-based categories, delineating the data by both the day of the week and the hour of the day (see Figure 36).

```r
# Function to produce the matrix of statistical data.
getMatrix <- function(data) {
  # Initialize a matrix to store the results
  results_matrix <- as.data.frame(matrix(nrow = nrow(data), ncol = 36))
  colnames(results_matrix) <- c("Sum_All_All", "Avg_All_All", "Max_All_All", "Min_All_All",
                                "Sum_Weekend_All", "Avg_Weekend_All", "Max_Weekend_All", "Min_Weekend_All",
                                "Sum_Weekday_All", "Avg_Weekday_All", "Max_Weekday_All", "Min_Weekday_All",
                                "Sum_All_Daytime", "Avg_All_Daytime", "Max_All_Daytime", "Min_All_Daytime",
                                "Sum_All_Nighttime", "Avg_All_Nighttime", "Max_All_Nighttime", "Min_All_Nighttime",
                                "Sum_Weekend_Daytime", "Avg_Weekend_Daytime", "Max_Weekend_Daytime", "Min_Weekend_Daytime",
                                "Sum_Weekend_Nighttime", "Avg_Weekend_Nighttime", "Max_Weekend_Nighttime", "Min_Weekend_Nighttime",
                                "Sum_Weekday_Daytime", "Avg_Weekday_Daytime", "Max_Weekday_Daytime", "Min_Weekday_Daytime",
                                "Sum_Weekday_Nighttime", "Avg_Weekday_Nighttime", "Max_Weekday_Nighttime", "Min_Weekday_Nighttime")

  # Get the full months sum, avg, max, and min with categories.
  results_matrix$Sum_All_All <- rowSums(data[, 1:ncol(data)])
  results_matrix$Avg_All_All <- round(rowSums(data[, 1:ncol(data)]) / ncol(data),5)
  results_matrix$Max_All_All <- apply(data, 1, max, na.rm=TRUE)
  results_matrix$Min_All_All <- apply(data, 1, min, na.rm=TRUE)

  # 1. Weekend Daytime
  # 2. Weekend Nighttime
  # 3. Weekday Daytime
  # 4. Weekday Nightime
  # Assuming 'category' is a vector that matches the columns of 'data'
  weekend_indices <- which(category == 1 | category == 2)
  weekday_indices <- which(category == 3 | category == 4)
  daytime_indices <- which(category == 1 | category == 3)
  nighttime_indices <- which(category == 2 | category == 4)
  weekend_daytime_indices <- which(category == 1)
  weekend_nighttime_indices <- which(category == 2)
  weekday_daytime_indices <- which(category == 3)
  weekday_nighttime_indices <- which(category == 4)
```

(a)

```
  # First, the simple averages by categories.
  results_matrix$Avg_Weekend_All <- round(rowSums(data[, ..weekend_indices]) / length(weekend_indices),5)
  results_matrix$Avg_Weekday_All <- round(rowSums(data[, ..weekday_indices]) / length(weekday_indices),5)
  results_matrix$Avg_All_Daytime <- round(rowSums(data[, ..daytime_indices]) / length(daytime_indices),5)
  results_matrix$Avg_All_Nighttime <- round(rowSums(data[, ..nighttime_indices]) / length(nighttime_indices),5)
  results_matrix$Avg_Weekend_Daytime <- round(rowSums(data[, ..weekend_daytime_indices]) / length(weekend_daytime_indices),5)
  results_matrix$Avg_Weekend_Nighttime <- round(rowSums(data[, ..weekend_nighttime_indices]) / length(weekend_nighttime_indices),5)
  results_matrix$Avg_Weekday_Daytime <- round(rowSums(data[, ..weekday_daytime_indices]) / length(weekday_daytime_indices),5)
  results_matrix$Avg_Weekday_Nighttime <- round(rowSums(data[, ..weekday_nighttime_indices]) / length(weekday_nighttime_indices),5)

  # The Weekend Data.
  temp <- data[, ..weekend_indices]
  results_matrix$Max_Weekend_All <- apply(temp, 1, max, na.rm=TRUE)
  results_matrix$Min_Weekend_All <- apply(temp, 1, min, na.rm=TRUE)

  # The Weekday Data
  temp <- data[, ..weekday_indices]
  results_matrix$Max_Weekday_All <- apply(temp, 1, max, na.rm=TRUE)
  results_matrix$Min_Weekday_All <- apply(temp, 1, min, na.rm=TRUE)

  # The Daytime Data
  temp <- data[, ..daytime_indices]
  results_matrix$Max_All_Daytime <- apply(temp, 1, max, na.rm=TRUE)
  results_matrix$Min_All_Daytime <- apply(temp, 1, min, na.rm=TRUE)

  # The Nighttime Data
  temp <- data[, ..nighttime_indices]
  results_matrix$Max_All_Nighttime <- apply(temp, 1, max, na.rm=TRUE)
  results_matrix$Min_All_Nighttime <- apply(temp, 1, min, na.rm=TRUE)

  # The Four Combinations of Daytime/Nighttime, Weekend/Weekday
  temp <- data[, ..weekend_daytime_indices]
  results_matrix$Max_Weekend_Daytime <- apply(temp, 1, max, na.rm=TRUE)
  results_matrix$Min_Weekend_Daytime <- apply(temp, 1, min, na.rm=TRUE)
  temp <- data[, ..weekday_daytime_indices]
  results_matrix$Max_Weekday_Daytime <- apply(temp, 1, max, na.rm=TRUE)
  results_matrix$Min_Weekday_Daytime <- apply(temp, 1, min, na.rm=TRUE)
  temp <- data[, ..weekend_nighttime_indices]
  results_matrix$Max_Weekend_Nighttime <- apply(temp, 1, max, na.rm=TRUE)
  results_matrix$Min_Weekend_Nighttime <- apply(temp, 1, min, na.rm=TRUE)
  temp <- data[, ..weekday_nighttime_indices]
  results_matrix$Max_Weekday_Nighttime <- apply(temp, 1, max, na.rm=TRUE)
  results_matrix$Min_Weekday_Nighttime <- apply(temp, 1, min, na.rm=TRUE)

  return(results_matrix)
}
```

(b)

Figure 37. Function 'getMatrix'

7. **Statistical Computation Function:** The getMatrix (Figure 37) function aggregates the data to produce sums, averages, maximums, and minimums for the various time categories, thus enabling the Project Team to perform a comprehensive analysis.
8. **Data Aggregation Workflow:** The Project Team utilizes a nested loop to iterate through the years and months, systematically processing hourly data into a structured statistical matrix. This approach allows for efficient categorization and computation of aggregated statistics.
9. **Spatial Data Integration:** In the integration process, the Project Team leverages an id_match.csv file, as indicated in Step 4 of the Data Conversion section, to map each precipitation data point to its corresponding geographic identifier.
10. **Seasonal Data Aggregation:** Acknowledging the significance of seasonal variations in weather patterns, the developed coding script groups data by meteorological seasons. This step involves recalculating aggregated statistics to capture variations during distinct periods.
11. **Data Exportation:** In the final phase of the process, the Project Team exports the synthesized data into CSV files. These files serve as a comprehensive and accessible record for subsequent analysis, reporting, and peer review.

3.5.4 Driveway Density

The Project Team used Longitudinal Employer-Household Dynamics (LEHD) (USCB, 2024) and Smart Location Database (SLD) (EPA, 2013) to determine driveway density type. An interactive tool (https://aitlab.shinyapps.io/DrivewayDensity_V03/) has been developed to show driveway types. Figure 38 shows the interface of the tool.

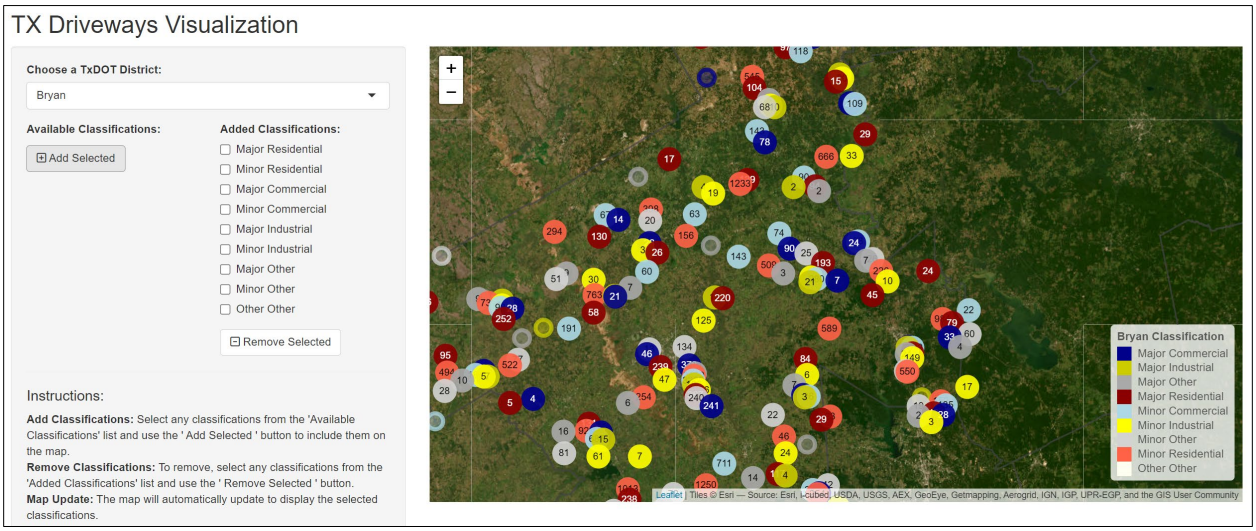


Figure 38. Interface of Driveway Density Type Interactive Tool.

3.6 SUMMARY

Chapter 3 focuses on the data preparation process for the safety evaluation of urban roadways. The Project Team gathered and integrated various datasets, including traffic crash data from the CRIS, roadway inventory data from the RHiNO, speed data from the NPMRDS and INRIX XD, and weather data from the C3S and CAMS. A comprehensive data conflation framework was developed to merge these diverse datasets. The preparation involved assigning speed measures, integrating crash data, and processing weather data, such as precipitation from Copernicus, to enhance the overall quality of the dataset. Additionally, both annual and short-duration data were processed to support long-term safety evaluation and real-time risk analysis.

CHAPTER 4:
DATA ANALYSIS AND MODEL DEVELOPMENT

4.1 INTRODUCTION

This chapter offers a concise overview of the exploratory data analysis conducted on the databases created for both annual and short-duration data. It also includes annual-level SPFs for urban two-lane and multilane roadways.

4.2 ANNUAL LEVEL DATABASES

The predictive method in HSM Chapter 12 addresses the following urban and suburban arterial facilities: two- and four-lane undivided facilities, four-lane divided facilities, and three- and five-lane facilities with center two-way left-turn lanes. Divided arterials are defined as nonfreeway facilities with travel lanes in both directions separated by a raised or depressed median and may include occasional grade-separated interchanges that are not the primary form of access. These predictive models exclude arterial sections within interchange limits that have free-flow ramp terminals. Arterials with a flush separator (painted median) between travel lanes are classified as undivided facilities, not divided. Separate prediction models are provided for arterials with a flush separator functioning as a center two-way left-turn lane. The specific site types are defined as follows:

- **Two-lane undivided arterial (2U):** A roadway consisting of two lanes with a continuous cross-section providing two directions of travel, in which the lanes are not physically separated by distance or a barrier.
- **Three-lane arterial (3T):** A roadway consisting of three lanes with a continuous cross-section providing two directions of travel, in which the center lane is a TWLTL.
- **Four-lane undivided arterial (4U):** A roadway consisting of four lanes with a continuous cross-section providing two directions of travel, in which the lanes are not physically separated by distance or a barrier.
- **Four-lane divided arterial (4D):** A roadway consisting of four lanes with a continuous cross-section providing two directions of travel, in which opposing directions are physically separated by a raised or depressed median or by distance.
- **Five-lane arterial with center TWLTL (5T):** A roadway consisting of five lanes with a continuous cross-section providing two directions of travel, in which the center lane is a TWLTL.

This section presents the descriptive statistics of the key variables used for the annual level analysis. Table 29 lists crash frequencies by different crash severity type levels and roadway facility types. Figure 39 displays the distribution of urban roadways by facility types. Table 29 through Table 34, list descriptive statistics of the key variables for annual level analysis.

Table 29. Total Road Length and Crash Counts by Facility Types for the Annual Model.

Facility Type	Total Length (miles)	Total Number of Crashes (By Crash Severity Type)				
		K	A	B	C	O
Two-lane undivided roadways (2U)	28,265.17	773	3,175	10,368	14,771	63,537
Three-lane roadways (3T)	115.47	24	54	217	340	1,197
Four-lane divided roadways (4D)	671.00	1,259	5,109	24,170	42,247	150,939
Four-lane undivided roadways (4U)	2,632.00	1,378	5,805	28,226	49,509	177,286
Five-lane roadways (5T)	170.00	216	702	3,017	4,448	15,148
All	31,853.63	3,650	14,845	65,998	111,315	408,107

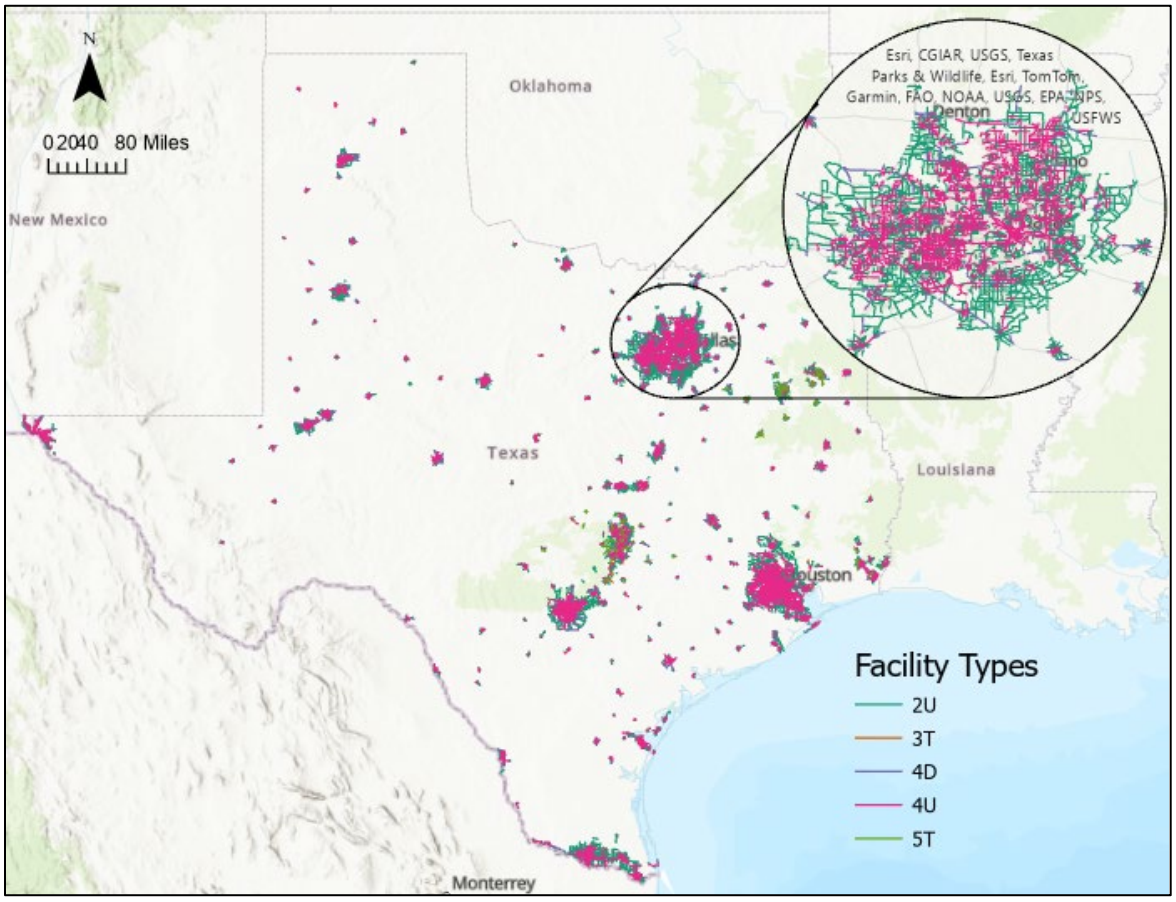


Figure 39. Distribution of Urban Roadways by Facility Types.

Table 30. Descriptive Statistics of Urban Two-lane Undivided Roadways (Annual Level Data).

Variables	Code	Mean	SD	Min	Max
AADT (vpd)	ADT_CUR	5003.42	4927.74	2	56535.00
Truck proportion	TRK_AAD	6.55	6.55	0	94.10
Lane width (ft)	LANE_WI	12.06	3.34	4	20.00
Inside shoulder width (ft)	S_WID_I	2.04	3.60	0	30.00

Outside shoulder width (ft)	S_WID_O	2.08	3.62	0	32.00
K-factor	K_FAC	10.79	2.88	3.9	99.90
Minor Commercial Driveway	MnrCmmr	1.87	3.33	0	55.00
Major Commercial Driveway	MjrCmmr	1.06	2.82	0	110.00
Minor Other Driveway	MnrOthr	0.07	0.55	0	31.00
Major Industrial Driveway	MjrInds	0.16	0.85	0	26.00
Minor Industrial Driveway	MnrInds	0.19	1.01	0	27.00
Minor Residential Driveway	MnrRsdn	0.50	1.53	0	25.00
Average operating speed (mph)	SpdAve	64.75	11.34	0	75.00
85 percentile speed (mph)	SEF	69.78	12.24	0	75.00
Sum of precipitation (in)	pS	34.23	20.13	0	66.60
Average of precipitation (in)	pA	0.01	0.00	0	0.02

Table 31. Descriptive Statistics of Urban Three-lane Roadways (Annual Level Data).

Variables	Code	Mean	SD	Min	Max
AADT (vpd)	ADT_CUR	12361.88	6649.54	1035	41353.00
Truck proportion	TRK_AAD	7.88	5.04	1.6	32.10
Lane width (ft)	LANE_WI	12.74	3.13	10	20.00
Inside shoulder width (ft)	S_WID_I	4.32	3.97	0	14.00
Outside shoulder width (ft)	S_WID_O	4.31	3.91	0	13.00
K-factor	K_FAC	9.63	1.99	6	21.70
Minor Commercial Driveway	MnrCmmr	0.09	0.47	0	5.00
Major Commercial Driveway	MjrCmmr	1.62	1.48	0	7.00
Minor Other Driveway	MnrOthr	0.01	0.12	0	1.00
Major Industrial Driveway	MjrInds	0.25	0.69	0	5.00
Minor Industrial Driveway	MnrInds	0.01	0.08	0	1.00
Minor Residential Driveway	MnrRsdn	0.12	0.45	0	4.00
Average operating speed (mph)	SpdAve	70.67	9.25	0	75.00
85 percentile speed (mph)	SEF	72.23	8.89	0	75.00
Sum of precipitation (in)	pS	29.37	18.79	0	45.78
Average of precipitation (in)	pA	0.02	0.01	0	0.05

Table 32. Descriptive Statistics of Urban Four-Lane Undivided Roadways (Annual Level Data).

Variables	Code	Mean	SD	Min	Max
AADT (vpd)	ADT_CUR	13357.15	8762.67	2	71576.00
Truck proportion	TRK_AAD	6.31	5.85	0	81.40
Lane width (ft)	LANE_WI	12.05	2.60	4	20.00
Inside shoulder width (ft)	S_WID_I	2.22	3.89	0	30.00
Outside shoulder width (ft)	S_WID_O	2.38	4.08	0	26.00
K-factor	K_FAC	9.81	1.94	4.3	50.00
Minor Commercial Driveway	MnrCmmr	0.75	2.47	0	47.00
Major Commercial Driveway	MjrCmmr	3.17	4.95	0	77.00
Minor Other Driveway	MnrOthr	0.04	0.35	0	12.00
Major Industrial Driveway	MjrInds	0.53	1.90	0	48.00
Minor Industrial Driveway	MnrInds	0.11	0.94	0	29.00

Minor Residential Driveway	MnrRsdn	0.12	0.57	0	17.00
Average operating speed (mph)	SpdAve	69.86	12.14	0	80.00
85 percentile speed (mph)	SEF	74.41	13.21	0	80.00
Sum of precipitation (in)	pS	33.31	15.13	0	51.00
Average of precipitation (in)	pA	0.03	0.01	0	0.02

Table 33. Descriptive Statistics of Urban Four-Lane Divided Roadways (Annual Level Data).

Variables	**Code**	**Mean**	**SD**	**Min**	**Max**
AADT (vpd)	ADT_CUR	20158.73	12155.84	250	140415.00
Truck proportion	TRK_AAD	9.48	8.01	0.1	81.40
Lane width (ft)	LANE_WI	12.41	1.56	6	20.00
Inside shoulder width (ft)	S_WID_I	6.07	5.43	0	30.00
Outside shoulder width (ft)	S_WID_O	12.30	7.64	0	44.00
K-factor	K_FAC	9.52	1.43	6	21.20
Minor Commercial Driveway	MnrCmmr	0.08	0.45	0	8.00
Major Commercial Driveway	MjrCmmr	1.73	2.88	0	31.00
Minor Other Driveway	MnrOthr	0.04	0.27	0	6.00
Major Industrial Driveway	MjrInds	1.12	1.89	0	23.00
Minor Industrial Driveway	MnrInds	0.02	0.22	0	5.00
Minor Residential Driveway	MnrRsdn	0.03	0.22	0	5.00
Average operating speed (mph)	SpdAve	70.23	11.45	0	80.00
85 percentile speed (mph)	SEF	72.12	12.75	0	80.00
Sum of precipitation (in)	pS	33.31	15.13	0	51.00
Average of precipitation (in)	pA	0.03	0.01	0	0.02

Table 34. Descriptive Statistics of Urban Five-lane Roadways (Annual Level Data).

Variables	Code	Mean	SD	Min	Max
AADT (vpd)	ADT_CUR	19439.50	11000.37	2099	71576.00
Truck proportion	TRK_AAD	7.509	4.739	1.3	51.50
Lane width (ft)	LANE_WI	13.354	2.581	10	20.00
Inside shoulder width (ft)	S_WID_I	4.344	4.379	0	18.00
Outside shoulder width (ft)	S_WID_O	4.859	4.496	0	20.00
K-factor	K_FAC	9.198	1.384	6.4	21.50
Minor Commercial Driveway	MnrCmmr	0.027	0.170	0	2.00
Major Commercial Driveway	MjrCmmr	1.570	1.818	0	12.00
Minor Other Driveway	MnrOthr	0.006	0.077	0	1.00
Major Industrial Driveway	MjrInds	0.529	1.102	0	8.00
Minor Industrial Driveway	MnrInds	0.004	0.067	0	1.00
Minor Residential Driveway	MnrRsdn	0.028	0.198	0	2.00
Average operating speed (mph)	SpdAve	71.22	10.45	0	80.00
85 percentile speed (mph)	SEF	73.27	12.71	0	80.00
Sum of precipitation (in)	pS	34.72	13.39	0	57
Average of precipitation (in)	pA	0.009	0.004	0	0.02

4.3 ANNUAL LEVEL SAFETY PERFORMANCE FUNCTIONS

The Project Team populated a list of speed measures that quantify speed with respect to different aspects. Given a long list, it is imperative to select a measure that is appropriate and meaningful to include in the SPF development. The team conducted the correlation analysis between speed measures for the year 2018 and calculated the Pearson correlation coefficient, as shown in Table 35. The speed measures are highly correlated, and it was decided to use the 85th percentile speed where the operational speeds are higher than the posted speed limit and the standard deviation in the speed (SpdStd) in SPFs, given their wide range of use.

Table 35. Correlation Analysis Results.

Variable	Spd Ave	Spd Std	Spd85	Spd Ave Day	Spd Std Day	Spd Ave Night	Spd Std Night	Spd Ave MMW T	Spd Std MMW T	Spd Ave FSS	Spd Std FSS	Spd FF Ave	PSL
Spd Ave	1												
Spd Std	0.437	1											
Spd85	0.995	0.516	1										
SpdAve Day	0.998	0.403	0.989	1									
SpdStd Day	0.439	0.975	0.511	0.405	1								
SpdAve Night	0.998	0.470	0.997	0.991	0.472	1							
SpdStd Night	0.444	0.948	0.521	0.419	0.871	0.469	1						
SpdAve MMWT	1.000	0.430	0.994	0.998	0.432	0.997	0.439	1					
SpdStd MMWT	0.428	0.995	0.505	0.393	0.978	0.462	0.929	0.420	1				
SpdAve FSS	1.000	0.447	0.995	0.997	0.449	0.998	0.451	0.999	0.438	1			
SpdStd FSS	0.441	0.983	0.520	0.408	0.943	0.472	0.960	0.435	0.960	0.448	1		
SpdFF Ave	0.994	0.510	0.998	0.987	0.504	0.996	0.517	0.993	0.500	0.994	0.514	1	
PSL	0.989	0.538	0.997	0.982	0.529	0.992	0.548	0.988	0.527	0.989	0.544	0.998	1

4.3.1 SPFs for Freeways

For the SPF development, the team developed adjustment factors for different number of lanes since each type of cross section has different safety performance. Since PDO crashes are usually under-reported, it was decided to develop separate models for FI and PDO crashes. The Project Team 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 below reflects the findings from several preliminary regression analyses. The same form is also used for modeling the PDO crashes, even if some variables are insignificant or counter-intuitive. The predicted crash frequency is calculated as follows using Equation (41).

$$N = L \times y \times e^{b_0 + b_6 I_6 + b_8 I_8 + b_{10} I_{10} + b_{aadt} \ln(AADT)} \times CMF_{tk} \times CMF_{lw} \times CMF_{isw} \times CMF_{osw} \times CMF_{mw} \times CMF_{spd} \quad (41)$$

With,

$$CMF_{tk} = e^{b_{tk}(tk\_perc)}$$

$$CMF_{lw} = \begin{cases} e^{b_{lw}(lw-12)}, & if\ lw \le 12ft \\ e^{b_{lw,2}}, & if\ lw > 12f \end{cases}$$

$$CMF_{isw} = \begin{cases} e^{b_{isw}(isw-4)}, & if\ isw \le 4ft \\ e^{b_{isw,2}}, & if\ isw > 4ft \end{cases}$$

$$CMF_{osw} = \begin{cases} e^{b_{osw}(osw-8)}, & if\ osw \le 8ft \\ e^{b_{osw,2}}, & if\ osw > 8ft \end{cases}$$

$$CMF_{mw} = \begin{cases} e^{b_{mb}(mw-30)}, & if\ median\ barrier\ is\ present \\ e^{b_{mw}(mw-30)}, & if\ median\ barrier\ is\ not\ present \end{cases}$$

$$CMF_{spd} = e^{b_{spd}(Sfef-PSL)}$$

Where,

N = Predicted annual average crash frequency,
L = Segment length, miles,
y = Number of years of crash data,
I_6 = Indicator variable for 6-lane section (1 if 6 lanes, 0 otherwise),
I_8 = Indicator variable for 8-lane section (1 if 8 lanes, 0 otherwise),
I_{10} = Indicator variable for 10+-lane section (1 if 10+ lanes, 0 otherwise),
$AADT$ = Average Annual Daily Traffic, vehicles per day,
CMF_{tk} = CMF for truck proportion in the traffic mix,
CMF_{lw} = CMF for lane width,
CMF_{isw} = Crash Modification Factor for inside shoulder width,
CMF_{osw} = CMF for outside shoulder width,
CMF_{mw} = CMF for median width,
CMF_{spd} = CMF for excess speed,
tk_perc = Percent of trucks in the traffic mix, %,
lw = Lane width, feet,
isw = Inside shoulder width, feet,
osw = Outside shoulder width, feet,
mw = Median width when barrier is not present, feet,
PSL = Posted speed limit, mph,
b_j = Calibrated coefficients.

Table 36 and Table 37 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, when the coefficient is not statistically significant but is intuitive and within logical boundaries, then those variables are considered as well. The NonLinear MIXED-effects models (NLMIXED) procedure in the Statistical Analysis System (SAS) software was used to estimate the proposed model coefficients. This procedure was used because the proposed predictive model is both nonlinear and discontinuous. The log-likelihood function for the al NB distribution was used to determine the best-fit model coefficients.

Table 36. Calibrated Coefficients for Fatal and Injury Crashes on Freeways.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-5.380	0.189	-28.51	<.0001
b_6	Adjustment factor for 6 lanes	0.216	0.026	8.29	<.0001
b_8	Adjustment factor for 8 lanes	0.359	0.035	10.18	<.0001
b_{10}	Adjustment factor for 10+ lanes	0.450	0.052	8.58	<.0001
b_{aadt}	AADT	0.700	0.017	41.04	<.0001
b_{tk}	Truck proportion	-0.022	0.001	-21.02	<.0001
b_{lw}	Lane width (if ≤12ft)	-0.101	0.051	-2.00	0.0457
$b_{lw,2}$	Lane width (if >12ft)	0.102	0.063	1.61	0.1084
b_{isw}	Inside shoulder width (if ≤4t)	-0.021	0.004	-5.22	<.0001
$b_{isw,2}$	Inside shoulder width (if >4ft)	-0.011	0.027	-0.39	0.6935
b_{osw}	Outside shoulder width (if ≤8ft)	-0.005	0.006	-0.92	0.3575
$b_{osw,2}$	Outside shoulder width (if >8ft)	0.024	0.027	0.88	0.3785
b_{mb}	Median width (if barrier is present)	-0.168	0.034	-4.88	<.0001
b_{mw}	Median width (if barrier is not present)	-0.481	0.038	-12.55	<.0001
b_{spd}	Excess speed	0.972	0.334	2.91	0.0036
k	Inverse dispersion parameter	1.714	0.017	100.77	<.0001

Table 37. Calibrated Coefficients for Property Damage Only Crashes on Freeways.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-4.545	0.187	-24.27	<.0001
b_6	Adjustment factor for 6 lanes	0.164	0.026	6.32	<.0001
b_8	Adjustment factor for 8 lanes	0.286	0.035	8.15	<.0001
b_{10}	Adjustment factor for 10+ lanes	0.470	0.053	8.84	<.0001
b_{aadt}	AADT	0.701	0.017	41.12	<.0001
b_{tk}	Truck proportion	-0.014	0.001	-14.51	<.0001
b_{lw}	Lane width (if ≤12ft)	-0.041	0.050	-0.82	0.4144
$b_{lw,2}$	Lane width (if >12ft)	0.120	0.064	1.87	0.0618
b_{isw}	Inside shoulder width (if ≤4t)	-0.013	0.004	-3.26	0.0011
$b_{isw,2}$	Inside shoulder width (if >4ft)	0.011	0.028	0.4	0.6896
b_{osw}	Outside shoulder width (if ≤8ft)	-0.011	0.006	-1.81	0.0696
$b_{osw,2}$	Outside shoulder width (if >8ft)	-0.030	0.027	-1.11	0.2659
b_{mb}	Median width (if barrier is present)	-0.201	0.032	-6.2	<.0001
b_{mw}	Median width (if barrier is not present)	-0.583	0.034	-17.01	<.0001
b_{spd}	Excess speed	0.103	0.324	0.32	0.7497
k	Inverse dispersion parameter	1.628	0.015	110.44	<.0001

A comparison of different calibrated freeway SPFs for four-lane freeways is shown in Figure 40 and Figure 41 for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the HSM-calibrated SPFs (Geedipally et al., 2022) and the Texas Roadway Safety Design Workbook (called as Texas WB going forward) (Bonneson and Pratt, 2008). The equations are plotted for the case of all CMFs equal to 1.0 (representing base conditions). It is important to note that the SPFs do not include the same set of base conditions and thus they are not directly comparable to each other. In addition, the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since Texas WB includes SPFs for FI crashes only, the comparison is made just with HSM SPFs for total crashes.

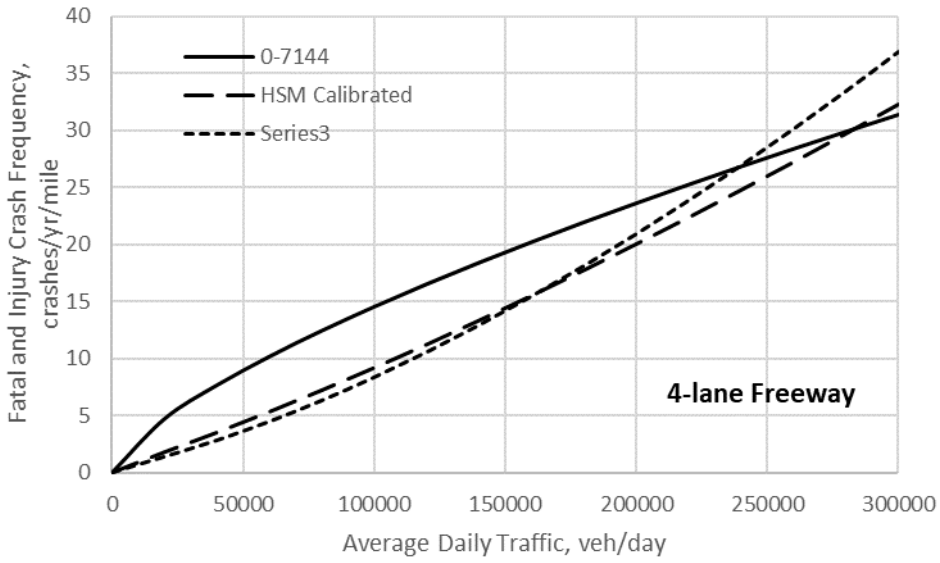


Figure 40. SPF comparison for Fatal and Injury Crashes on Freeways.

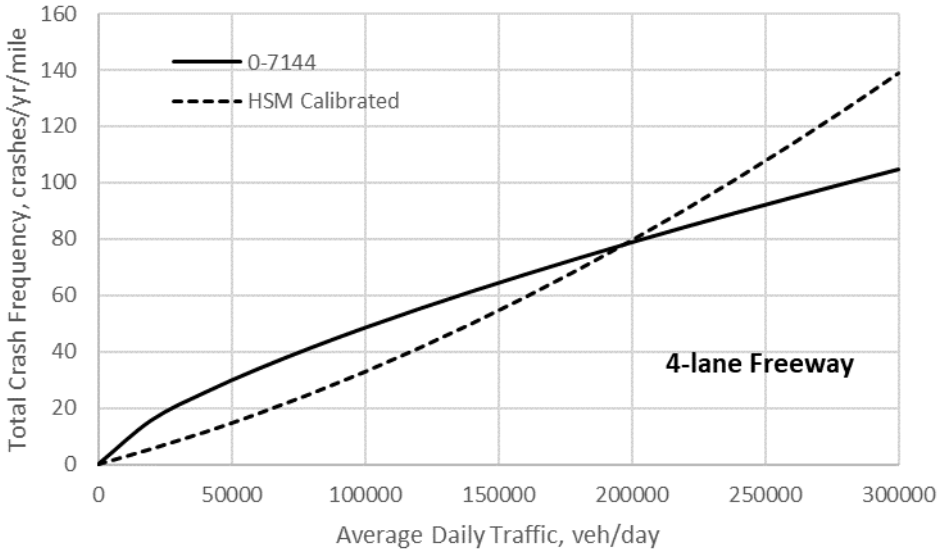


Figure 41. SPF comparison for Total Crashes on Freeways.

Crash Modification Factors (CMF)

CMFs were calibrated in conjunction with the SPFs. All of them were calibrated using the FI crash data. Collectively, they describe the relationship between various operational and geometric factors and crash frequency. These CMFs are described in this section and, where possible, compared with the findings from HSM and Texas WB as a means of model validation.

Truck Proportion CMF

The truck proportion CMF is described using Equation 42:

$$CMF_{tk} = e^{-0.022(tk\_perc)} \quad (42)$$

The base condition for this CMF is no trucks in the traffic mix. The truck proportion CMF is shown in Figure 42. The CMF for truck proportion for urban arterials from Texas WB is used for comparison. Both CMFs show similar trends. The CMF shows that the crashes decrease as the proportion of trucks in traffic increases. Although this may seem counterintuitive, trucks usually travel on high-standard roads and that is reflected in this CMF.

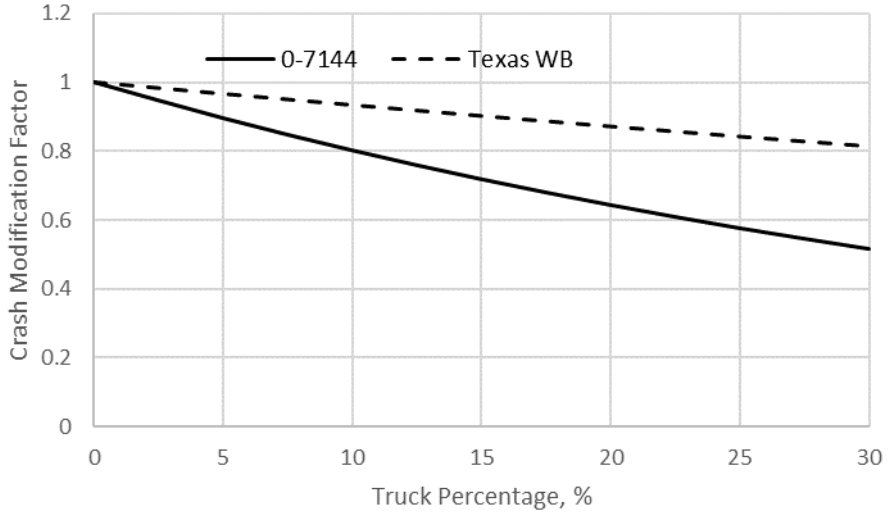


Figure 42. CMF for Truck Proportion on Freeways.

Lane Width CMF

The lane width CMF is described using Equation 43:

$$CMF_{lw} = \begin{cases} e^{-0.101\times(lw-12)}, & if\ lw \leq 12ft \\ 1.0, & if\ lw > 12ft \end{cases} \quad (43)$$

The base condition for this CMF is a 12-ft lane width. The lane width used in this CMF is an average for all through lanes on the segment. The lane width CMF developed in this study is shown in Figure 43 using a solid trend line. The lane widths used to calibrate this CMF range from 10 to 13 ft. The coefficient for lane widths greater than 12ft is found to be statistically insignificant and counterintuitive, so a value of 1.0 is used. Figure 43 also presents the CMFs from HSM and the Texas WB. Broken lines are used to differentiate these CMFs from the one proposed in this research project. The proposed CMF is shown to be more sensitive to lane width than the CMFs in HSM and Texas WB.

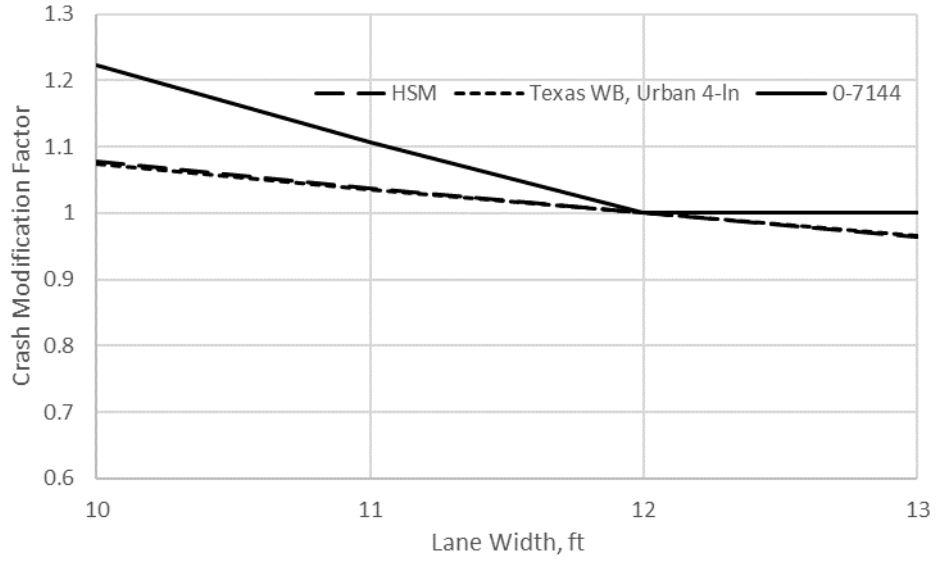


Figure 43. CMF for Lane Width on Freeways.

Inside Shoulder Width CMF

The inside shoulder width CMF is described using Equation 44:

$$CMF_{isw} = \begin{cases} e^{-0.021\times(isw-4)}, & if\ isw \leq 4ft \\ 0.99, & if\ isw > 4ft \end{cases} \quad (44)$$

The base condition for this CMF is a 4-ft inside shoulder width. The width used in this CMF is an average for inside shoulders in both directions. The inside shoulder width CMF developed in this study is shown in Figure 44 using a solid trend line. The inside shoulder widths used to calibrate this CMF range from 2 to 8 ft. Also shown in Figure 44 are CMFs presented in the HSM and Texas WB. Broken lines are used to differentiate these CMFs from the one proposed in this research project. The proposed CMF closely tracks the CMFs presented in the HSM and Texas WB, however this study found that the rate of improvement with increased shoulder width diminishes after 4ft.

Figure 44. CMF for Inside Shoulder Width on Freeways.

Outside Shoulder Width CMF

The outside shoulder width CMF is described using Equation 45:

$$CMF_{osw} = \begin{cases} e^{-0.05\times(osw-8)}, & if\ osw \leq 8ft \\ 1.0, & if\ osw > 8ft \end{cases} \quad (45)$$

The base condition for this CMF is an 8-ft outside shoulder width. The width used in this CMF is average for outside shoulders in both directions. The outside shoulder width CMF developed in this study is shown in Figure 45 using a solid trend line. The outside shoulder widths used to calibrate this CMF range from 6 to 12 ft. The coefficient for shoulder widths greater than 8ft is found to be statistically insignificant and counterintuitive, so a value of 1.0 is used. Also shown in Figure 45 are CMFs presented in HSM and Texas WB. Since HSM and Texas WB have a 10-ft base condition, the CMF developed in this study is adjusted accordingly. Broken lines are used to differentiate these CMFs from the one proposed in this research project. The proposed CMF closely tracks the CMFs presented in the Texas WB. The CMF presented in HSM is shown to be more sensitive to outside shoulder width than the proposed CMF or the one in the Texas WB.

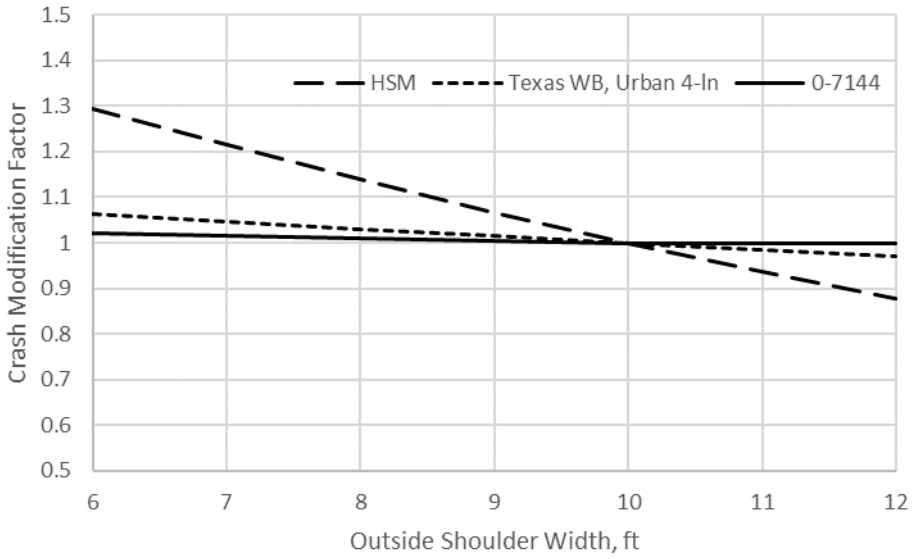


Figure 45. CMF for Outside Shoulder Width on Freeways.

Median Width CMF

The median width CMF is described using Equation 46.

$$CMF_{mw} = \begin{cases} e^{-0.168\times0.01\times(mw-30)}, & if\ median\ barrier\ is\ present \\ e^{-0.481\times0.01\times(mw-30)}, & if\ median\ barrier\ is\ not\ present \end{cases} \quad (46)$$

The base condition for this CMF is a 30-ft median width. The median width CMF is shown in Figure 46 using a solid trend line. The CMF proposed in this research is compared with the CMF in HSM and Texas WB in Figure 46. Broken lines are used to differentiate these CMFs from the one proposed in this research project. The proposed CMF closely tracks the CMFs presented in the HSM for medians with barriers present.

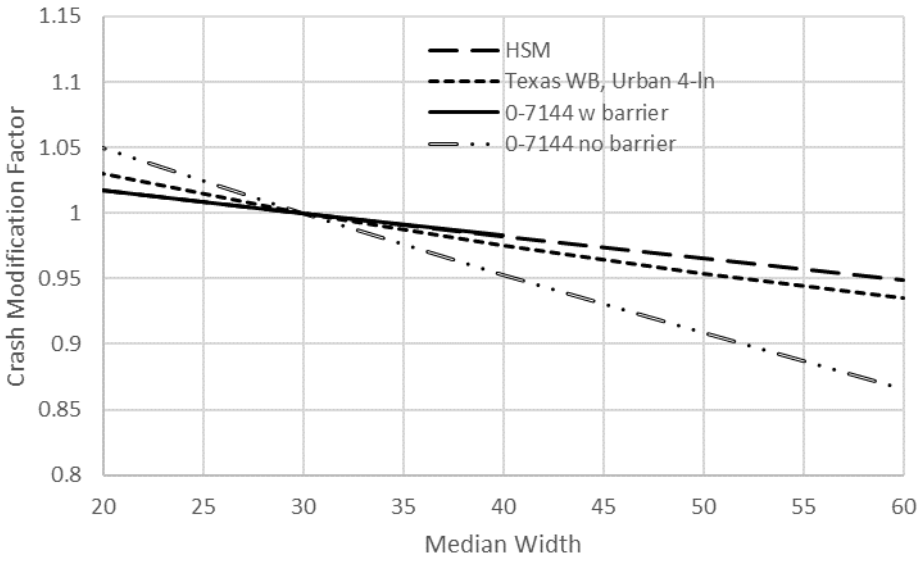


Figure 46. CMF for Median Width on Freeways.

Excess Speed CMF

The excess speed CMF is described using Equation 47.

$$CMF_{spd} = e^{0.972 \times 0.01 \times (Sfef - PSL)} \quad (47)$$

The base condition for this CMF varies according to the posted speed limit. Since the operating speed CMF does not exist in HSM or Texas WB, a comparison could not be made. The CMF shows that exceeding the PSL by 10 mph increases the crashes by 10% (see Figure 47).

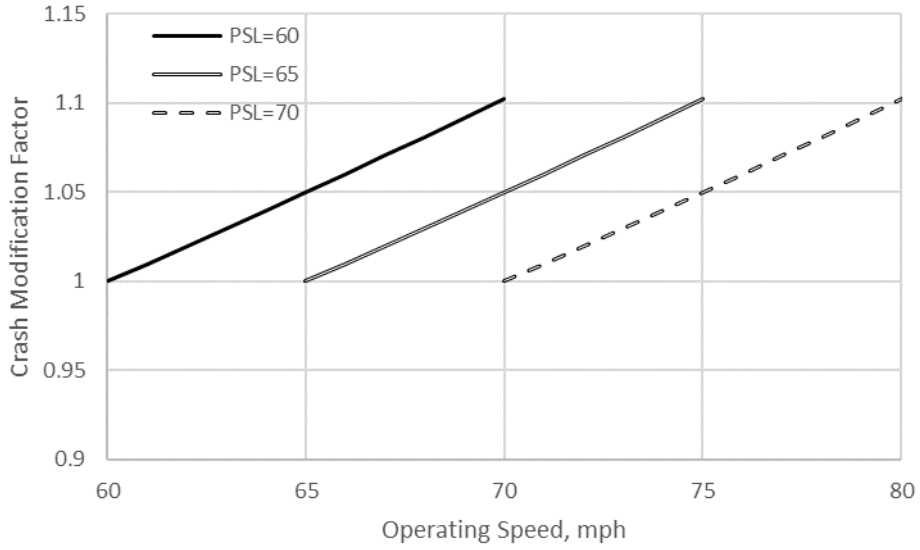


Figure 47. CMF for Operating Speeds on Freeways.

4.3.2 SPFs for Multi-lane Divided Highways

For the SPF development, the team developed adjustment factors for different number of lanes since each type of cross section has different safety performance. The predicted crash frequency is calculated as follows in Equation (48);

$$N = L \times y \times e^{b_0+b_6I_6+b_8I_8+b_{aadt}\ln(AADT)} \times CMF_{tk} \times CMF_{lw} \times CMF_{isw} \times CMF_{osw} \times CMF_{mw} \times CMF_{dw} \times CMF_{spd}; \quad (48)$$

With,

$$CMF_{tk} = e^{b_{tk}(tk\_perc)}$$

$$CMF_{lw} = \begin{cases} e^{b_{lw}(lw-12)}, & if\ lw \le 12ft \\ e^{b_{lw,2}}, & if\ lw > 12ft \end{cases}$$

$$CMF_{isw} = \begin{cases} e^{b_{isw}(isw-4)}, & if\ isw \le 4ft \\ e^{b_{isw,2}}, & if\ isw > 4ft \end{cases}$$

$$CMF_{osw} = \begin{cases} e^{b_{osw}(osw-8)}, & if\ osw \le 8ft \\ e^{b_{osw,2}}, & if\ osw > 8ft \end{cases}$$

$$CMF_{mw} = \begin{cases} e^{b_{mb}(mw-30)}, & if\ median\ barrier\ is\ present \\ e^{b_{mw}(mw-30)}, & if\ median\ barrier\ is\ not\ present \end{cases}$$

$$CMF_{dw} = e^{b_{dw}(0.01\times DD)}$$

$$CMF_{spd} = e^{b_{spd}(Sfef-PSL)}$$

Where,

CMF_{dw} = Crash Modification Factor for driveway density,
DD = Driveway density.

Table 38 and Table 39 provide calibrated coefficients for FI crashes and PDO crashes respectively.

Table 38. Calibrated Coefficients for FI Crashes on Multi-lane Divided Highways.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-5.083	0.221	-22.98	<.0001
b_6	Adjustment factor for 6 lanes	0.330	0.032	10.15	<.0001
b_8	Adjustment factor for 8+ lanes	0.541	0.113	4.78	<.0001
b_{aadt}	AADT	0.647	0.022	29.78	<.0001
b_{tk}	Truck proportion	-0.017	0.002	-9.18	<.0001
b_{lw}	Lane width (if ≤12ft)	-0.098	0.032	-3.06	0.0022
$b_{lw,2}$	Lane width (if >12ft)	0.100	0.041	2.44	0.0147
b_{isw}	Inside shoulder width (if ≤4t)	-0.030	0.004	-7.84	<.0001
$b_{isw,2}$	Inside shoulder width (if >4ft)	0.110	0.047	2.34	0.0193
b_{osw}	Outside shoulder width (if ≤8ft)	-0.034	0.004	-7.79	<.0001
$b_{osw,2}$	Outside shoulder width (if >8ft)	0.074	0.032	2.33	0.0199
b_{mb}	Median width (if barrier is present)	-0.265	0.051	-5.21	<.0001
b_{mw}	Median width (if barrier is not present)	-0.213	0.030	-7.18	<.0001
b_{dw}	Driveway density	0.016	0.002	6.32	<.0001
b_{spd}	Excess speed	3.463	0.460	7.52	<.0001
k	Inverse dispersion parameter	1.394	0.018	75.92	<.0001

Table 39. Calibrated Coefficients for PDO Crashes on Multi-lane Divided Highways.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-4.723	0.203	-23.27	<.0001
b_6	Adjustment factor for 6 lanes	0.320	0.031	10.23	<.0001
b_8	Adjustment factor for 8+ lanes	0.629	0.111	5.67	<.0001
b_{aadt}	AADT	0.678	0.020	33.94	<.0001
b_{tk}	Truck proportion	-0.011	0.002	-6.5	<.0001
b_{lw}	Lane width (if ≤12ft)	-0.140	0.031	-4.53	<.0001
$b_{lw,2}$	Lane width (if >12ft)	0.130	0.040	3.28	0.001
b_{isw}	Inside shoulder width (if ≤4t)	-0.033	0.004	-9.01	<.0001
$b_{isw,2}$	Inside shoulder width (if >4ft)	0.082	0.046	1.79	0.073
b_{osw}	Outside shoulder width (if ≤8ft)	-0.033	0.004	-8.03	<.0001
$b_{osw,2}$	Outside shoulder width (if >8ft)	0.090	0.030	2.97	0.0029
b_{mb}	Median width (if barrier is present)	-0.190	0.051	-3.74	0.0002
b_{mw}	Median width (if barrier is not present)	-0.255	0.028	-9.07	<.0001
b_{dw}	Driveway density	0.021	0.002	8.85	<.0001
b_{spd}	Excess speed	3.280	0.435	7.54	<.0001
k	Inverse dispersion parameter	1.376	0.016	87.87	<.0001

A comparison of different calibrated multi-lane divided SPFs is shown in Figure 48 and Figure 49 for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the calibrated HSM SPFs and Texas WB. The equations are plotted for the case of all CMFs equal to 1.0 (representing base conditions). It is important to note that the SPFs do not include the same set of base conditions and thus they are not directly comparable to each other. In addition, the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since Texas WB includes SPFs for FI crashes only, the comparison is made just with HSM SPFs for total crashes.

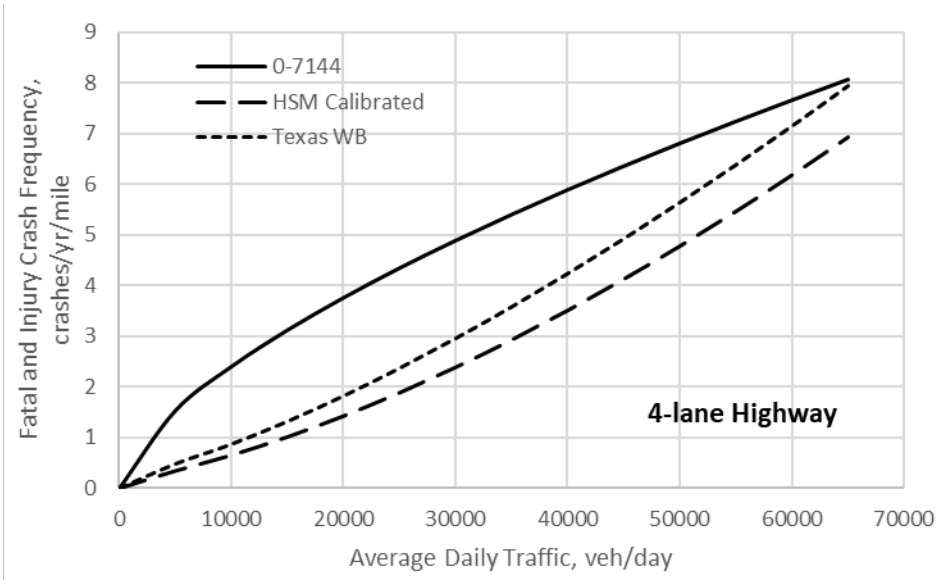


Figure 48. SPF Comparison for FI Crashes on Multi-lane Divided Highways.

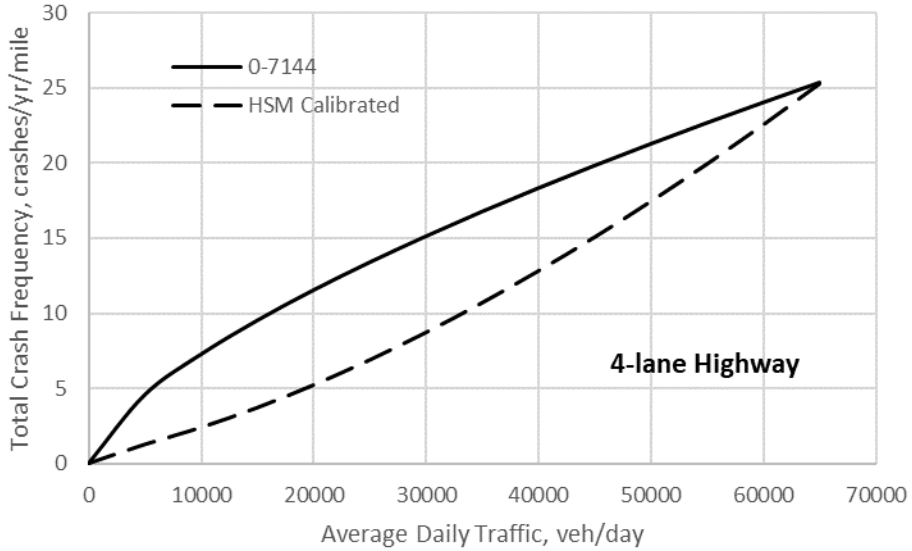


Figure 49. SPF Comparison for Total Crashes on Multi-lane Divided Highways.

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 Texas WB as a means of model validation.

Truck Proportion CMF

The truck proportion CMF is described using Equation 49:

$$CMF_{tk} = e^{-0.017(tk\_perc)} \quad (49)$$

The base condition for this CMF is no trucks in the traffic mix. The truck proportion CMF is shown in Figure 50. The CMF for truck proportion for urban arterials from Texas WB is used for comparison. Both CMFs show similar trends. The CMF shows that the crashes decrease with the increase in truck proportion in traffic. Although this may seem counterintuitive, trucks usually travel on high-standard roads, and that is reflected in this CMF.

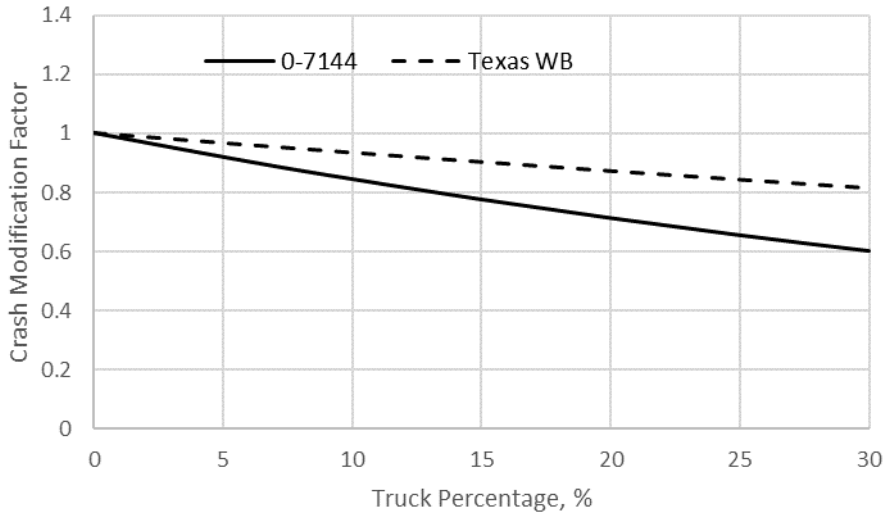


Figure 50. CMF for Truck Proportion on Multi-Lane Divided Highways.

Lane Width CMF

The lane width CMF is described using Equation 50:

$$CMF_{lw} = \begin{cases} e^{-0.098(lw-12)}, & if\ lw \leq 12ft \\ 1.0, & if\ lw > 12ft \end{cases} \quad (50)$$

The base condition for this CMF is a 12-ft lane width. The lane width used in this CMF is an average for all through lanes on the segment. The lane width CMF developed in this study is shown in Figure 51 using a solid trend line. The lane widths used to calibrate this CMF range from 10 to 13 ft. The coefficient for lane widths greater than 12ft is found to be marginally significant and counterintuitive, so a value of 1.0 is used. Also shown in Figure 51 is the CMF presented in the Texas WB. Broken lines are used to differentiate this CMF from the one proposed in this research project. It is important to note that the HSM does not include a CMF for lane width for urban arterials. The proposed CMF is shown to be more sensitive to lane width than the CMF in the Texas WB.

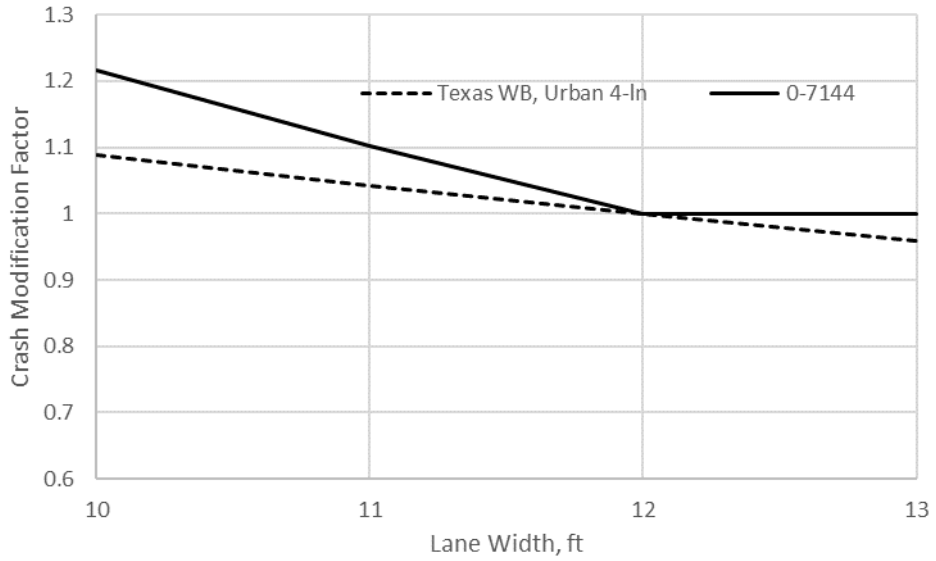


Figure 51. CMF for Lane Width on Multi-Lane Divided Highways.

Inside Shoulder Width CMF

The inside shoulder width CMF is described using Equation 51:

$$CMF_{isw} = \begin{cases} e^{-0.030\times(isw-4)}, & if\ isw \leq 4ft \\ 1.0, & if\ isw > 4ft \end{cases} \quad (51)$$

The base condition for this CMF is a 4-ft inside shoulder width. The width used in this CMF is an average for inside shoulders in both directions. The inside shoulder width CMF developed in this study is shown in Figure 52 using a solid trend line. The coefficient for inside shoulder widths greater than 4ft is found to be marginally significant and counterintuitive, so a value of 1.0 is used. The inside shoulder widths used to calibrate this CMF range from 2 to 10 ft. Also shown in Figure 52 is the CMF presented in the Texas WB. It is important to note that the HSM does not include a CMF for inside shoulder width for urban arterials. The proposed CMF tracks well with the CMF in the Texas WB.

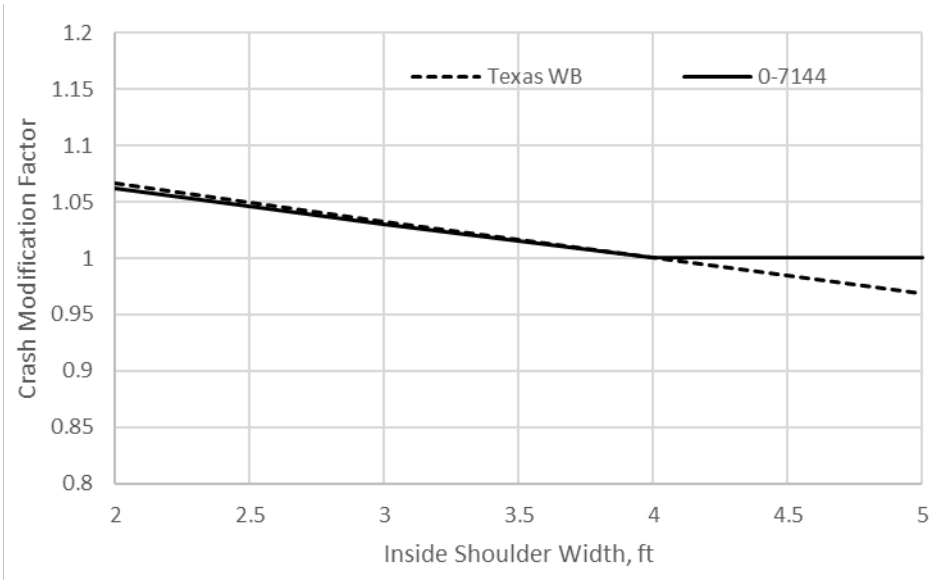


Figure 52. CMF for Inside Shoulder Width on Multi-lane Divided Highways.

Outside Shoulder Width CMF

The outside shoulder width CMF is described using Equation 52:

$$CMF_{osw} = \begin{cases} e^{-0.034\times(osw-8)}, & if\ osw \leq 8ft \\ 1.0, & if\ osw > 8ft \end{cases} \quad (52)$$

The base condition for this CMF is an 8-ft outside shoulder width. The width used in this CMF is an average for outside shoulders in both directions. The outside shoulder width CMF developed in this study is shown in Figure 53 using a solid trend line. The outside shoulder widths used to calibrate this CMF range from 6 to 12 ft. The coefficient for outside shoulder widths greater than 8ft is found to be marginally significant and counterintuitive, so a value of 1.0 is used. Also shown in Figure 53 are CMF presented in the Texas WB. Broken lines are used to differentiate this CMF from the one proposed in this research project. It is important to note that the HSM does not include a CMF for outside shoulder width for urban arterials. The proposed CMF is more sensitive to the outside shoulder width than the Texas WB CMF.

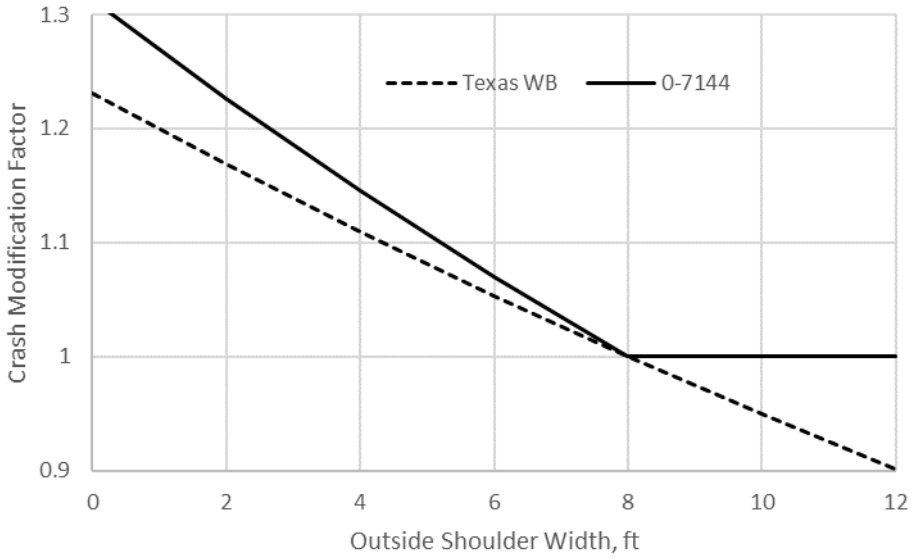


Figure 53. CMF for Outside Shoulder Width on Multi-lane Divided Highways.

Median Width CMF

The median width CMF is described using Equation 53.

$$CMF_{mw} = \begin{cases} e^{-0.265\times0.01\times(mw-30)}, & if\ median\ barrier\ is\ present \\ e^{-0.213\times0.01\times(mw-30)}, & if\ median\ barrier\ is\ not\ present \end{cases} \quad (53)$$

The base condition for this CMF is a 30-ft median width. The median width CMF is shown in Figure 54 using a solid trend line for median with barriers and with dotted lines for median without barriers. The CMF proposed in this research is compared with the CMF in HSM. The proposed CMFs are more sensitive to the median width than the CMF presented in the HSM.

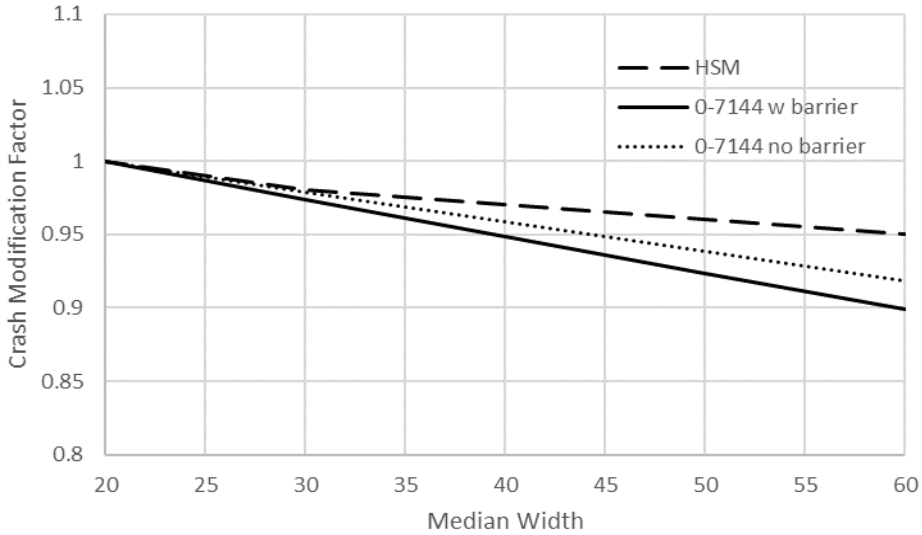


Figure 54. CMF for Median Width on Multi-lane Divided Highways.

Excess Speed CMF

The excess speed CMF is described using Equation 54.

$$CMF_{spd} = e^{3.463 \times 0.01 \times (Sfef - PSL)} \quad (54)$$

The base condition for this CMF varies according to the posted speed limit. Figure 55 shows the CMF for excess speeds on multi-lane divided highways. Since the operating speed CMF does not exist in HSM or Texas WB, a comparison could not be made. The CMF shows that exceeding the PSL by 10 mph increases the crashes by 40%. Over speeding has a more pronounced effect on multi-lane divided arterials than on freeways.

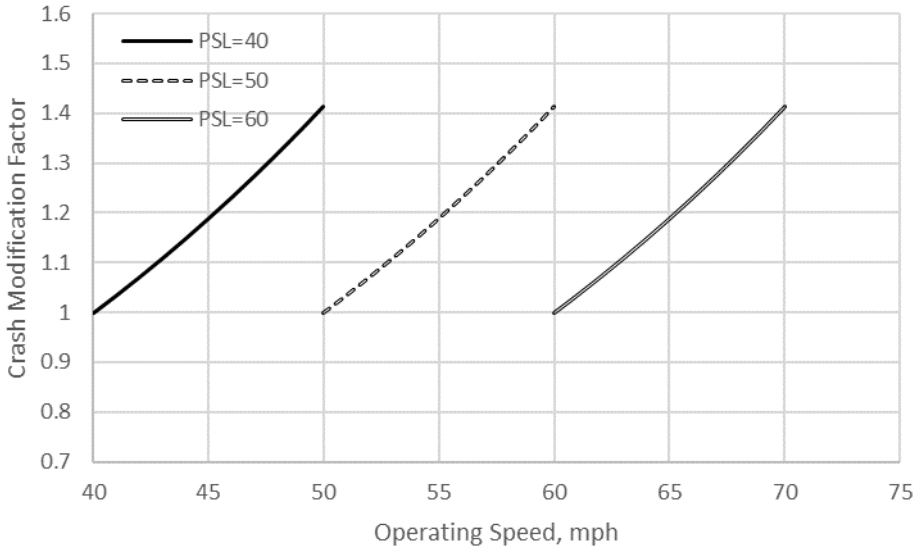


Figure 55. CMF for Excess Speeds on Multi-Lane Divided Highways.

Driveway CMF

The driveway CMF is described using Equation 55.

$$CMF_{std} = e^{0.016(0.1 \times DD)} \quad (55)$$

The base condition for this CMF is no driveways. The driveway CMF is shown in Figure 56 using a solid trend line for residential driveway density. The Project Team developed an equation to convert industrial and commercial driveways into equivalent residential driveways based on traffic volumes in their previous research (Geedipally et al., 2021). Mainly, it was found that one industrial driveway is equivalent to 3 residential and one commercial driveway is equivalent to 12 residential driveways.

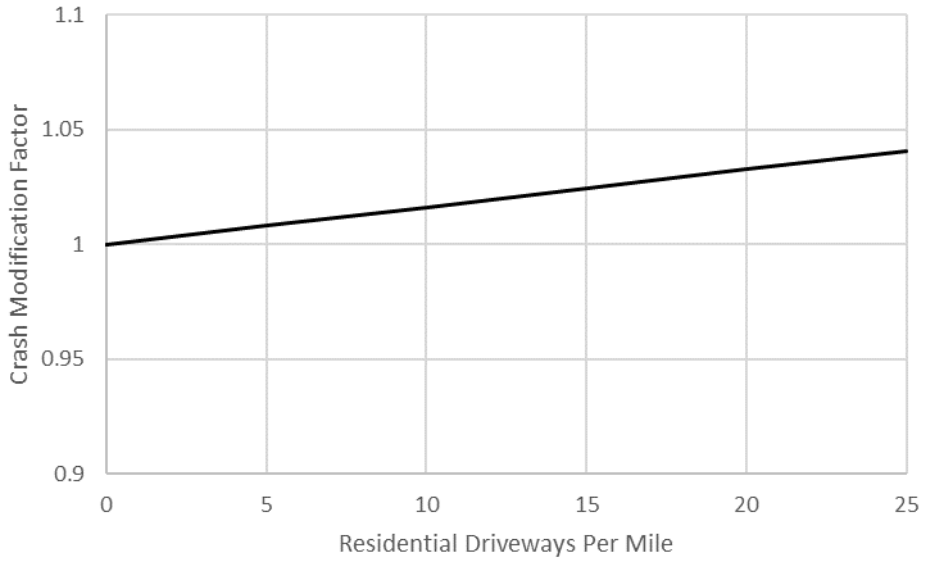


Figure 56. CMF for Driveway Density on Multi-lane Divided Highways.

4.3.3 SPFs for Multi-lane Undivided Highways

For the SPF development, the team just considered the four-lane highways, given the large sample size in that category. The predicted crash frequency is calculated as follows in Equation (56).

$$N = L \times y \times e^{b_0 + b_{aadt} \ln(AADT)} \times CMF_{tk} \times CMF_{lw} \times CMF_{sw} \times CMF_{spd} \times CMF_{dw} \times CMF_{pre}; \quad (56)$$

With,

$$CMF_{lw} = e^{b_{lw}(lw-12)}$$
$$CMF_{sw} = e^{b_{sw}(sw-6)}$$
$$CMF_{spd} = e^{b_{spd}(Sfef-PSL)}$$
$$CMF_{dw} = e^{b_{dw}(0.01 \times DD)}$$
$$CMF_{pre} = e^{b_{pre}(pre) \times 0.1}$$

Where,

CMF_{sw} = Crash Modification Factor for average shoulder width,
CMF_{pre} = Crash Modification Factor for average daily precipitation,
sw = Average shoulder width, feet,
pre = Average yearly precipitation, inches,

Table 40 and Table 41 provide calibrated coefficients for FI crashes and PDO crashes, respectively.

Table 40. Calibrated Coefficients for FI Crashes on Multi-lane Undivided Highways.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-6.679	0.168	-39.75	<.0001
b_{aadt}	AADT	0.863	0.017	50.28	<.0001
b_{tk}	Truck proportion	-0.017	0.002	-10.49	<.0001
b_{lw}	Lane width	-0.015	0.013	-1.14	0.2528
b_{sw}	Shoulder width	-0.020	0.002	-9.32	<.0001
b_{spd}	Excess speed	0.025	0.003	8.1	<.0001
b_{dw}	Driveway density	0.011	0.002	5.66	<.0001
b_{pre}	Precipitation	0.076	0.019	4	<.0001
k	Inverse dispersion parameter	1.563	0.016	99.5	<.0001

Table 41. Calibrated Coefficients for PDO Crashes on Multi-lane Undivided Highways.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-5.829	0.156	-37.4	<.0001
b_{aadt}	AADT	0.857	0.016	53.71	<.0001
b_{tk}	Truck proportion	-0.005	0.002	-3.48	0.0005
b_{lw}	Lane width	-0.069	0.013	-5.34	<.0001
b_{sw}	Shoulder width	-0.020	0.002	-9.57	<.0001
b_{spd}	Excess speed	0.018	0.003	5.87	<.0001
b_{dw}	Driveway density	0.011	0.002	6.17	<.0001
b_{pre}	Precipitation	0.007	0.018	0.38	0.7072
k	Inverse dispersion parameter	1.520	0.013	117.21	<.0001

A comparison of different calibrated multi-lane undivided SPFs is shown in Figure 57 and Figure 58 for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the HSM-calibrated SPFs and Texas WB. The equations are plotted for the case of all CMFs equal to 1.0 (representing base conditions). It is important to note that the SPFs do not include the same set of base conditions and thus they are not directly comparable to each other. In addition, the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since Texas WB includes SPFs for FI crashes only, the comparison is made just with HSM SPFs for total crashes.

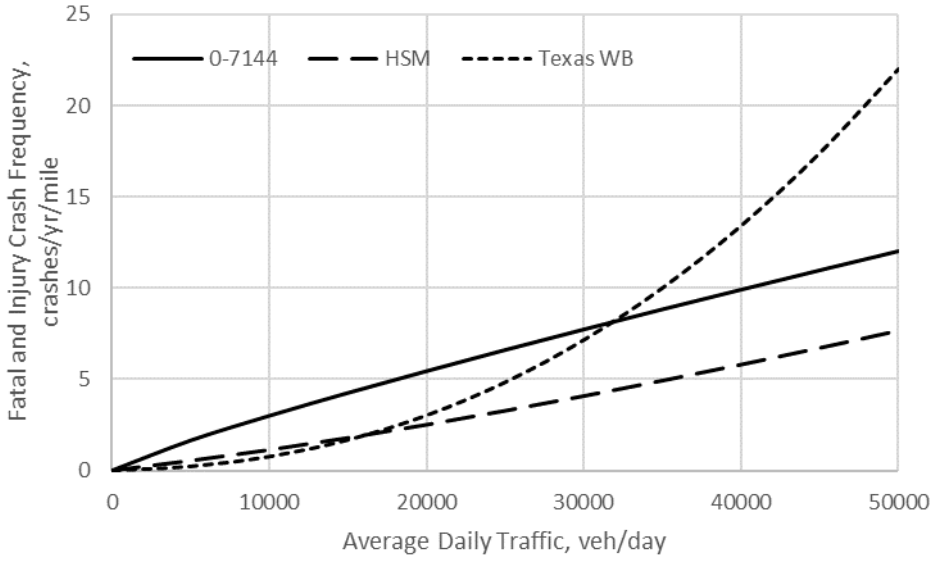


Figure 57. SPF comparison for Fatal and Injury Crashes on Multi-lane Undivided Highways.

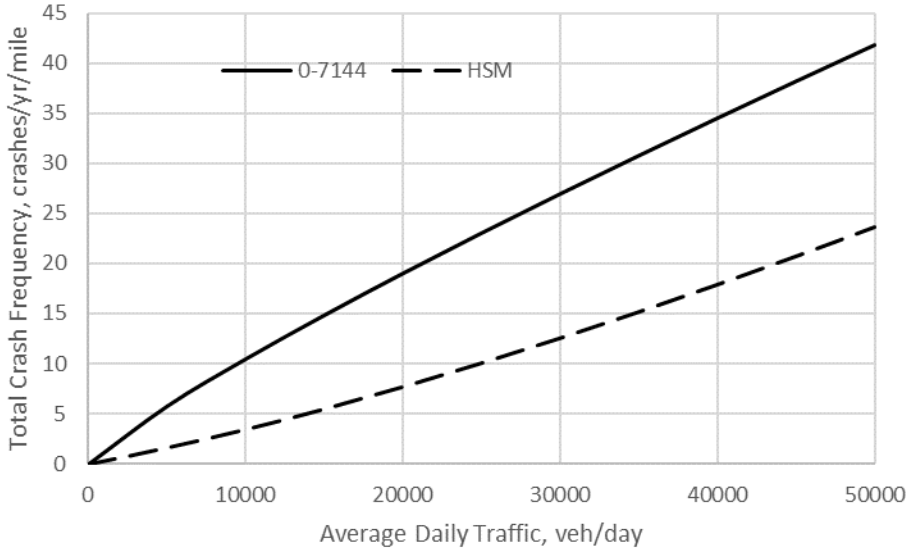


Figure 58. SPF comparison for Total Crashes on Multi-lane Undivided Highways.

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 HSM and Texas WB as a means of model validation.

Truck Proportion CMF

The truck proportion CMF is described using Equation 57:

$$CMF_{tk} = e^{-0.017(tk\_perc)} \quad (57)$$

The base condition for this CMF is no trucks in the traffic mix. The truck proportion CMF is shown in Figure 59. The CMF for truck proportion for urban arterials from Texas WB is used for comparison. Both CMFs show similar trends. The CMF shows that the crashes decrease with the increase in truck proportion in traffic. Although this may seem counterintuitive, trucks usually travel on high-standard roads.

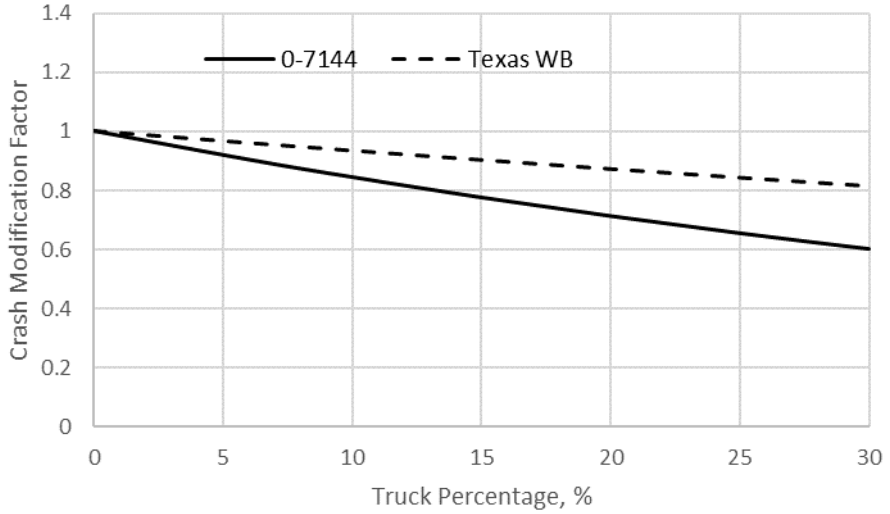


Figure 59. CMF for Truck Proportion on Multi-Lane Undivided Highways.

Lane Width CMF

The lane width CMF is described using Equation 58:

$$CMF_{lw} = e^{-0.015(lw-12)} \quad (58)$$

The base condition for this CMF is a 12-ft lane width. The lane width used in this CMF is an average for all through lanes on the segment. The lane width CMF developed in this study is shown in Figure 60 using a solid trend line. The lane widths used to calibrate this CMF range from 10 to 13 ft. The coefficient for lane widths greater than 12ft is found to be marginally significant and counterintuitive, so a value of 1.0 is used. Also shown in Figure 60 is the CMF presented in the Texas WB. Broken lines are used to differentiate this CMF from the one proposed in this research project. It should be noted that the HSM does not include CMF for lane width for urban arterials. The proposed CMF is shown to be less sensitive to lane width than the CMF in the Texas WB.

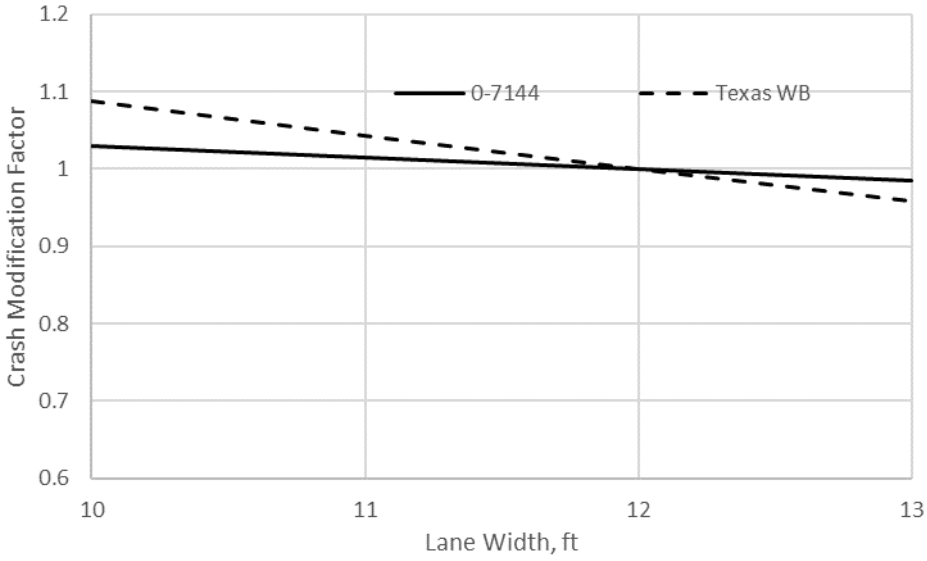


Figure 60. CMF for Lane Width on Multi-Lane Undivided Highways.

Shoulder Width CMF

The shoulder width CMF is described using Equation 59:

$$CMF_{isw} = e^{-0.020\times(sw-6)} \quad (59)$$

The base condition for this CMF is a 6-ft shoulder width. The width used in this CMF is an average for shoulders in both directions. The shoulder width CMF developed in this study is shown in Figure 61 using a solid trend line. The shoulder widths used to calibrate this CMF range from 0 to 8 ft. Also shown in Figure 61 is the CMF presented in the Texas WB. It should be noted that HSM does not include CMF for shoulder width for urban arterials. The proposed CMF is shown to be less sensitive to shoulder width than the CMF in the Texas WB.

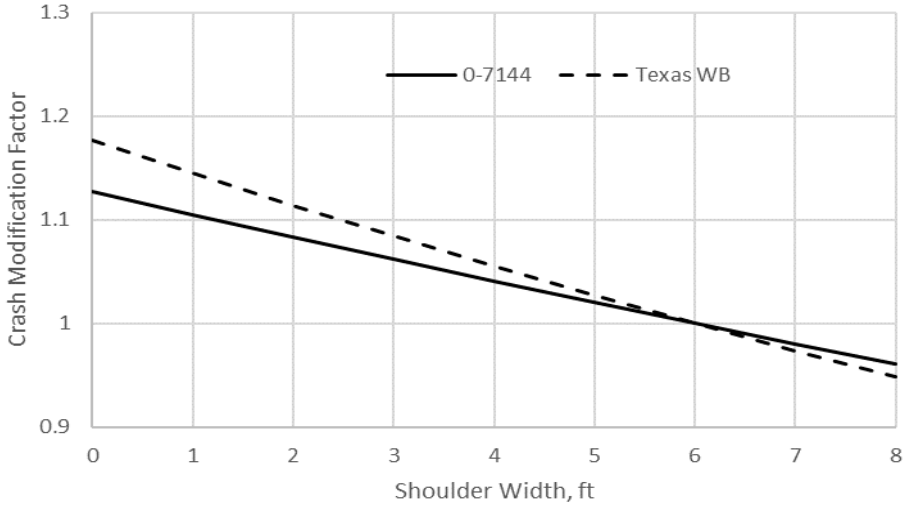


Figure 61. CMF for Shoulder Width on Multi-lane Undivided Highways.

Excess Speed CMF

The excess speed CMF is described using Equation 60.

$$CMF_{spd} = e^{0.025\times(Sfef-PSL)} \quad (60)$$

The base condition for this CMF varies according to the posted speed limit. Figure 62 shows the CMF for excess speeds on multi-lane undivided highways. Since the operating speed CMF does not exist in HSM or Texas WB, a comparison could not be made. The CMF shows that exceeding the PSL by 10 mph increases the crashes by 30%.

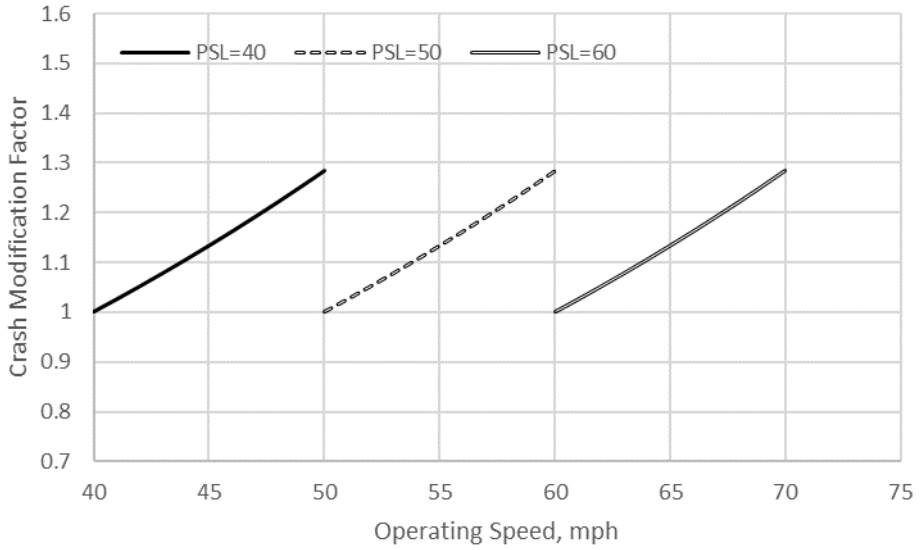


Figure 62. CMF for Excess Speeds on Multi-Lane Undivided Highways.

Driveway CMF

The driveway CMF is described using Equation 61.

$$CMF_{std} = e^{0.011(0.1\times DD)} \quad (61)$$

The base condition for this CMF is no driveways. The driveway CMF is shown in Figure 63 for residential driveway density.

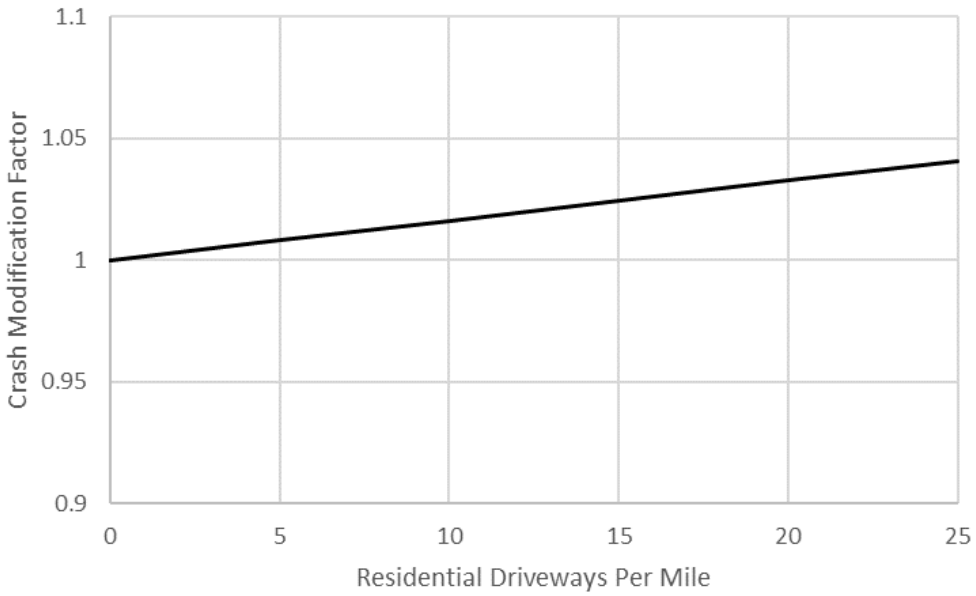


Figure 63. CMF for Driveway Density on Multi-lane Undivided Highways.

Precipitation CMF

The precipitation CMF is described using Equation 62.

$$CMF_{pre} = e^{0.076(0.1 \times pre)} \quad (62)$$

The base condition for this CMF is an annual average precipitation of 0 inches. Since the precipitation CMF does not exist in HSM or Texas WB, a comparison could not be made. The precipitation CMF is shown in Figure 64 using a solid trend line. The CMF shows that precipitation has a greater influence on the occurrence of crashes. For instance, every 1 inch of rainfall increases crashes by 1 percent.

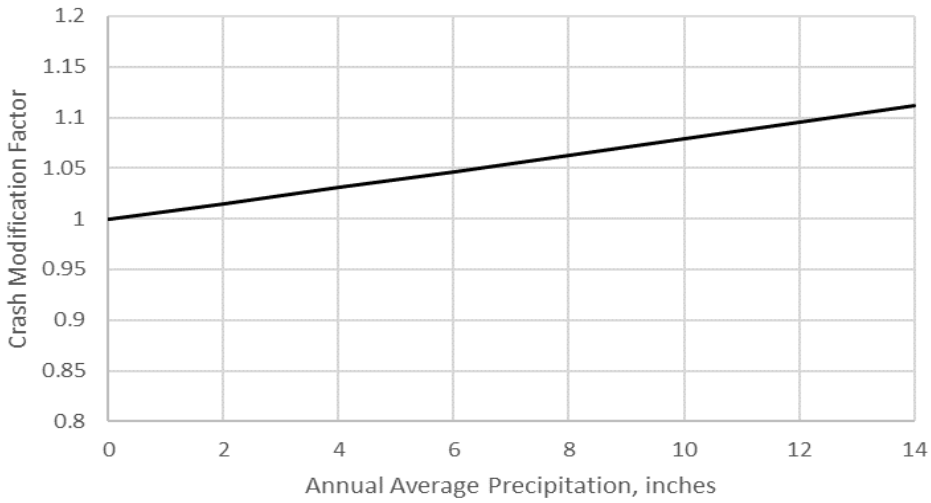


Figure 64. CMF for Precipitation on Multi-lane Undivided Highways.

4.3.4 SPFs for Multi-lane Undivided Highways with Continuous Left Turn Lane

For the SPF development, the team considered all multi-lane highways that had a continuous left turn lane. The predicted crash frequency is calculated as follows.

$$N = L \times y \times e^{b_0 + b_{aadt} \ln(AADT)} \times CMF_{lw} \times CMF_{sw} \times CMF_{spd} \times CMF_{dw}; \quad (63)$$

With,

$$CMF_{lw} = e^{b_{lw}(lw-12)}$$
$$CMF_{sw} = e^{b_{sw}(sw-6)}$$
$$CMF_{spd} = e^{b_{spd}(Sfef-PSL)}$$
$$CMF_{dw} = e^{b_{dw}(0.01 \times DD)}$$

Table 42 and Table 43 provide calibrated coefficients for FI crashes and PDO crashes, respectively.

Table 42. Calibrated Coefficients for FI Crashes on Multi-lane Undivided Highways with Continuous Turn Lane.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-9.029	0.550	-16.41	<.0001
b_{aadt}	AADT	1.030	0.056	18.30	<.0001
b_{lw}	Lane width	-0.028	0.011	-2.52	0.0117
b_{sw}	Shoulder width	-0.056	0.006	-8.69	<.0001
b_{spd}	Excess speed	0.038	0.012	3.18	0.0015
b_{dw}	Driveway density	0.030	0.006	4.92	<.0001
k	Inverse dispersion parameter	1.671	0.049	34.19	<.0001

Table 43. Calibrated Coefficients for PDO Crashes on Multi-lane Undivided Highways with Continuous Turn Lane.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-8.666	0.532	-16.28	<.0001
b_{aadt}	AADT	1.061	0.055	19.44	<.0001
b_{lw}	Lane width	-0.040	0.010	-3.91	<.0001
b_{sw}	Shoulder width	-0.057	0.006	-9.15	<.0001
b_{spd}	Excess speed	0.025	0.012	2.09	0.0366
b_{dw}	Driveway density	0.037	0.006	6.52	<.0001
k	Inverse dispersion parameter	1.656	0.041	39.99	<.0001

A comparison of different calibrated multi-lane undivided with continuous turn lane SPFs is shown in Figure 65 and Figure 66 for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the calibrated HSM SPFs and Texas WB. The equations are plotted for the case of all CMFs equal to 1.0 (representing base conditions). It is important to note that the SPFs do not include the same set of base conditions and thus they are not directly comparable to each other. In addition, the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since Texas WB includes SPFs for FI crashes only, the comparison is made just with HSM SPFs for total crashes.

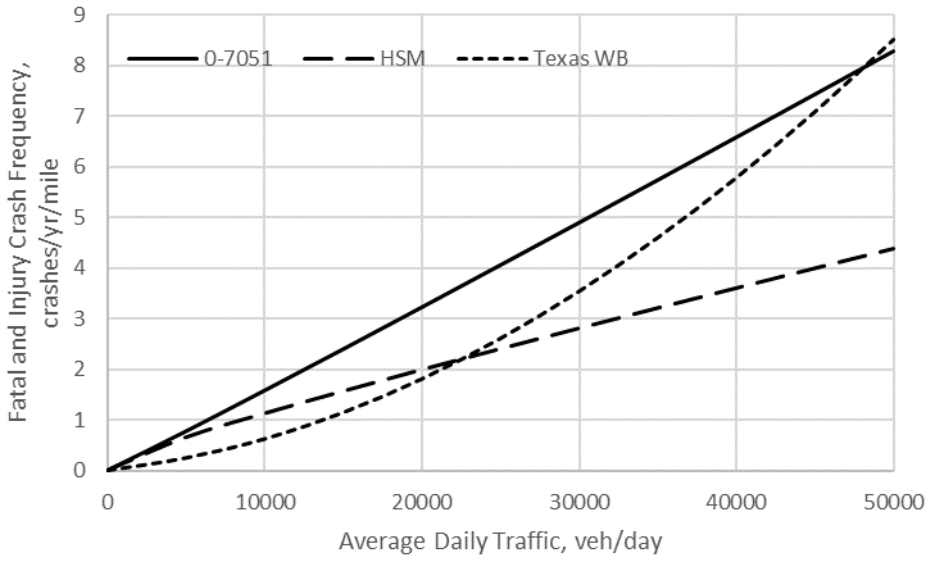


Figure 65. SPF comparison for Fatal and Injury Crashes on Multi-lane Undivided Highways with Continuous Left Turn Lane.

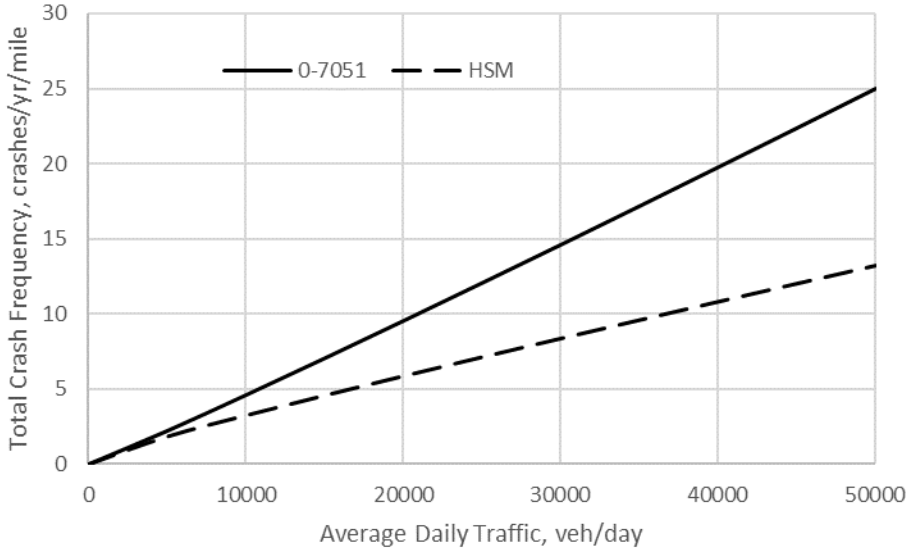


Figure 66. SPF comparison for Total Crashes on Multi-lane Undivided Highways with Continuous Left Turn Lane.

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 HSM and Texas WB as a means of model validation.

Lane Width CMF

The lane width CMF is described using Equation 64:

$$CMF_{lw} = e^{-0.015(lw-12)} \quad (64)$$

The base condition for this CMF is a 12-ft lane width. The lane width used in this CMF is an average for all through lanes on the segment. The lane width CMF developed in this study is shown in Figure 67 using a solid trend line. The lane widths used to calibrate this CMF range from 10 to 13 ft. Also shown in Figure 67 is the CMF presented in the Texas WB. Broken lines are used to differentiate these CMFs from the one proposed in this research project. It should be noted that the HSM does not include CMF for the lane width for urban arterials. The proposed CMF is shown to be less sensitive to lane width than the CMF in the Texas WB.

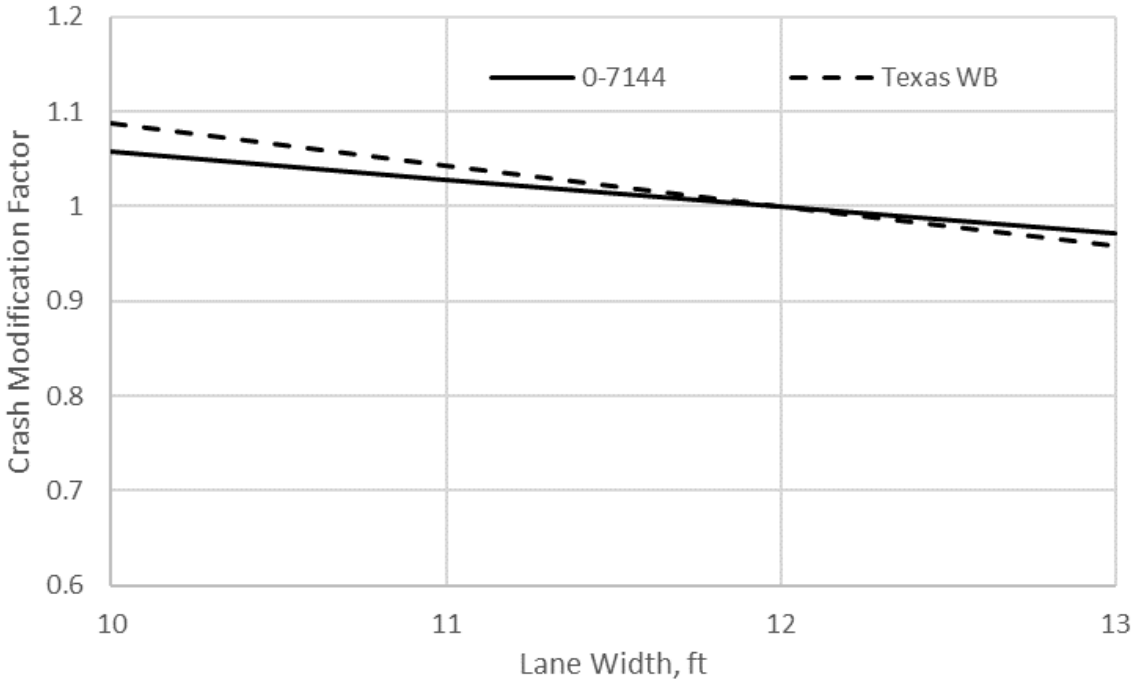


Figure 67. CMF for Lane Width on Multi-Lane Undivided Highways with Continuous Left Turn Lane.

Shoulder Width CMF

The shoulder width CMF is described using Equation 65:

$$CMF_{isw} = e^{-0.056\times(sw-6)} \quad (65)$$

The base condition for this CMF is a 6-ft shoulder width. The width used in this CMF is an average for shoulders in both directions. The shoulder width CMF developed in this study is shown in Figure 68 using a solid trend line. The inside shoulder widths used to calibrate this CMF range from 0 to 8 ft. Also shown in Figure 68 is the CMF presented in the Texas WB. It

should be noted that the HSM does not include CMF for the shoulder width for urban arterials. The proposed CMF is shown to be less sensitive to shoulder width than the CMF in the Texas WB.

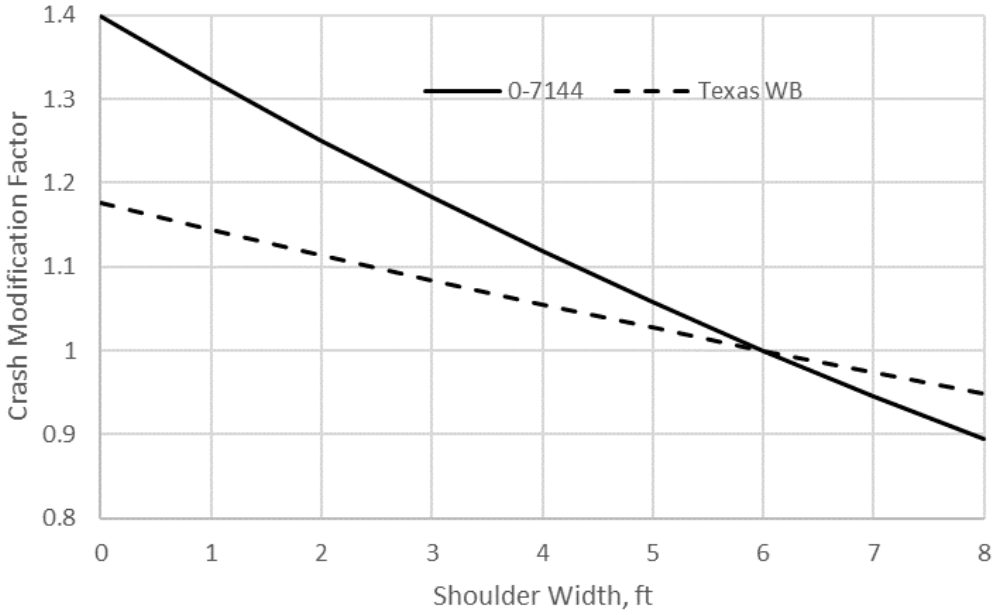


Figure 68. CMF for Shoulder Width on Multi-lane Undivided Highways with Continuous Left Turn Lane.

Excess Speed CMF

The excess speed CMF is described using Equation 66.

$$CMF_{spd} = e^{0.038\times(Sfef-PSL)} \quad (66)$$

The base condition for this CMF varies according to the posted speed limit. Figure 69 shows the CMF for excess speeds on multi-lane, undivided highways with a continuous left turn lane. Since the operating speed CMF does not exist in HSM or Texas WB, a comparison could not be made. The CMF shows that exceeding the PSL by 10 mph increases the crashes by 45%.

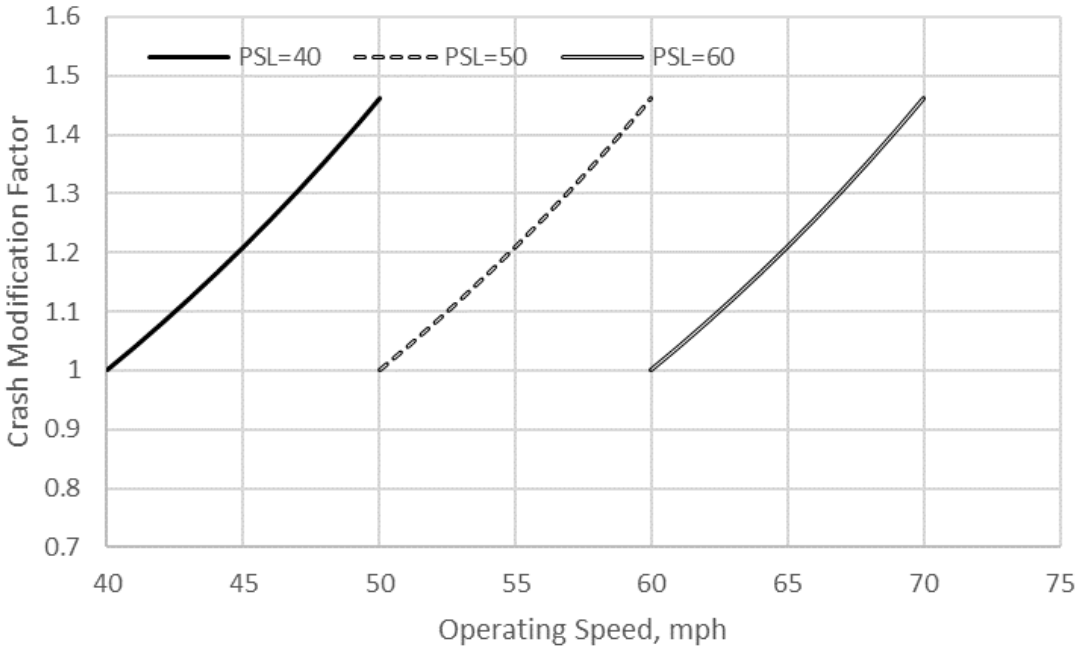


Figure 69. CMF for Excess Speeds on Multi-Lane Undivided Highways with Continuous Left Turn Lane.

Driveway CMF

The driveway CMF is described using Equation 67.

$$CMF_{std} = e^{0.03(0.1 \times DD)} \tag{67}$$

The base condition for this CMF is no driveways. The driveway CMF is shown in Figure 70 for residential driveway density.

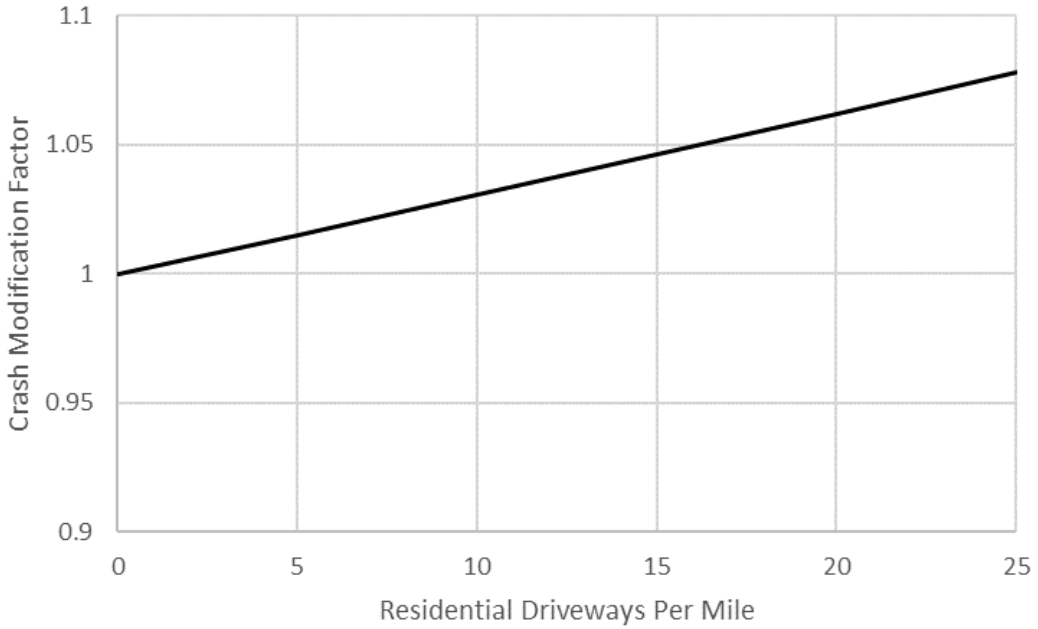


Figure 70. CMF for Driveway Density on Multi-lane Undivided Highways with Continuous Left Turn Lane.

SPFs for Two-lane Undivided Highways with Continuous Left Turn Lane

For the SPF development, the team considered all two-lane highways that had a continuous left turn lane. The predicted crash frequency is calculated as follows.

$$N = L \times y \times e^{b_0 + b_{aadt} \ln(AADT)} \times CMF_{sw} \times CMF_{spd} \times CMF_{dw}; \quad (68)$$

With,

$$CMF_{sw} = e^{b_{sw}(sw-6)}$$
$$CMF_{spd} = e^{b_{spd}(Sfef-PSL)}$$
$$CMF_{dw} = e^{b_{dw}(0.01 \times DD)}$$

Table 44 and Table 45 provide calibrated coefficients for FI crashes and PDO crashes, respectively.

Table 44. Calibrated Coefficients for FI Crashes on Two-lane Undivided Highways with Continuous Turn Lane.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-7.946	1.365	-5.82	<.0001
b_{aadt}	AADT	0.895	0.146	6.14	<.0001
b_{sw}	Shoulder width	-0.036	0.020	-1.84	0.0664
b_{spd}	Excess speed	0.018	0.018	0.98	0.3279
b_{dw}	Driveway density	0.048	0.014	3.56	0.0004
k	Inverse dispersion parameter	1.655	0.139	11.88	<.0001

Table 45. Calibrated Coefficients for PDO Crashes on Two-lane Undivided Highways with Continuous Turn Lane.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-6.349	1.183	-5.37	<.0001
b_{aadt}	AADT	0.808	0.128	6.33	<.0001
b_{sw}	Shoulder width	-0.031	0.017	-1.82	0.0698
b_{spd}	Excess speed	--	--	--	--
b_{dw}	Driveway density	0.050	0.012	4.04	<.0001
k	Inverse dispersion parameter	1.746	0.109	16.01	<.0001

A comparison of different calibrated two-lane undivided with continuous turn lane SPFs is shown in Figure 71 and Figure 72 for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the calibrated HSM SPFs and Texas WB. The equations are plotted for the case of all CMFs equal to 1.0 (representing base conditions). It is important to note that the SPFs do not include the same set of base conditions and thus they are not directly comparable to each other. In addition, the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since Texas WB includes SPFs for FI crashes only, the comparison is made just with HSM SPFs for total crashes.

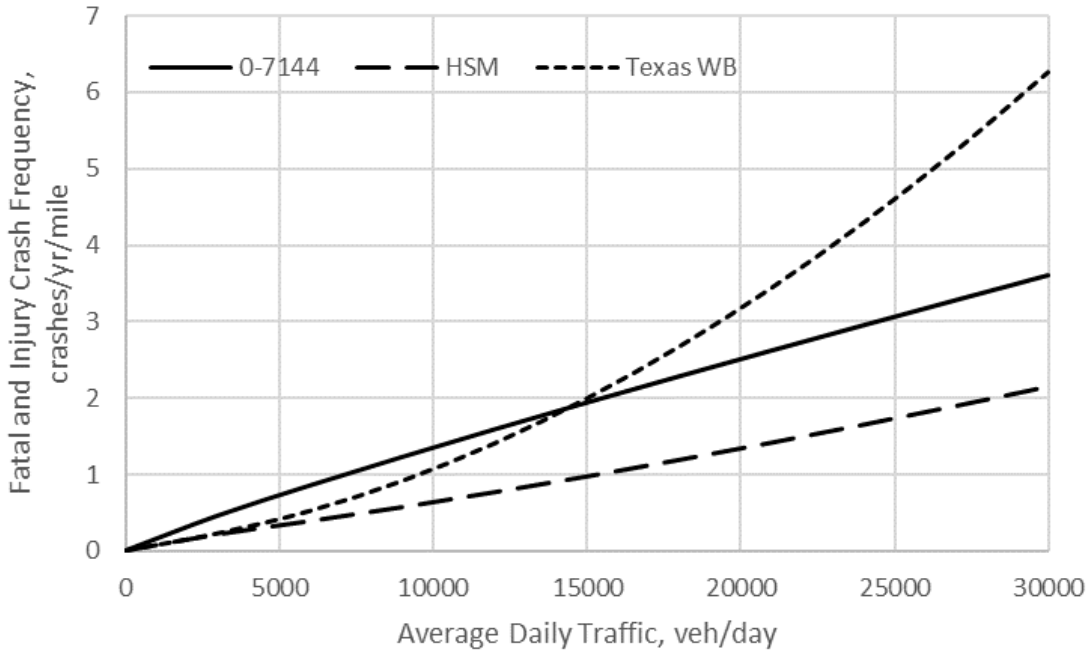


Figure 71. SPF comparison for Fatal and Injury Crashes on Two-lane Undivided Highways with Continuous Left Turn Lane.

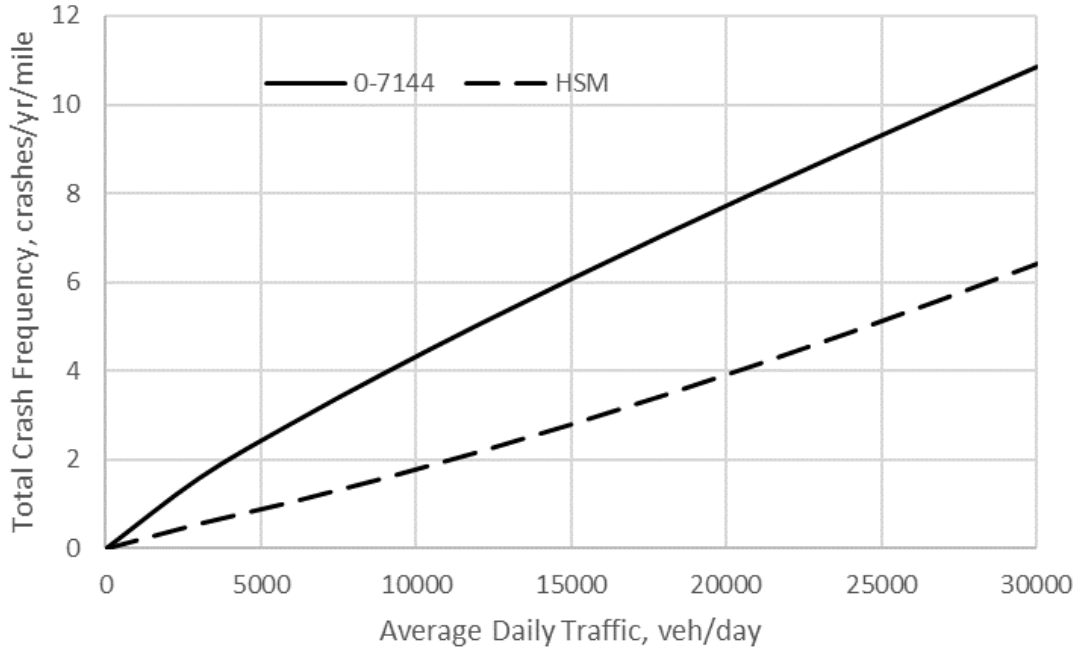


Figure 72. SPF comparison for Total Crashes on Two-lane Undivided Highways with Continuous Left Turn Lane.

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 HSM and Texas WB as a means of model validation.

Shoulder Width CMF

The shoulder width CMF is described using Equation 69:

$$CMF_{sw} = e^{-0.036\times(sw-6)} \quad (69)$$

The base condition for this CMF is a 6-ft shoulder width. The width used in this CMF is an average for shoulders in both directions. The shoulder width CMF developed in this study is shown in Figure 73 using a solid trend line. The inside shoulder widths used to calibrate this CMF range from 0 to 8 ft. Also shown in Figure 73 is the CMF presented in the Texas WB. The proposed CMF is shown to be less sensitive to shoulder width than the CMF in the Texas WB.

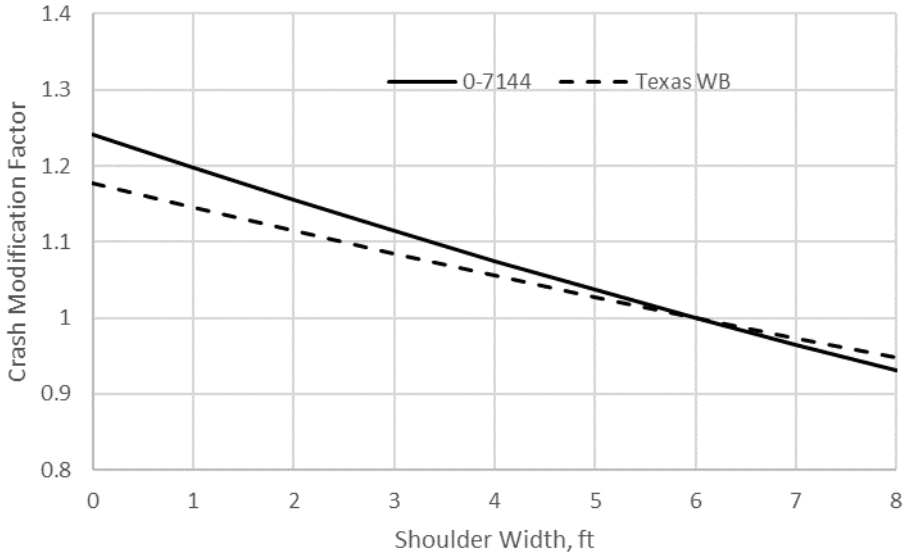


Figure 73. CMF for Shoulder Width on Two-lane Undivided Highways with Continuous Left Turn Lane.

Excess Speed CMF

The excess speed CMF is described using Equation 70.

$$CMF_{spd} = e^{0.018\times(Sfef-PSL)} \quad (70)$$

The base condition for this CMF varies according to the posted speed limit. Figure 74 shows the CMF for excess speeds on multi-lane undivided highways. Since the operating speed CMF does

not exist in HSM or Texas WB, a comparison could not be made. The CMF shows that exceeding the PSL by 10 mph increases the crashes by 20%.

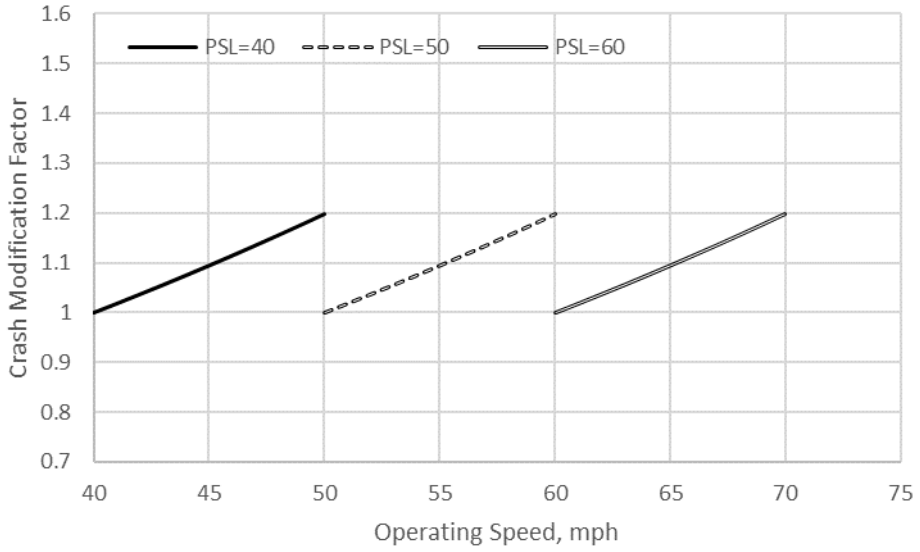


Figure 74. CMF for Excess Speeds on Two-Lane Undivided Highways with Continuous Left Turn Lane.

Driveway CMF

The driveway CMF is described using Equation 71.

$$CMF_{std} = e^{0.03(0.1 \times DD)} \tag{71}$$

The base condition for this CMF is no driveways. The driveway CMF is shown in Figure 75 for residential driveway density.

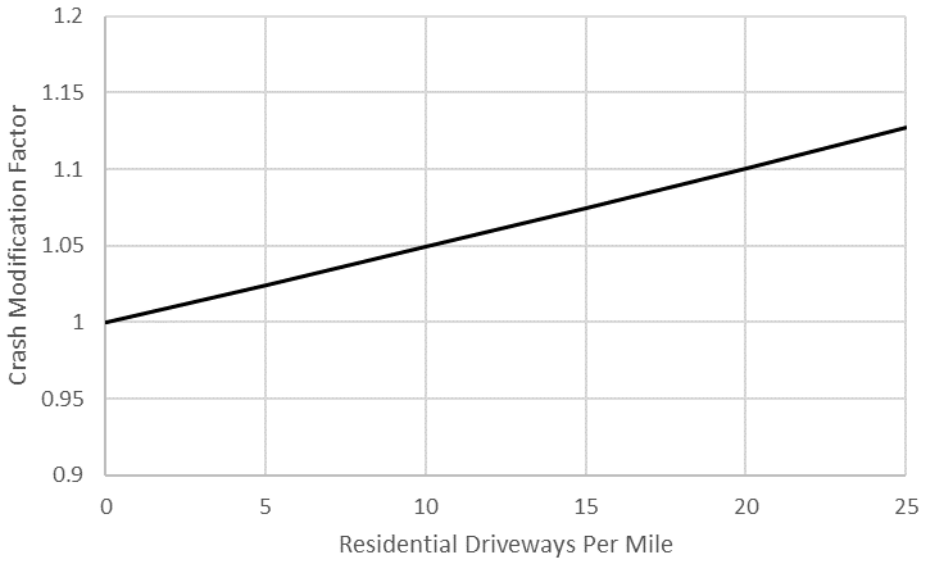


Figure 75. CMF for Driveway Density on Two-lane Undivided Highways with Continuous Turn Lane.

4.3.5 SPFs for Two-lane Highways

The predicted crash frequency is calculated as follows.

$$N = L \times y \times e^{b_0 + b_{aadt} \ln(AADT)} \times CMF_{tk} \times CMF_{lw} \times CMF_{sw} \times CMF_{spd} \times CMF_{dw} \times CMF_{pre}; \quad (72)$$

With,

$$CMF_{lw} = e^{b_{lw}(lw-12)}$$
$$CMF_{sw} = e^{b_{sw}(sw-6)}$$
$$CMF_{spd} = e^{b_{spd}(Sfef-PSL)}$$
$$CMF_{dw} = e^{b_{dw}(0.01 \times DD)}$$
$$CMF_{pre} = e^{b_{pre}(pre) \times 0.1}$$

Table 46 and Table 47 provide calibrated coefficients for FI crashes and PDO crashes, respectively.

Table 46. Calibrated Coefficients for FI Crashes on Two-lane Highways.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-6.611	0.147	-45.07	<.0001
b_{aadt}	AADT	0.797	0.016	48.68	<.0001
b_{tk}	Truck proportion	-0.006	0.002	-3.55	0.0004
b_{lw}	Lane width	-0.025	0.043	-0.59	0.5523
b_{sw}	Shoulder width	-0.010	0.003	-3.30	0.001
b_{spd}	Excess speed	--	--	--	--
b_{dw}	Driveway density	0.034	0.003	10.82	<.0001
b_{pre}	Precipitation	0.114	0.021	5.34	<.0001
k	Inverse dispersion parameter	1.230	0.022	56.13	<.0001

Table 47. Calibrated Coefficients for PDO Crashes on Two-lane Highways.

Coefficient	Variable	Value	Std. Dev	t-statistic	p-value
b_0	Intercept	-6.102	0.132	-46.18	<.0001
b_{aadt}	AADT	0.834	0.015	57.22	<.0001
b_{tk}	Truck proportion	-0.002	0.001	-1.15	0.2497
b_{lw}	Lane width	0.090	0.014	6.28	<.0001
b_{sw}	Shoulder width	-0.023	0.003	-7.81	<.0001
b_{spd}	Excess speed	-0.007	0.004	-1.59	0.1111
b_{dw}	Driveway density	0.042	0.003	14.39	<.0001
b_{pre}	Precipitation	0.017	0.020	0.88	0.3814
k	Inverse dispersion parameter	1.164	0.016	72.03	<.0001

A comparison of different calibrated multi-lane undivided SPFs is shown in Figure 76 and Figure 77 for FI crashes and total crashes, respectively. The SPFs developed in this project are compared with the calibrated HSM SPFs and Texas WB. The equations are plotted for the case of all CMFs equal to 1.0 (representing base conditions). It is important to note that the SPFs do not include the same set of base conditions and thus they are not directly comparable to each other. In addition, the Texas WB SPFs are not calibrated to the current time period. The SPFs are shown for illustration purposes only. Since Texas WB includes SPFs for FI crashes only, the comparison is made just with HSM SPFs for total crashes.

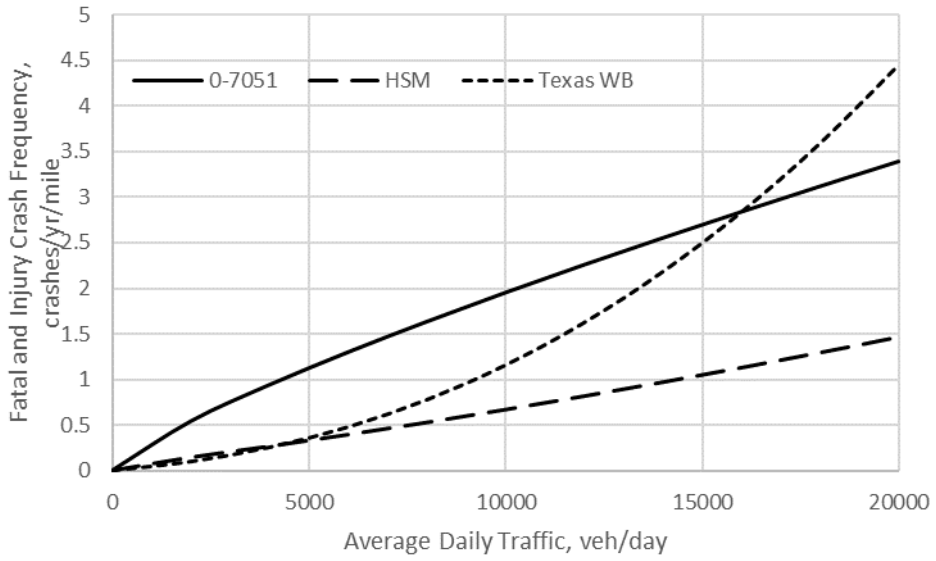


Figure 76. SPF comparison for FI Crashes on Two-lane Undivided Highways.

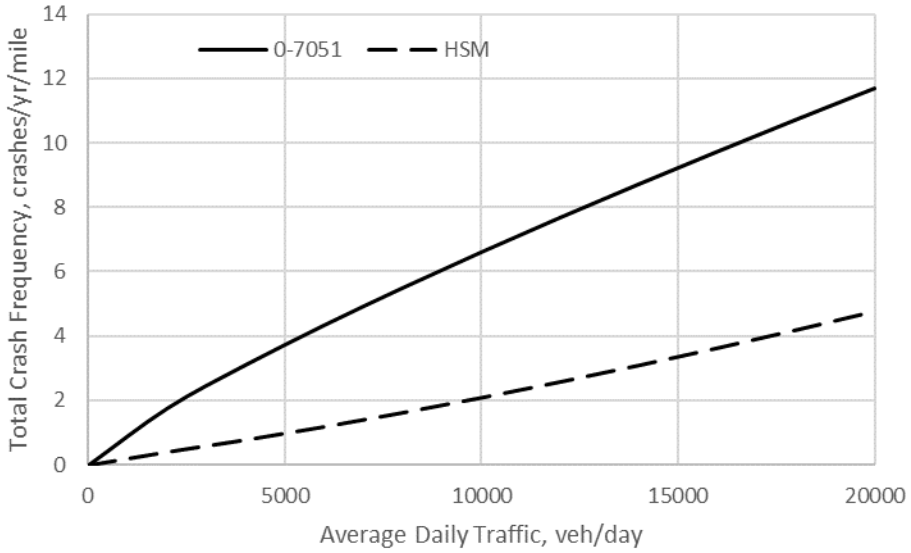


Figure 77. SPF comparison for Total Crashes on Two-lane Undivided Highways.

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 Texas WB as a means of model validation.

Truck Proportion CMF

The truck proportion CMF is described using Equation 73:

$$CMF_{tk} = e^{-0.006(tk\_perc)} \quad (73)$$

The base condition for this CMF is no trucks in the traffic mix. The truck proportion CMF is shown in Figure 78. The CMF for truck proportion for urban arterials from Texas WB is used for comparison. Both CMFs show similar trends. The CMF shows that the crashes decrease with the increase in truck proportion in traffic. Although this may seem counterintuitive, trucks usually travel on high standard roads.

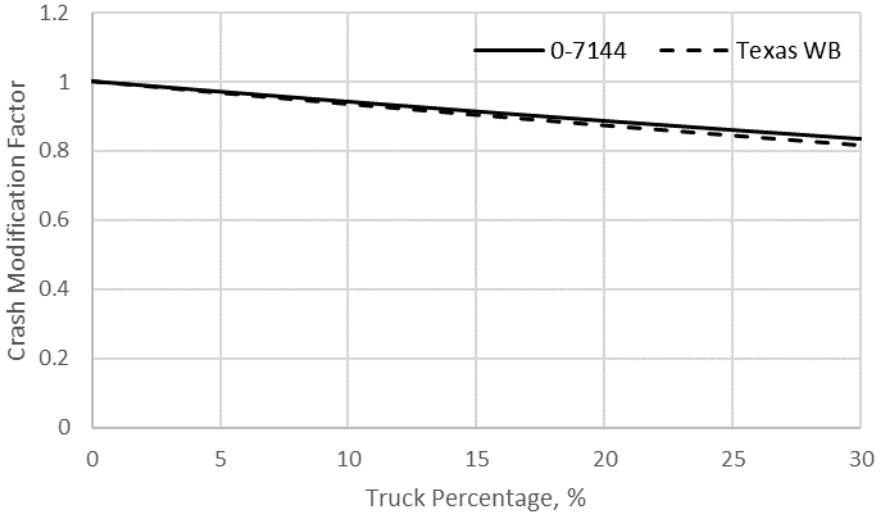


Figure 78. CMF for Truck Proportion on Two-Lane Undivided Highways.

Lane Width CMF

The lane width CMF is described using Equation 74:

$$CMF_{lw} = e^{-0.025(lw-12)} \quad (74)$$

The base condition for this CMF is a 12-ft lane width. The lane width used in this CMF is an average for all through lanes on the segment. The lane width CMF developed in this study is shown in Figure 79 using a solid trend line. The lane widths used to calibrate this CMF range from 10 to 13 ft. Also shown in Figure 79 is the CMF presented in the Texas WB. Broken lines are used to differentiate this CMF from the one proposed in this research project. The proposed CMF is shown to be less sensitive to lane width than the CMF in the Texas WB.

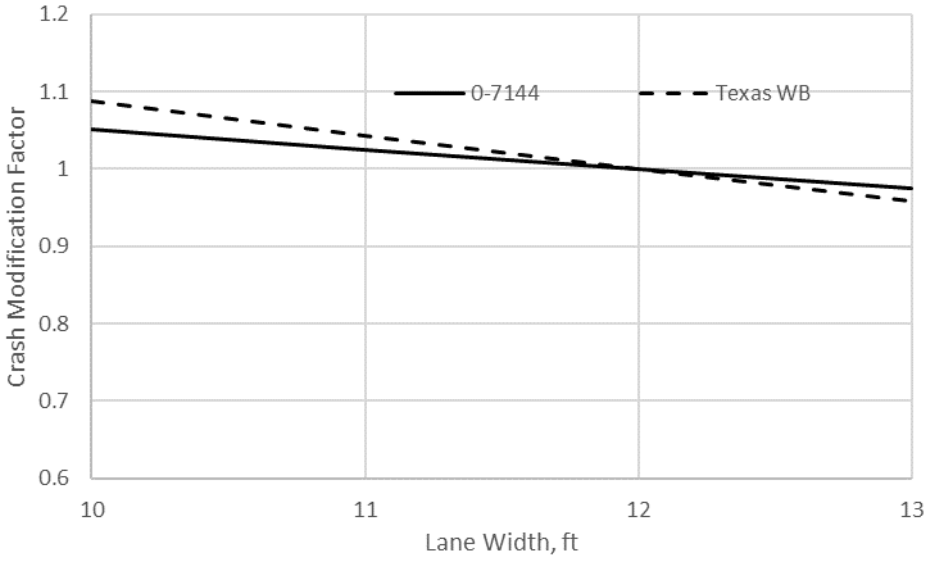


Figure 79. CMF for Lane Width on Two-Lane Undivided Highways.

Shoulder Width CMF

The shoulder width CMF is described using Equation 75:

$$CMF_{isw} = e^{-0.01\times(sw-6)} \quad (75)$$

The base condition for this CMF is a 6-ft shoulder width. The width used in this CMF is an average for shoulders in both directions. The shoulder width CMF developed in this study is shown in Figure 80 using a solid trend line. The inside shoulder widths used to calibrate this CMF range from 0 to 8 ft. Also shown in Figure 80 is the CMF presented in the Texas WB. The proposed CMF is shown to be less sensitive to shoulder width than the CMF in the Texas WB.

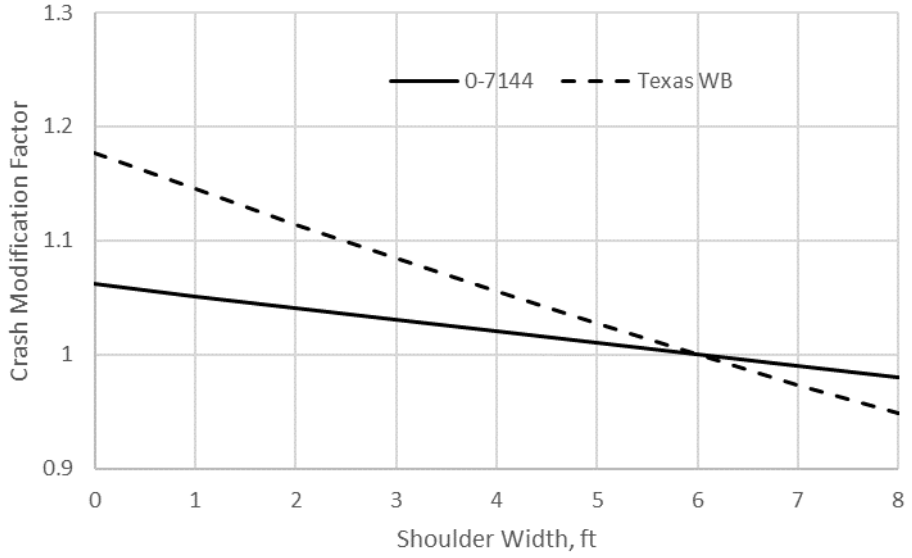


Figure 80. CMF for Shoulder Width on Two-lane Undivided Highways.

Driveway CMF

The driveway CMF is described using Equation 76.

$$CMF_{std} = e^{0.034(0.1 \times DD)} \tag{76}$$

The base condition for this CMF is no driveways. The driveway CMF is shown in Figure 81 for residential driveway density.

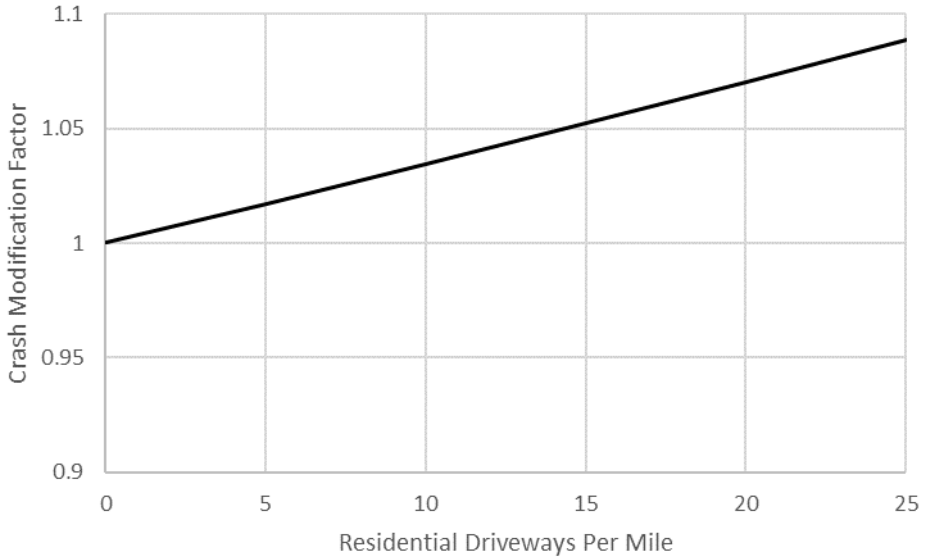


Figure 81. CMF for Driveway Density on Two-lane Divided Highways.

Precipitation CMF

The precipitation CMF is described using Equation 77.

$$CMF_{pre} = e^{0.114(0.1 \times pre)} \tag{77}$$

The base condition for this CMF is an annual average precipitation of 0 inches. Since the precipitation CMF does not exist in HSM or Texas WB, a comparison could not be made. The precipitation CMF is shown in Figure 82 using a solid trend line. The CMF shows that precipitation has a greater influence on the occurrence of crashes. For instance, every 2 inches of rainfall increases crashes by 2 percent.

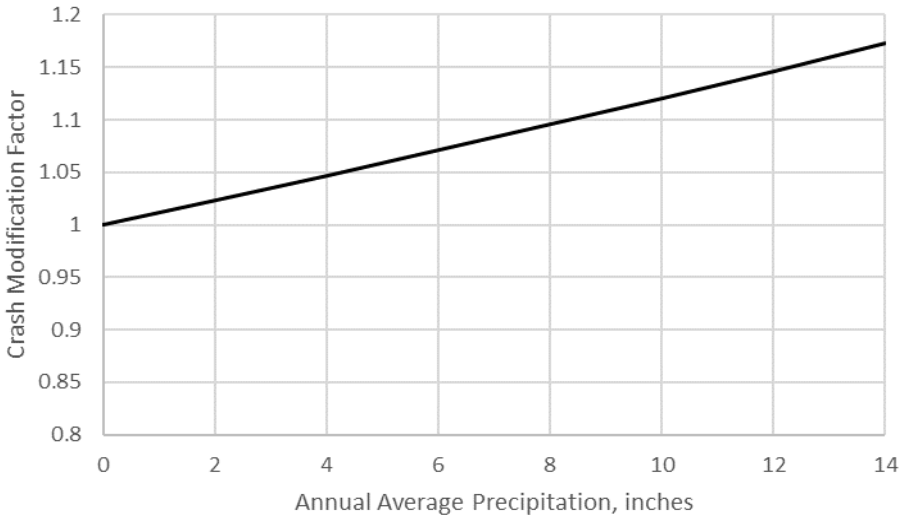


Figure 82. CMF for Precipitation on Two-lane Undivided Highways.

4.4 SENSITIVITY ANALYSIS

The relationship between crash frequency and traffic volume for all facilities, as obtained from the SPFs, is illustrated in Figure 83 for a 1-mile urban highway segment. The trend lines shown in Figure 83 indicate that the 4-lane undivided (4U) highway has the worst performance and the two-lane highway with a continuous left turn lane (2T) has the best performance, probably due to lower speeds. The 4-lane divided (4D) highway also has a much lower crash frequency compared to an undivided highway. It should be noted that the freeway SPF does not include other influential variables such as ramp presence, distance to nearest ramps, and presence of managed lanes, so the crashes on 4-lane freeways (4F) may be overestimated.

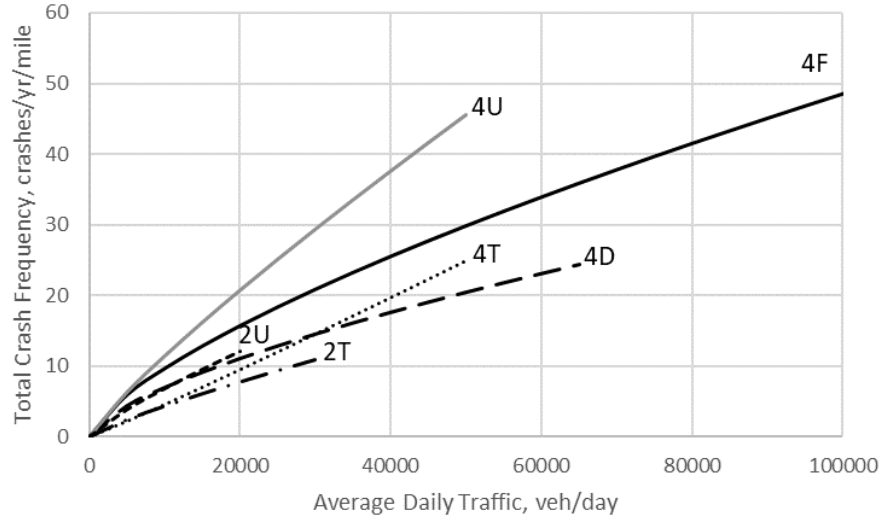


Figure 83. SPF Comparison for All Facilities.

Figure 84 shows the CMFs for the excess speed for all facilities. The trend lines shown in Figure 84 indicate that the excess speed has a more pronounced effect on 4T and the least on 4F. Researchers were not able to quantify the effect of excess speed on 2U facilities.

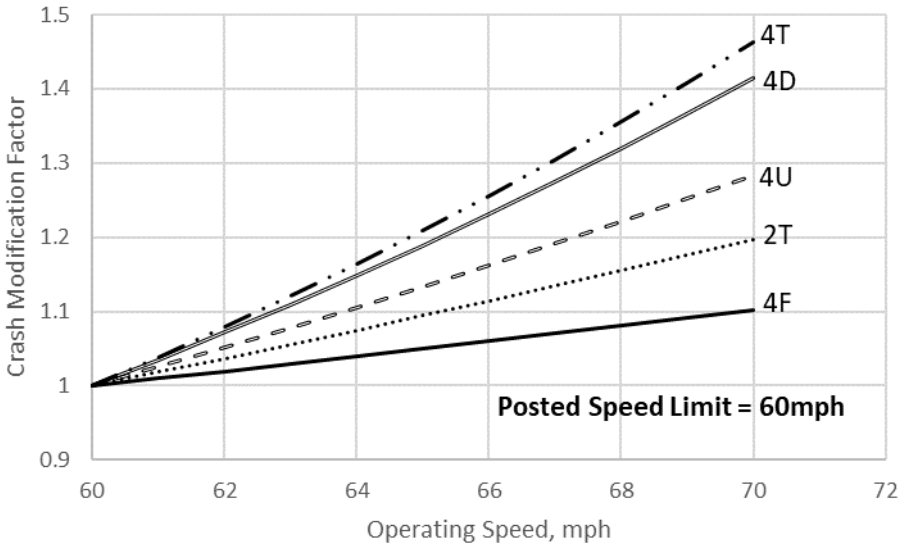


Figure 84. Excess Speed CMF Comparison Among All Facilities.

Figure 85 shows the CMFs for the precipitation for all facilities. The trend lines shown in Figure 85 indicate that precipitation has a more pronounced effect on 2U than on 4U. Researchers were not able to quantify the effect of precipitation on other facilities.

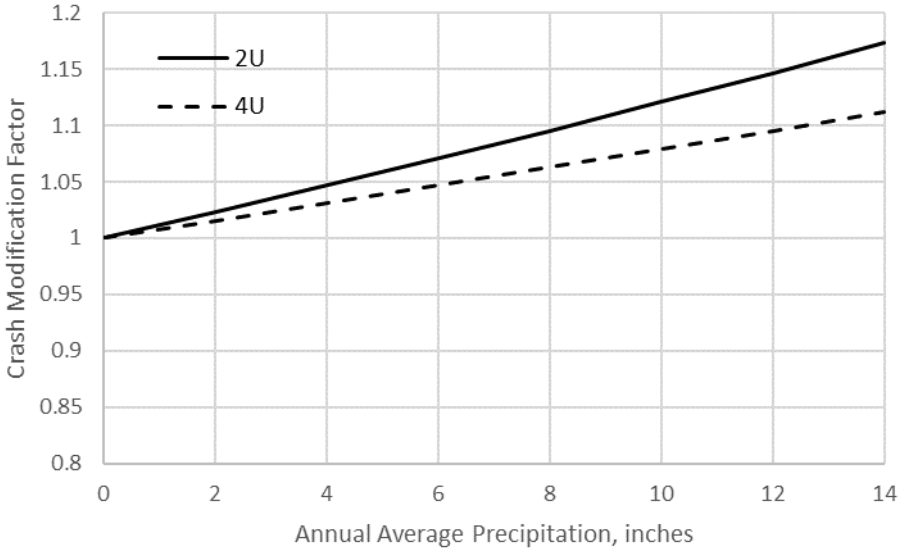


Figure 85. Precipitation CMF Comparison Among Undivided Facilities.

4.5 SHORT DURATION LEVEL DATABASES

This section presents the descriptive statistics of key variables categorized by season, day of the week, and time of day. Table 48 shows a sample of short-duration weather and speed measurements for a specific road segment, with 45 rows per segment. Table 49 to Table 53 provide detailed descriptive statistics of the key variables for short-duration analysis.

Table 48. Sample Short-Duration Weather and Speed Measurements for a Specific Road Segment.

Year	Season	DOW	TOD	Prcip_Sum	Prcip_Avg	Prcip_Max	SpdAvg	SpdStd	PSL	TTAvg
2019	All	All	All	1.84	0.00	0.01	46.60	6.39	50	49.47
2019	All	All	Daytime	0.64	0.00	0.01	48.94	5.77	50	46.88
2019	All	All	Nighttime	1.20	0.00	0.01	43.83	5.96	50	52.53
2019	All	Weekday	All	1.49	0.00	0.01	47.42	6.05	50	48.50
2019	All	Weekday	Daytime	0.53	0.00	0.01	49.79	5.47	50	46.06
2019	All	Weekday	Nighttime	0.96	0.00	0.01	44.63	5.49	50	51.37
2019	All	Weekend	All	0.35	0.00	0.01	44.54	6.73	50	51.92
2019	All	Weekend	Daytime	0.11	0.00	0.00	46.83	5.94	50	48.94
2019	All	Weekend	Nighttime	0.24	0.00	0.01	41.83	6.60	50	55.45
2019	Fall	All	All	0.70	0.00	0.01	47.99	7.02	50	48.25
2019	Fall	All	Daytime	0.25	0.00	0.01	50.66	6.04	50	45.29
2019	Fall	All	Nighttime	0.45	0.00	0.01	44.83	6.78	50	51.76
2019	Fall	Weekday	All	0.64	0.00	0.01	48.90	6.50	50	47.12
2019	Fall	Weekday	Daytime	0.23	0.00	0.01	51.56	5.61	50	44.43
2019	Fall	Weekday	Nighttime	0.41	0.00	0.01	45.77	6.08	50	50.30
2019	Fall	Weekend	All	0.06	0.00	0.00	45.71	7.71	50	51.10
2019	Fall	Weekend	Daytime	0.02	0.00	0.00	48.43	6.48	50	47.43
2019	Fall	Weekend	Nighttime	0.04	0.00	0.00	42.49	7.82	50	55.43
2019	Spring	All	All	0.50	0.00	0.00	44.74	5.15	50	51.18
2019	Spring	All	Daytime	0.20	0.00	0.00	46.76	4.58	50	48.85
2019	Spring	All	Nighttime	0.30	0.00	0.00	42.36	4.74	50	53.93
2019	Spring	Weekday	All	0.38	0.00	0.00	45.59	4.92	50	50.18
2019	Spring	Weekday	Daytime	0.16	0.00	0.00	47.67	4.35	50	47.92
2019	Spring	Weekday	Nighttime	0.22	0.00	0.00	43.13	4.39	50	52.85
2019	Spring	Weekend	All	0.12	0.00	0.00	42.60	5.09	50	53.72
2019	Spring	Weekend	Daytime	0.04	0.00	0.00	44.45	4.35	50	51.22
2019	Spring	Weekend	Nighttime	0.08	0.00	0.00	40.40	5.03	50	56.68
2019	Summer	All	All	0.52	0.00	0.01	47.53	6.71	50	48.58
2019	Summer	All	Daytime	0.15	0.00	0.00	50.09	6.08	50	45.94
2019	Summer	All	Nighttime	0.37	0.00	0.01	44.50	6.15	50	51.71
2019	Summer	Weekday	All	0.36	0.00	0.01	48.21	6.45	50	47.84
2019	Summer	Weekday	Daytime	0.10	0.00	0.00	50.88	5.67	50	45.19
2019	Summer	Weekday	Nighttime	0.26	0.00	0.01	45.06	5.87	50	50.97
2019	Summer	Weekend	All	0.16	0.00	0.01	45.88	7.04	50	50.37

2019	Summer	Weekend	Daytime	0.05	0.00	0.00	48.19	6.59	50	47.72
2019	Summer	Weekend	Nighttime	0.11	0.00	0.01	43.15	6.57	50	53.50
2019	Winter	All	All	0.12	0.00	0.00	46.14	5.99	50	49.87
2019	Winter	All	Daytime	0.04	0.00	0.00	48.26	5.39	50	47.45
2019	Winter	All	Nighttime	0.08	0.00	0.00	43.63	5.69	50	52.73
2019	Winter	Weekday	All	0.11	0.00	0.00	47.01	5.68	50	48.82
2019	Winter	Weekday	Daytime	0.04	0.00	0.00	49.07	5.30	50	46.68
2019	Winter	Weekday	Nighttime	0.08	0.00	0.00	44.57	5.14	50	51.35
2019	Winter	Weekend	All	0.01	0.00	0.00	43.87	6.18	50	52.58
2019	Winter	Weekend	Daytime	0.00	0.00	0.00	46.15	5.05	50	49.43
2019	Winter	Weekend	Nighttime	0.01	0.00	0.00	41.19	6.31	50	56.30

Table 49. Descriptive Statistics of Urban Two-lane Undivided Roadways (Short Duration).

Variables	Code	Mean	SD	Min	Max
AADT	ADT_CUR	19439	11000	2099	71576
Truck proportion	TRK_AAD	7.509	4.739	1.3	51.50
Lane width	LANE_WI	13.354	2.581	10	20.00
Inside shoulder width	S_WID_I	4.344	4.379	0	18.00
Outside shoulder width	S_WID_O	4.859	4.496	0	20.00
K-factor	K_FAC	9.198	1.384	6.4	21.50
Minor Commercial Driveway	MnrCmmr	0.027	0.170	0	2.00
Major Commercial Driveway	MjrCmmr	1.570	1.818	0	12.00
Minor Other Driveway	MnrOthr	0.006	0.077	0	1.00
Major Industrial Driveway	MjrInds	0.529	1.102	0	8.00
Minor Industrial Driveway	MnrInds	0.004	0.067	0	1.00
Minor Residential Driveway	MnrRsdn	0.028	0.198	0	2.00
Daytime 85 percentile speed (mph)	Spd85D	78.88	14.17	0	85.00
Nighttime 85 percentile speed (mph)	Spd 85N	76.96	14.61	0	85.00
Weekday 85 percentile speed (mph)	Spd 85WD	78.08	14.39	0	85.00
Weekend 85 percentile speed (mph)	Spd 85WE	79.08	14.63	0	85.00
Daytime Sum of precipitation (in)	PreD	8.76	4.40	0	1.74
Nighttime Sum of precipitation (in)	PreN	11.12	5.86	0	2.27

Table 50. Descriptive Statistics of Urban Three-lane Roadways (Short Duration).

Variables	Code	Mean	SD	Min	Max
AADT	ADT_CUR	12361	6649	1035	41353
Truck proportion	TRK_AAD	7.88	5.04	1.6	32.10
Lane width	LANE_WI	12.74	3.13	10	20.00
Inside shoulder width	S_WID_I	4.32	3.97	0	14.00
Outside shoulder width	S_WID_O	4.31	3.91	0	13.00
K-factor	K_FAC	9.63	1.99	6	21.70
Minor Commercial Driveway	MnrCmmr	0.09	0.47	0	5.00
Major Commercial Driveway	MjrCmmr	1.62	1.48	0	7.00
Minor Other Driveway	MnrOthr	0.01	0.12	0	1.00
Major Industrial Driveway	MjrInds	0.25	0.69	0	5.00
Minor Industrial Driveway	MnrInds	0.01	0.08	0	1.00
Minor Residential Driveway	MnrRsdn	0.12	0.45	0	4.00
Daytime 85 percentile speed (mph)	Spd85D	68.77	41.05	0	80.00

Nighttime 85 percentile speed (mph)	Spd 85N	67.60	42.58	0	80.00
Weekday 85 percentile speed (mph)	Spd 85WD	68.25	41.79	0	80.00
Weekend 85 percentile speed (mph)	Spd 85WE	68.69	42.49	0	80.00
Daytime Sum of precipitation (in)	PreD	8.41	5.00	0	1.34
Nighttime Sum of precipitation (in)	PreN	11.05	6.85	0	1.92

Table 51. Descriptive Statistics of Urban Four-Lane Undivided Roadways (Short Duration).

Variables	Code	Mean	SD	Min	Max
AADT	ADT_CUR	13357	8762	2	71576
Truck proportion	TRK_AAD	6.31	5.85	0	81.40
Lane width	LANE_WI	12.05	2.60	4	20.00
Inside shoulder width	S_WID_I	2.22	3.89	0	30.00
Outside shoulder width	S_WID_O	2.38	4.08	0	26.00
K-factor	K_FAC	9.81	1.94	4.3	50.00
Minor Commercial Driveway	MnrCmmr	0.75	2.47	0	47.00
Major Commercial Driveway	MjrCmmr	3.17	4.95	0	77.00
Minor Other Driveway	MnrOthr	0.04	0.35	0	12.00
Major Industrial Driveway	MjrInds	0.53	1.90	0	48.00
Minor Industrial Driveway	MnrInds	0.11	0.94	0	29.00
Minor Residential Driveway	MnrRsdn	0.12	0.57	0	17.00
Daytime 85 percentile speed (mph)	Spd85D	78.47	32.02	0	85.00
Nighttime 85 percentile speed (mph)	Spd 85N	85.47	33.12	0	95.00
Weekday 85 percentile speed (mph)	Spd 85WD	78.00	32.53	0	85.00
Weekend 85 percentile speed (mph)	Spd 85WE	78.39	33.15	0	85.00
Daytime Sum of precipitation (in)	PreD	8.76	4.39	0	1.82
Nighttime Sum of precipitation (in)	PreN	11.05	5.81	0	2.37

Table 52. Descriptive Statistic of Urban Four-Lane Divided Roadways (Short Duration).

Variables	Code	Mean	SD	Min	Max
AADT	ADT_CUR	20158	12155	250	140415
Truck proportion	TRK_AAD	9.48	8.01	0.1	81.40
Lane width	LANE_WI	12.41	1.56	6	20.00
Inside shoulder width	S_WID_I	6.07	5.43	0	30.00
Outside shoulder width	S_WID_O	12.30	7.64	0	44.00
K-factor	K_FAC	9.52	1.43	6	21.20
Minor Commercial Driveway	MnrCmmr	0.08	0.45	0	8.00
Major Commercial Driveway	MjrCmmr	1.73	2.88	0	31.00
Minor Other Driveway	MnrOthr	0.04	0.27	0	6.00
Major Industrial Driveway	MjrInds	1.12	1.89	0	23.00
Minor Industrial Driveway	MnrInds	0.02	0.22	0	5.00
Minor Residential Driveway	MnrRsdn	0.03	0.22	0	5.00
Daytime 85 percentile speed (mph)	Spd85D	78.90	48.29	0	90.00
Nighttime 85 percentile speed (mph)	Spd 85N	76.49	48.94	0	85.00
Weekday 85 percentile speed (mph)	Spd 85WD	77.89	48.66	0	85.00
Weekend 85 percentile speed (mph)	Spd 85WE	78.54	49.32	0	85.00
Daytime Sum of precipitation (in)	PreD	8.74	4.23	0	1.70
Nighttime Sum of precipitation (in)	PreN	11.05	5.65	0	2.21

Table 53. Descriptive Statistic of Urban Five-lane Roadways (Short Duration).

Variables	Code	Mean	SD	Min	Max
AADT	ADT_CUR	19439	11000	2099	71576
Truck proportion	TRK_AAD	7.509	4.739	1.3	51.50
Lane width	LANE_WI	13.354	2.581	10	31.00
Inside shoulder width	S_WID_I	4.344	4.379	0	18.00
Outside shoulder width	S_WID_O	4.859	4.496	0	20.00
K-factor	K_FAC	9.198	1.384	6.4	21.50
Minor Commercial Driveway	MnrCmmr	0.027	0.170	0	2.00
Major Commercial Driveway	MjrCmmr	1.570	1.818	0	12.00
Minor Other Driveway	MnrOthr	0.006	0.077	0	1.00
Major Industrial Driveway	MjrInds	0.529	1.102	0	8.00
Minor Industrial Driveway	MnrInds	0.004	0.067	0	1.00
Minor Residential Driveway	MnrRsdn	0.028	0.198	0	2.00
Daytime 85 percentile speed (mph)	Spd85D	78.47	32.02	0	85.00
Nighttime 85 percentile speed (mph)	Spd 85N	85.47	33.12	0	90.00
Weekday 85 percentile speed (mph)	Spd 85WD	78.00	32.53	0	85.00
Weekend 85 percentile speed (mph)	Spd 85WE	78.39	33.15	0	85.00
Daytime Sum of precipitation (in)	PreD	8.76	4.39	0	1.82
Nighttime Sum of precipitation (in)	PreN	11.05	5.81	0	2.37

4.6 SHORT DURATION SAFETY PERFORMANCE FUNCTIONS

The negative binomial Lindley model is preferred over the traditional negative binomial model for datasets associated with preponderant zeros. To begin, we document the formulation of the NB. The NB model can either be formulated as a sequence of Bernoulli trials or a mixture of the Poisson and gamma distributions. The Poisson-gamma mixture is normally used for analyzing crash data. The NB Generalized Linear Model (GLM) is formulated as follows:

$$P(Y = y; \mu, \phi) = NB(y;\ \mu, \phi)\ = \frac{\Gamma(\phi + y)}{\Gamma(\phi)\Gamma(y+1)}\left(\frac{\phi}{\mu+\phi}\right)^{\phi}\left(\frac{\mu}{\mu+\phi}\right)^{y};\ \phi > 0, \mu > 0 \quad (78)$$

where, μ = mean response of the observation and ϕ = inverse of the dispersion parameter. The μ is assumed to have a loglinear relationship with the covariates as shown in Equation (79):

$$\ln(\mu) = \beta_0 + \sum_{j=1}^{m} \beta_j X \quad (79)$$

where X= explanatory variables considered for the study, β_0 is the intercept and β_j is the regression coefficient for the j-th covariate, and m = total number of covariates in the model.

For NBL models, the Lindley distributions provide extra flexibility to address the excess zeros problems. The NBL is the mixture of the NB and Lindley distributions. The NB-L model can be formulated as follows:

$$P(Y = y;\ \mu, \phi, \theta) = \int NB(y; \phi, \varepsilon\mu)\,\text{Lindley}(\varepsilon\,; \theta)\, d\varepsilon \quad (80)$$

where ε is the frailty term and θ is the Lindley parameter. Equation (80) however does not have a closed form. Therefore, it is often rewritten as a multi-level hierarchical structure, as shown in Equation (81) to Equation (84):

$$y_i|\varepsilon_i\mu_i,\phi \sim NB(\varepsilon_i\mu_i,\phi) \quad (81)$$

$$\varepsilon_i|z_I,\theta \sim \text{gamma}(1+z_I,\theta) \quad (82)$$

$$z_i|\theta \sim \text{Bernoulli}\left(\frac{1}{1+\theta}\right) \quad (83)$$

$$\ln(\mu_i|\beta_0,\beta_1\cdots\beta_m) = \beta_0 + \sum_{j=1}^{m}\beta_j X_{ij} \quad (84)$$

As short duration weather data is not available for 2018, the Project Team developed yearly models for the following urban roadways for 2022, 2021, 2020, and 2019 separately.

4.6.1 Short Duration SPFs (2022 Data)

2U Roadways

Considering crashes occurring on urban 2-way undivided roads, the crash variables were separated into two parts. The variables that were fixed over the months, like shoulder width and surface width and variables that were observed to change over months, like standard deviation of speed, standard deviation of travel time, monthly average daily traffic (MADT) and sum of monthly precipitation. As shown in Table 54, the variable for shoulder width had a negative relationship with crash count, indicating that an increase in shoulder width is likely to result in a reduction in the number of crashes. On the other hand, an increase in road surface width (i.e., cross-sectional width of the road without shoulders) is likely to increase the number of crashes. For variables that changed with months, the standard deviation of speed and travel time had a negative relationship with crash frequency, meaning an increase in the standard deviation of speed and travel time is likely to reduce the number of crashes. The MADT and the sum of precipitation have a positive association with the number of crashes, indicating that an increase in those variables increases the number of crashes. Finally, the variable for minor commercial driveways had a positive association with crash counts. The results indicate the number of crashes is likely to increase in minor commercial areas.

Table 54. Short duration SPFs (2022 Data) for 2U Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.5%	97.5%
Intercept	-9.925	0.145	-10.200	-9.641
Shoulder width	-0.013	0.001	-0.016	-0.010
Surface width	0.009	0.001	0.007	0.012
Speed standard deviation	-0.001	0.009	-0.018	0.015
Travel time standard deviation	-0.004	0.002	-0.009	0.000
Log (MADT)	1.001	0.016	0.971	1.033
Sum precipitation	0.077	0.017	0.044	0.107
MnrCmmr (1 if MnrCmmr is 4, 0 otherwise)	0.206	0.126	-0.021	0.504
Dispersion parameter	6.047	0.909	4.614	8.111
Model Performance				
DIC	71,020			

WAIC	73,030			

4U Roadways

For short duration SPFs for four-lane undivided (4U) roadways, as presented in Table 55, the analysis indicates that shoulder width has a negligible effect on crash counts, evidenced by a mean close to zero and a credible interval that crosses zero. This contrasts with the findings for 2U roadways, where shoulder width showed a more significant impact. The average speed of vehicles has a negative relationship with crash counts, as indicated by the negative mean value. This suggests that higher average speeds are associated with a reduction in the number of crashes. In contrast, the standard deviation of speed shows a positive relationship with crash counts. A higher variability in speed among vehicles is associated with an increase in crashes. The standard deviation of travel time has a negative relationship with crash frequency. This means that greater variability in travel time is linked to a decrease in the number of crashes, possibly due to drivers adjusting their behavior in response to varying conditions. (MADT) shows a positive association with crash counts. This indicates that as traffic volume increases, so does the likelihood of crashes. Similarly, the sum of precipitation is positively associated with crash counts, suggesting that more precipitation leads to a higher number of crashes, likely due to reduced visibility and more slippery road conditions.

Table 55. Short duration SPFs (2022 Data) for 4U Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-6.124	0.138	-6.388	-5.871
Shoulder width	-2.49E-04	9.52E-04	-0.002	0.002
Average speed	-0.051	8.53E-04	-0.052	-0.049
Speed standard deviation	0.102	0.008	0.087	0.118
Travel time standard deviation	-0.007	0.002	-0.01	-0.004
Log (MADT)	0.817	0.014	0.790	0.845
Sum precipitation	0.071	0.012	0.047	0.092
Dispersion Parameter	93.980	22.200	58.680	143.900
Model Performance				
DIC	71,210			
WAIC	92,710			

4D Roadways

The results of short-duration SPFs for four-lane divided (4D) roadways, as shown in Table 56, highlight key factors related to crash risk. Wider road surfaces and higher average speeds are linked to fewer crashes. In contrast, greater speed variability increases crash risk, emphasizing the importance of maintaining consistent speeds. Average travel time does not significantly affect crash counts, as its impact is minimal. However, MADT is associated with an increase in crashes, pointing to the influence of traffic density on safety. Additionally, increased

precipitation correlates with higher crash counts, likely due to reduced visibility and slippery conditions in wet weather.

Table 56. Short duration SPFs (2022 Data) for 4D Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-6.300	0.222	-6.732	-5.859
Surface width	-0.004	0.001	-0.006	-0.002
Average speed	-0.038	0.001	-0.039	-0.036
Speed standard deviation	0.091	0.006	0.078	0.102
Average travel time	0.001	0.001	-0.001	0.002
Log (MADT)	0.790	0.021	0.748	0.832
Sum Precipitation	0.062	0.020	0.023	0.099
Dispersion parameter	22.590	7.907	12.270	43.490
Model Performance				
DIC	37280			
WAIC	47260			

3T Roadways

As shown in Table 57, various factors affect crash counts on three-lane roadways with a center left turn lane. The results indicate that wider road surfaces are linked to more crashes, while higher average speeds tend to reduce crash counts, possibly due to better road design or management. However, increased variability in vehicle speeds leads to more crashes, highlighting the importance of maintaining consistent speeds for safety. Higher traffic volumes are also associated with an increase in crashes, emphasizing the role of traffic density in crash risk. In contrast, the width of the shoulder does not significantly impact crash counts. In summary, these findings suggest that maintaining consistent speeds and managing traffic volumes are essential for enhancing safety on 3T roadways.

Table 57. Short duration SPFs (2022 Data) for 3T Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-14.040	1.643	-17.300	-10.930
Surface width	0.031	0.007	0.016	0.045
Shoulder width	0.008	0.012	-0.016	0.032
Average speed	-0.022	0.009	-0.039	-0.006
Speed standard deviation	0.222	0.059	0.110	0.340
Log (MADT)	1.293	0.161	0.992	1.623
Dispersion Parameter	12.040	8.268	3.051	34.820
Model Performance				
DIC	1,245			
WAIC	1,270			

5T Roadways

The factors influencing crash counts on five-lane roadways with a center turn lane indicate that higher average speeds are associated with fewer crashes, suggesting that effective road design or traffic management may contribute to safer conditions at these speeds. However, increased variability in vehicle speeds is linked to more crashes, emphasizing the need for consistent driving speeds to enhance safety. Lane width does not significantly impact crash counts, as indicated by the near-zero mean and credible interval crossing zero. In contrast, MADT are associated with a greater number of crashes, underscoring the impact of traffic density on crash risk. These findings are presented in Table 58.

Table 58. Short duration SPFs (2022 Data) for 5T Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-6.545	0.527	-7.531	-5.536
Lane width	-0.004	0.010	-0.024	0.017
Average speed	-0.052	0.002	-0.057	-0.047
Speed standard deviation	0.090	0.021	0.051	0.132
Log (MADT)	0.851	0.050	0.757	0.953
Dispersion Parameter	37.790	14.310	17.340	71.500
Model Performance				
DIC	10,900			
WAIC	11,060			

4.6.2 Short Duration SPFs (2021 Data)

2U Roadways

The factors affecting crash counts on two-lane undivided roadways show that increasing shoulder width is associated with fewer crashes, suggesting that wider shoulders can enhance safety. Variability in vehicle speed (speed standard deviation) also has a negative relationship with crash counts, indicating that more consistent driving speeds can reduce crash risk. Conversely, higher ADT and increased precipitation are linked to more crashes, emphasizing the influence of traffic volume and weather conditions on safety. Lane width has a slight positive relationship with crash counts, though its impact is uncertain as the credible interval includes zero. The presence of minor commercial driveways is positively associated with crash counts, suggesting these areas may have a higher risk of crashes, though this effect is not particularly strong. These findings are outlined in Table 59.

Table 59. Short duration SPFs (2021 Data) for 2U Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-9.178	0.151	-9.480	-8.894

Shoulder width	-0.014	0.001	-0.017	-0.011
Average speed	0.001	0.001	-0.001	0.003
Speed standard deviation	-0.040	0.008	-0.056	-0.027
Lane width	0.007	0.005	-0.002	0.017
Average travel time	-0.002	0.000	-0.003	-0.001
MnrCmmr (1 if MnrCmmr is 4, 0 otherwise)	0.209	0.138	-0.068	0.469
Log (MADT)	0.975	0.015	0.945	1.003
Sum precipitation	0.045	0.014	0.020	0.074
Dispersion Parameter	4.835	0.639	3.813	6.313
Model Performance				
DIC	69,250			
WAIC	71,240			

4U Roadways

The factors influencing crash counts on four-lane undivided roadways indicate that both surface width and shoulder width have a slight negative relationship with crash counts, suggesting that wider surfaces and shoulders may help reduce crashes, though these effects are minimal and less certain. Higher average speeds are associated with fewer crashes, possibly due to better road conditions or traffic management, while variability in vehicle speed (speed standard deviation) is linked to more crashes, highlighting the importance of consistent driving speeds. Average travel time is negatively related to crash counts, implying that longer travel times might be associated with fewer crashes, potentially due to more careful driving or less congestion. Higher traffic volumes, indicated by the log of MADT, correlate with more crashes, underscoring the role of traffic density in crash risk. Increased precipitation is also associated with a higher number of crashes, pointing to the influence of weather conditions on safety. Interestingly, the presence of minor industrial areas has a slight negative association with crash counts, suggesting that these areas may experience fewer crashes, although this effect is not particularly strong. These findings are shown in Table 60.

Table 60. Short duration SPFs (2021 Data) for 4U Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-5.662	0.142	-5.959	-5.387
Surface width	-0.001	0.001	-0.003	0.000
Shoulder width	-0.002	0.001	-0.004	0.000
Average speed	-0.046	0.001	-0.048	-0.044
Speed standard deviation	0.062	0.007	0.049	0.075
Average travel time	-0.002	0.000	-0.002	-0.001
Log (MADT)	0.788	0.015	0.760	0.818
Sum precipitation	0.028	0.010	0.008	0.047
MnrInds (1 if MnrInds is 1, 0 otherwise)	-0.146	0.073	-0.294	-0.005
Dispersion Parameter	74.910	19.080	44.250	119.200
Model Performance				

DIC	91290			
WAIC	92990			

4D Roadways

The factors affecting crash counts on four-lane divided roadways show that both surface width and shoulder width have a slight negative relationship with crash counts, suggesting that wider road surfaces and shoulders can help reduce crashes, although these effects are relatively small. Higher average speeds are associated with fewer crashes, which may indicate that roads designed for higher speeds have better safety features or conditions. However, greater variability in speed (speed standard deviation) is linked to more crashes, emphasizing that inconsistent driving speeds increase crash risk. Average travel time shows a slight positive relationship with crash counts, indicating that longer travel times might be associated with a higher likelihood of crashes, potentially due to driver fatigue or other factors. Increased MADT is positively associated with crash counts, demonstrating that higher traffic volumes lead to more crashes. The sum of precipitation has a small positive relationship with crash counts, suggesting that weather conditions might slightly influence crash risk. The presence of minor commercial areas is also positively associated with crash counts, implying that these areas could experience more crashes, possibly due to increased traffic and turning movements. These findings are outlined in Table 61.

Table 61. Short duration SPFs (2021 Data) for 4D Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-5.796	0.217	-6.230	-5.347
Surface width	-0.003	0.001	-0.005	-0.001
Shoulder width	-0.002	0.001	-0.005	0.000
Average speed	-0.036	0.001	-0.038	-0.034
Speed standard deviation	0.076	0.007	0.063	0.089
Average travel time	0.001	0.001	0.000	0.003
Log (MADT)	0.732	0.021	0.690	0.774
Sum precipitation	0.021	0.016	-0.007	0.052
MnrCmmr (1 if MnrCmmr is 1, 0 otherwise)	0.170	0.059	0.053	0.282
Dispersion Parameter	6.243	0.927	4.761	8.396
Model Performance				
DIC	46,740			
WAIC	48,050			

3T Roadways

For three-lane roadways with a center turn lane, increasing surface and shoulder width is associated with more crashes, suggesting that wider roads may lead to riskier driving behavior. Higher average speeds are linked to fewer crashes, indicating these roads may be designed for safer high-speed travel. However, greater variability in speed increases crash risk, highlighting the importance of consistent driving speeds. Greater variability in travel time is slightly associated with fewer crashes, possibly encouraging more cautious driving. Higher traffic

volumes are linked to an increase in crashes, emphasizing the role of traffic density in crash risk. These findings are presented in Table 62.

Table 62. Short duration SPFs (2021 Data) for 3T Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-9.523	1.515	-12.48	-6.51
Surface width	0.02718	0.007503	0.012	0.04181
Shoulder width	0.03241	0.01304	0.007037	0.05775
Average speed	-0.04703	0.007619	-0.06185	-0.03174
Speed standard deviation	0.1928	0.06302	0.06364	0.3119
Travel time standard deviation	-0.02939	0.01865	-0.06618	0.007677
Log (MADT)	0.9531	0.1586	0.6477	1.264
Dispersion Parameter	6.994	6.061	1.543	22.88
Model Performance				
DIC	1,236			
WAIC	1,267			

5T Roadways

The factors influencing crash counts on three-lane roadways with a center turn lane suggest that lane width has a slight negative relationship with crash counts, indicating that wider lanes might help reduce crash risk. However, the effect is not strong, and the credible interval crosses zero, showing some uncertainty. Higher average speeds are linked to fewer crashes, suggesting that these roads may be better designed or offer conditions that support safer travel at higher speeds. The travel time standard deviation shows a slight positive relationship with crash counts, but this effect is weak, and the credible interval includes zero, indicating that variability in travel time has a limited impact on crash risk. Higher traffic volumes are associated with more crashes, highlighting the impact of traffic density on crash occurrence. These findings are presented in Table 63.

Table 63. Short duration SPFs (2021 Data) for 5T Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-6.338	0.512	-7.375	-5.363
Lane width	-0.016	0.010	-0.036	0.004
Average speed	-0.045	0.003	-0.050	-0.040
Travel time standard deviation	0.007	0.004	-0.002	0.015
Log (MADT)	0.853	0.048	0.762	0.948
Dispersion Parameter	33.920	13.400	15.120	66.950
Model Performance				
DIC	10600			
WAIC	10810			

4.6.3 Short Duration SPFs (2020 Data)

2U Roadways

The factors affecting crash counts on two-lane undivided roadways suggest that wider shoulders and more consistent driving speeds help reduce crashes, while wider lanes might increase crash risk. Average speed and travel time have minimal effects on crash counts, with slight indications that longer travel times might improve safety. Higher traffic volumes are linked to more crashes, highlighting the role of traffic density. Precipitation does not have a significant impact on crash counts. These findings are outlined in Table 64.

Table 64. Short duration SPFs (2020 Data) for 2U Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-9.066	0.169	-9.385	-8.732
Shoulder width	-0.013	0.002	-0.016	-0.010
Average speed	-0.002	0.001	-0.004	0.000
Speed standard deviation	-0.019	0.008	-0.035	-0.004
Lane width	0.028	0.005	0.016	0.038
Average travel time	-0.002	0.000	-0.003	-0.001
Log (MADT)	0.930	0.017	0.896	0.962
Sum precipitation	0.008	0.016	-0.024	0.042
Dispersion Parameter	4.494	0.637	3.473	6.018
Model Performance				
DIC	61100			
WAIC	62870			

4U Roadways

The factors influencing crash counts on four-lane undivided roadways suggest that increasing surface and shoulder widths slightly reduces crashes. Higher average speeds are linked to fewer crashes, while greater speed variability increases crash risk. More variability in travel time may be associated with safer driving, possibly due to cautious behavior in changing conditions. Higher traffic volumes lead to more crashes, highlighting the impact of traffic density. Precipitation has a slightly negative effect on crash counts but is not a major factor. The presence of minor industrial areas has a weak negative association with crash risk, but this impact is not significant. These findings are presented in Table 65.

Table 65. Short duration SPFs (2020 Data) for 4U Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-5.440	0.149	-5.728	-5.144
Surface width	-0.001	0.001	-0.002	0.000
Shoulder width	-0.002	0.001	-0.004	0.000
Average speed	-0.050	0.001	-0.051	-0.048

Speed standard deviation	0.110	0.006	0.098	0.123
Travel time standard deviation	-0.011	0.002	-0.014	-0.008
Log (MADT)	0.743	0.016	0.712	0.773
Sum precipitation	-0.018	0.013	-0.043	0.005
MnrInds(1 if MnrInds is 1, 0 otherwise)	-0.088	0.078	-0.243	0.075
Dispersion Parameter	71.300	18.510	42.750	112.400
Model Performance				
DIC	81700			
WAIC	83390			

4D Roadways

The factors affecting crash counts on four-lane divided roadways suggest that wider road and shoulder widths may slightly reduce crashes, although the effects are minimal. Higher average speeds are linked to fewer crashes, likely due to better road design or management, while increased speed variability raises crash risk, highlighting the need for consistent speeds. Variability in travel time is slightly associated with more crashes, possibly due to unpredictable traffic conditions. Higher traffic volumes also lead to more crashes, emphasizing the role of traffic density in crash risk. Minor commercial areas are associated with higher crash counts, likely due to increased traffic and turning movements. These findings are outlined in Table 66.

Table 66. Short duration SPFs (2020 Data) for 4D Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-6.836	0.244	-7.331	-6.350
Surface width	-0.002	0.001	-0.005	0.000
Shoulder width	-0.001	0.001	-0.004	0.001
Average speed	-0.033	0.001	-0.035	-0.031
Speed standard deviation	0.096	0.009	0.078	0.114
Travel time standard deviation	0.005	0.003	0.000	0.011
Log (MADT)	0.803	0.024	0.755	0.853
MnrCmmr (1 if MnrCmmr is 1, 0 otherwise)	0.205	0.059	0.091	0.322
Dispersion Parameter	4.608	0.602	3.617	5.974
Model Performance				
DIC	42,090			
WAIC	43,310			

3T Roadways

Wider road surfaces on three-lane roadways with a center turn lane are associated with more crashes, while higher average speeds tend to reduce crash counts. Variability in driving speeds increases crash risk, highlighting the importance of maintaining consistent speeds. Variations in travel time may slightly reduce crashes, but this effect is weak. Higher traffic volumes lead to more crashes, underlining the need to manage traffic levels on these roads. These findings are indicated in Table 67.

Table 67. Short duration SPFs (2020 Data) for 3T Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-7.106	2.164	-11.490	-2.936
Surface width	0.023	0.009	0.005	0.040
Shoulder width	0.008	0.014	-0.018	0.035
Average speed	-0.030	0.009	-0.046	-0.011
Speed standard deviation	0.140	0.079	-0.012	0.299
Travel time standard deviation	-0.033	0.024	-0.084	0.014
Log(MADT)	0.667	0.227	0.226	1.126
Dispersion Parameter	7.850	6.856	1.491	26.720
Model Performance				
DIC	950.5			
WAIC	972.6			

5T Roadways

For five-lane roadways with a center turn lane, higher average speeds are linked to fewer crashes, suggesting these roads can safely handle higher speeds. Variability in travel time increases crash risk, indicating that inconsistent travel times may lead to more crashes. Lane width has little impact on crash counts, while higher traffic volumes are associated with more crashes, highlighting the need to manage traffic density. These findings are shown in Table 68.

Table 68. Short duration SPFs (2020 Data) for 5T Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-7.454	0.577	-8.532	-6.239
Lane width	-0.012	0.010	-0.032	0.008
Average speed	-0.044	0.003	-0.050	-0.039
Travel time standard deviation	0.024	0.005	0.015	0.034
Log (MADT)	0.938	0.055	0.831	1.043
Dispersion Parameter	27.580	11.680	11.580	57.720
Model Performance				
DIC	9,613			
WAIC	9,828			

4.6.4 Short Duration SPFs (2019 Data)

2U Roadways

For two-lane undivided roadways, increasing shoulder width is linked to fewer crashes, suggesting that wider shoulders improve safety. Higher average speeds have little impact on crash counts. Variability in speed slightly increases crash risk, while greater variability in travel time is associated with fewer crashes, possibly due to safer driving behaviors. Wider road surfaces are linked to more crashes, indicating potential for riskier driving. Higher traffic

volumes also lead to more crashes, emphasizing the impact of traffic density on safety. Increased precipitation is associated with more crashes, highlighting the role of weather conditions in crash risk. These findings are indicated in Table 69.

Table 69. Short duration SPFs (2019 Data) for 2U Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-9.642	0.149	-9.932	-9.354
Shoulder width	-0.009	0.002	-0.013	-0.006
Average speed	-0.001	0.001	-0.003	0.001
Speed standard deviation	0.012	0.007	-0.001	0.025
Surface width	0.007	0.001	0.004	0.010
Travel time standard deviation	-0.007	0.002	-0.011	-0.003
Log (MADT)	0.978	0.016	0.946	1.009
Sum precipitation	0.036	0.015	0.004	0.067
Dispersion Parameter	9.902	2.786	6.265	16.980
Model Performance				
DIC	63210			
WAIC	65010			

4U Roadways

Increasing shoulder width on four-lane undivided roadways is linked to fewer crashes, suggesting that wider shoulders enhance safety, while surface width has no significant impact. Higher average speeds are associated with fewer crashes, indicating these roads may handle higher speeds safely. However, variability in speed leads to more crashes, highlighting the risk of inconsistent driving speeds. Greater variability in travel time appears to reduce crashes, possibly by encouraging safer driving behaviors. Higher traffic volumes are linked to more crashes, emphasizing the role of traffic density in crash risk. The impact of precipitation on crash counts is minimal. These findings are highlighted in Table 70.

Table 70. Short duration SPFs (2019 Data) for 4U Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-6.922	0.153	-7.211	-6.630
Surface width	0.000	0.001	-0.001	0.001
Shoulder width	-0.004	0.001	-0.006	-0.001
Average speed	-0.051	0.001	-0.052	-0.049
Speed standard deviation	0.099	0.007	0.086	0.112
Travel time standard deviation	-0.006	0.001	-0.009	-0.003
Log (MADT)	0.896	0.016	0.866	0.925
Sum precipitation	0.001	0.011	-0.022	0.023
Dispersion Parameter	62.840	17.470	36.110	103.400
Model Performance				
DIC	89,360			
WAIC	91,110			

4D Roadways

Increasing both surface width and shoulder width on four-lane divided roadways is linked to a reduction in crashes, suggesting that wider roads and shoulders enhance safety. Higher average speeds are associated with fewer crashes, indicating these roads can safely accommodate higher speeds. However, increased speed variability is linked to more crashes, highlighting the need for consistent driving speeds. The effect of the average travel time on crash counts is minimal, with a slight tendency to reduce crashes. Higher traffic volumes lead to more crashes, underscoring the impact of traffic density on safety. The presence of minor commercial areas is associated with more crashes, likely due to increased traffic and turning movements. These findings are outlined in Table 71.

Table 71. Short duration SPFs (2019 Data) for 4D Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-6.441	0.250	-6.927	-5.972
Surface width	-0.002	0.001	-0.005	0.000
Shoulder width	-0.004	0.001	-0.006	-0.002
Average speed	-0.035	0.001	-0.037	-0.033
Speed standard deviation	0.103	0.007	0.088	0.116
Average travel time	-0.001	0.001	-0.002	0.000
Log (MADT)	0.783	0.025	0.736	0.833
MnrCmmr (1 if MnrCmmr is 1, 0 otherwise)	0.138	0.059	0.022	0.253
Dispersion Parameter	4.018	0.446	3.268	4.988
Model Performance				
DIC	43900			
WAIC	45180			

3T Roadways

Higher average speeds on three-lane roadways with a center turn lane are linked to fewer crashes, suggesting these roads may allow safer driving at higher speeds. In contrast, speed variability increases crash risk, indicating that inconsistent speeds are a concern. Greater variability in travel time is associated with fewer crashes, possibly encouraging more cautious driving. The effects of surface width and shoulder width on crash counts are positive but not strong, with credible intervals including zero, indicating uncertainty in their impact. Higher traffic volumes lead to more crashes, emphasizing the need to manage traffic density on these roads. These findings are shown in Table 72.

Table 72. Short duration SPFs (2019 Data) for 3T Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-11.000	1.922	-14.730	-7.347
Surface width	0.011	0.009	-0.006	0.029
Shoulder width	0.015	0.013	-0.011	0.041
Average speed	-0.021	0.006	-0.034	-0.009

Speed standard deviation	0.258	0.064	0.134	0.378
Travel time standard deviation	-0.039	0.018	-0.074	-0.004
Log (MADT)	1.061	0.190	0.688	1.437
Dispersion Parameter	2.842	2.558	0.875	9.657
Model Performance				
DIC	1,185			
WAIC	1,218			

5T Roadways

Higher traffic volumes, measured by the log of MADT, are associated with an increase in crashes, emphasizing the importance of managing traffic density to enhance safety on these roads. Higher average speeds on five-lane roadways with a center turn lane are associated with fewer crashes, suggesting these roads may support safer high-speed driving. Increased variability in travel time leads to more crashes, highlighting the risk posed by inconsistent travel times. Lane width has minimal impact on crash counts, indicating it does not significantly affect safety. Higher traffic volumes are linked to more crashes, underscoring the need to manage traffic density to improve safety on these roads. These findings are identified in Table 73.

Table 73. Short duration SPFs (2019 Data) for 5T Roadways.

Variables	Mean	Standard deviation	Credible interval	
			2.50%	97.50%
Intercept	-8.979	0.523	-10.030	-7.966
Lane width	-0.009	0.009	-0.027	0.010
Average speed	-0.042	0.002	-0.046	-0.037
Travel time standard deviation	0.022	0.003	0.015	0.028
Log (MADT)	1.069	0.048	0.976	1.161
Dispersion Parameter	32.400	12.810	14.760	61.790
Model Performance				
DIC	11,170			
WAIC	11,400			

4.7 SUMMARY

This chapter provides an in-depth analysis of SPFs for various urban roadways in Texas, exploring factors influencing crash frequencies across different roadway configurations, such as two-lane undivided, four-lane divided, and multi-lane facilities with center turn lanes. The analysis includes annual-level data for SPFs and examines the influence of roadway characteristics, traffic volumes, and environmental conditions on crash occurrences. The study identifies critical variables, such as lane width, shoulder width, average speed, speed variability, and precipitation, that significantly affect crash frequencies. Additionally, the chapter discusses the calibration of CMFs for different road types and presents comparisons with existing models to validate findings. In addition, short duration yearly models were developed for five urban roadway facilities. The annual and short-duration SPFs are used for developing the decision support tool.

CHAPTER 5
DECISION SUPPORT TOOL

5.1 INTRODUCTION

This chapter presents a brief overview of the developed decision support tool (P2). The Project Team developed a geographic information system (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

The Project Team used the open-source software platform Shiny to develop the decision support tool. Two versions of the tool were developed to accommodate different user needs: one with a data upload option and another without it. The version with the data upload option is accessible via the following link: https://aitlab.shinyapps.io/0_7144_V03/. The version without the data upload option can be accessed here: https://aitlab.shinyapps.io/7144Tool/. This chapter can be considered as a software manual that guides the use of the interactive tool developed for this project.

5.2.1 Interface

Figure 86 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.

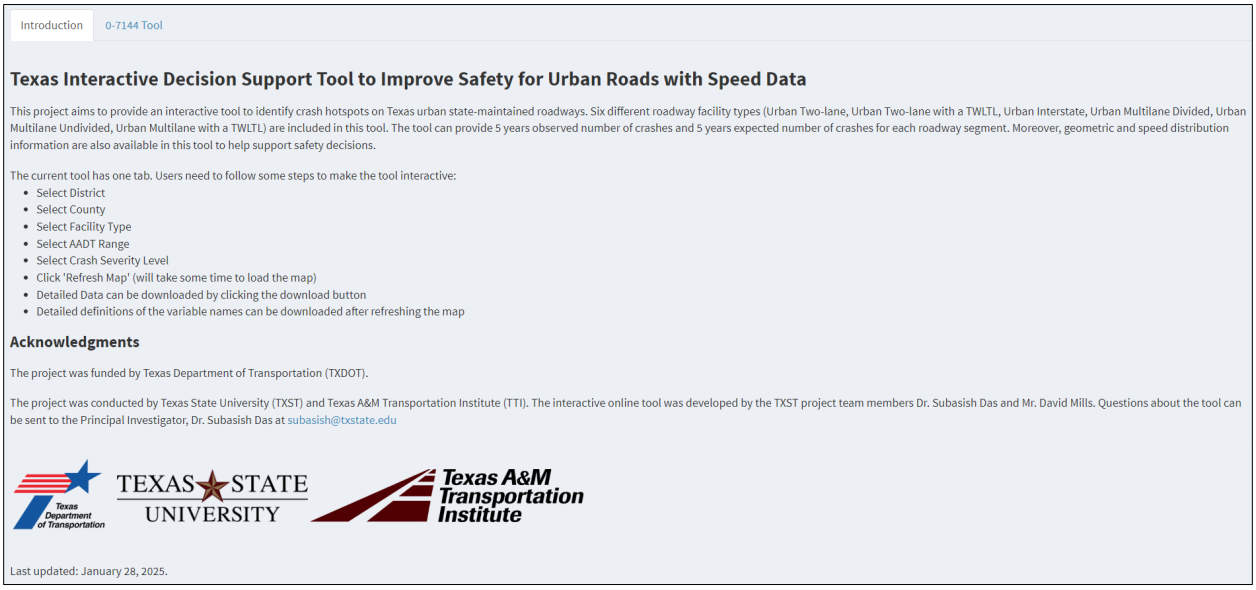


Figure 86. Interface of the 0-7144 Decision Support Tool.

The web interface has two tabs:

- Introduction (the interface shown in Figure 86)
- 0-7144 Tool (users can go to this page by clicking on the tab)

5.2.2 Decision Support Tool

The team developed two versions of the decision support tool: one without a data upload option and another with a data upload option. The version without data upload enables the visualization of urban roads using existing data and predictions. The version with data upload allows users to update the existing dataset and generate estimates based on the newly uploaded information.

Without Data Upload Option

Figure 87 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 selecting the filters and clicking 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.
- *Note:* Two notes providing instruction on the data dictionary and interactive table.
- *Zoom in/out in 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 (will show up after selecting the filters and map refreshing).

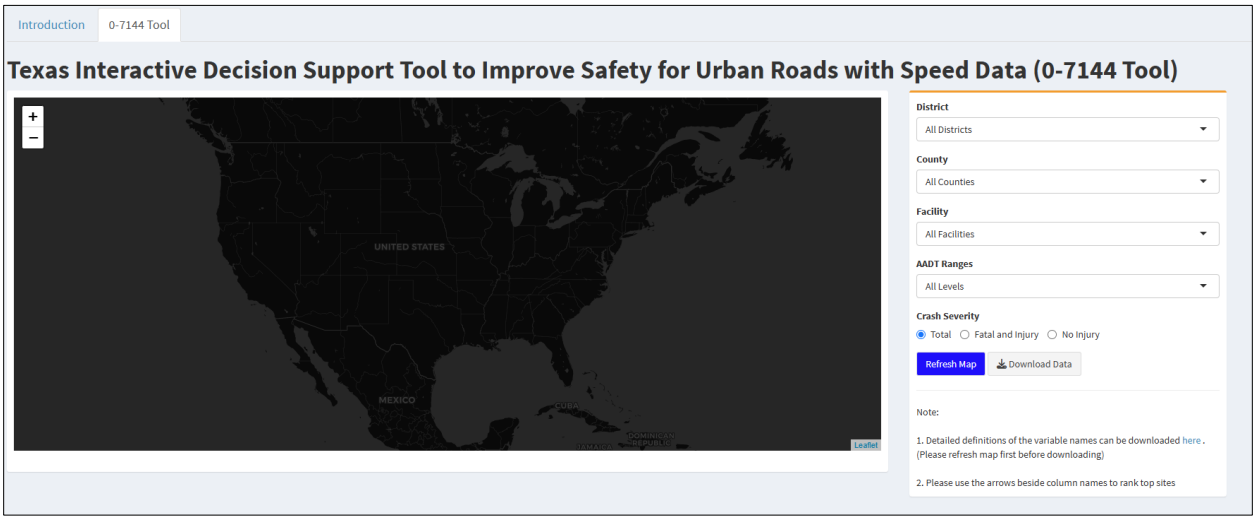


Figure 87. Decision Support Tool without Data Upload Option.

With Data Upload Option

Compared to the data upload version, the version (see Figure 88) with uploaded data includes the following features in the top panel after selecting 'District':

- *Upload Your CSV:* data uploading from a local device

Figure 88. Decision Support Tool with Data Upload Option.

5.2.3 Map Generation Steps

Detailed explanation of the tool

Six different roadway facility types are included in this tool: Urban Two Lane, Urban Two Lane with a TWLTL, Urban Interstate Highway, Urban Multilane Divided, Urban Multilane Undivided, and Urban Multilane with a TWLTL. The tool provides the 5-year observed number of crashes and the 5-year expected number of crashes for each roadway segment. The tool is based on 5 years of crash data (2018-2022) from CRIS (Texas Department of Transportation, 2025b). Other sources of data include roadway inventory data from the RHiNO database, weather data from the NOAA, and operating speed data from the NPMRDS/INRIX XD (Texas A&M Transportation Institute, 2025; Texas Department of Transportation, 2025a).

The results can be filtered by the following in the 0-7144 Tool tab:

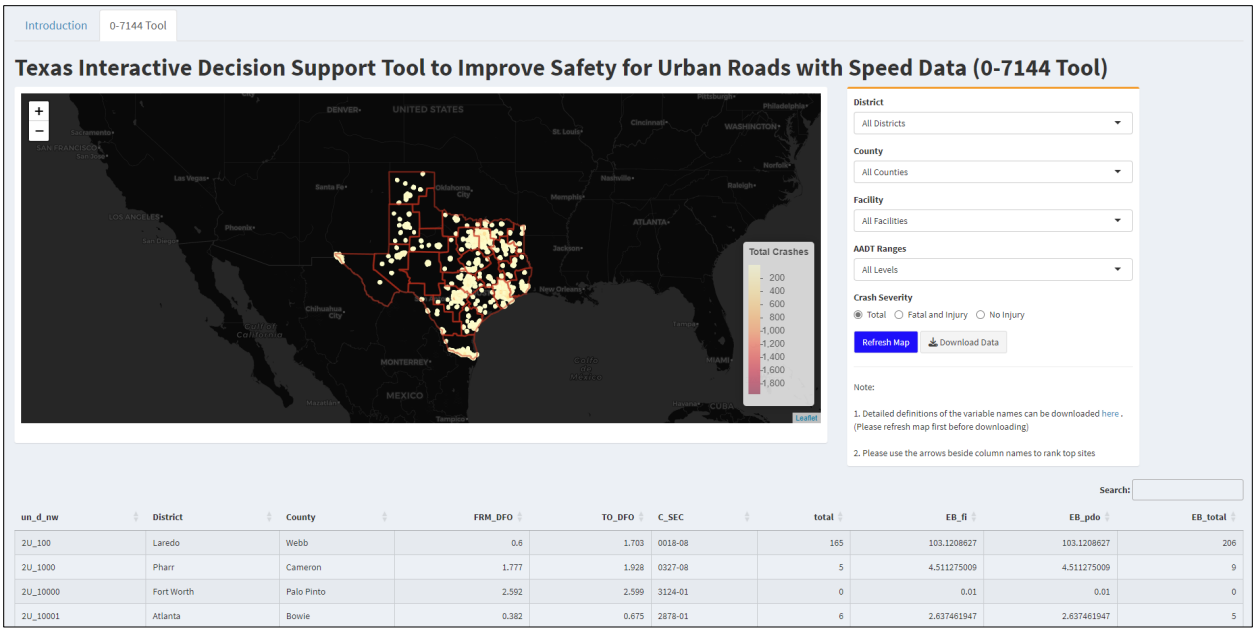
- District (TxDOT districts)
- County (Counties in Texas)
- Facility Type (All facilities; Urban Two Lane, Urban Two Lane with a TWLTL, Urban Interstate Highway, Urban Multilane Divided, Urban Multilane Undivided, and Urban Multilane with a TWLTL)

- AADT Ranges (Less than 2000 vehicles per day or vpd, 2,001-10,000 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 is important to note that it may take some time to load the map]

The results can be filtered by the following in the 0-7144 Tool tab:
- Detailed data can be downloaded by clicking the ‘Download Data’ button (grey box below blue box).
- Data dictionary (see Appendix B: Data Dictionary) can be downloaded after refreshing the map (see Note 1 in the 0-7144 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 for the results.

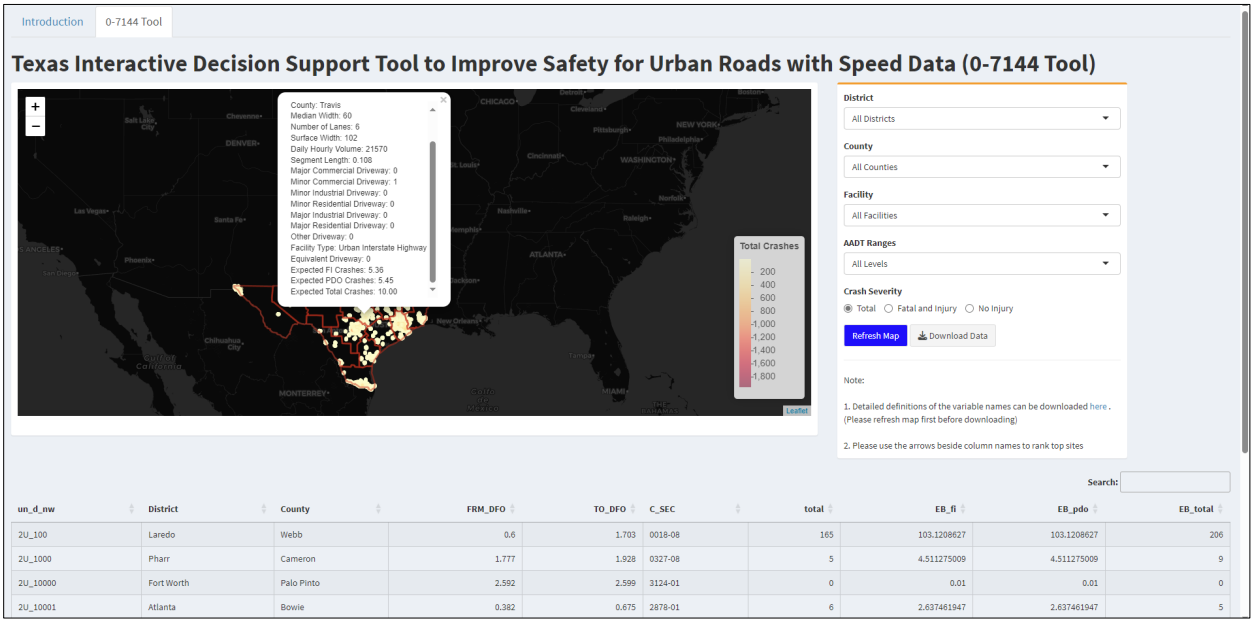
Figure 89 shows the image of the interface after selecting ‘all’ from the four drop-down panels. The map shows the entire state-maintained urban roadway networks. The color of the segments is based on the number of estimated/expected crashes on the individual segment. Below the interactive map, an interactive table produces the result of the final selection. The column names and associated descriptions can be downloaded by clicking the ‘Data Dictionary’ button ‘here’ 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 below.



un_d_nw	District	County	FRM_DFO	TO_DFO	C_SEC	total	EB_fi	EB_pdo	EB_total
2U_100	Laredo	Webb	0.6	1.703	0018-08	165	103.1208627	103.1208627	206
2U_1000	Pharr	Cameron	1.777	1.928	0327-08	5	4.511275009	4.511275009	9
2U_10000	Fort Worth	Palo Pinto	2.592	2.599	3124-01	0	0.01	0.01	0
2U_10001	Atlanta	Bowie	0.382	0.675	2878-01	6	2.637461947	2.637461947	5

Figure 89. Interface of the Tool after Selecting 'all' from the Four Drop-down Panels.

After the generation of the map, the user can hover over a segment to see the segment-specific information (see Figure 90).



un_d_nw	District	County	FRM_DFO	TO_DFO	C_SEC	total	EB_fi	EB_pdo	EB_total
2U_100	Laredo	Webb	0.6	1.703	0018-08	165	103.1208627	103.1208627	206
2U_1000	Pharr	Cameron	1.777	1.928	0327-08	5	4.511275009	4.511275009	9
2U_10000	Fort Worth	Palo Pinto	2.592	2.599	3124-01	0	0.01	0.01	0
2U_10001	Atlanta	Bowie	0.382	0.675	2878-01	6	2.637461947	2.637461947	5

Figure 90. Hovering Option Details.

Example #1

A Safety Engineer from the Houston District wants to explore the tool to understand the safety condition of the median-volume urban two-lane roadways.

The safety engineer needs to select the following options:

- *District:* Houston
- *County:* All Counties
- *Facility:* Urban Two-lane
- *AADT Ranges:* 2,001 to 10,000

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 91 displays the map generated 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 the darker red indicates a higher number of expected crashes.

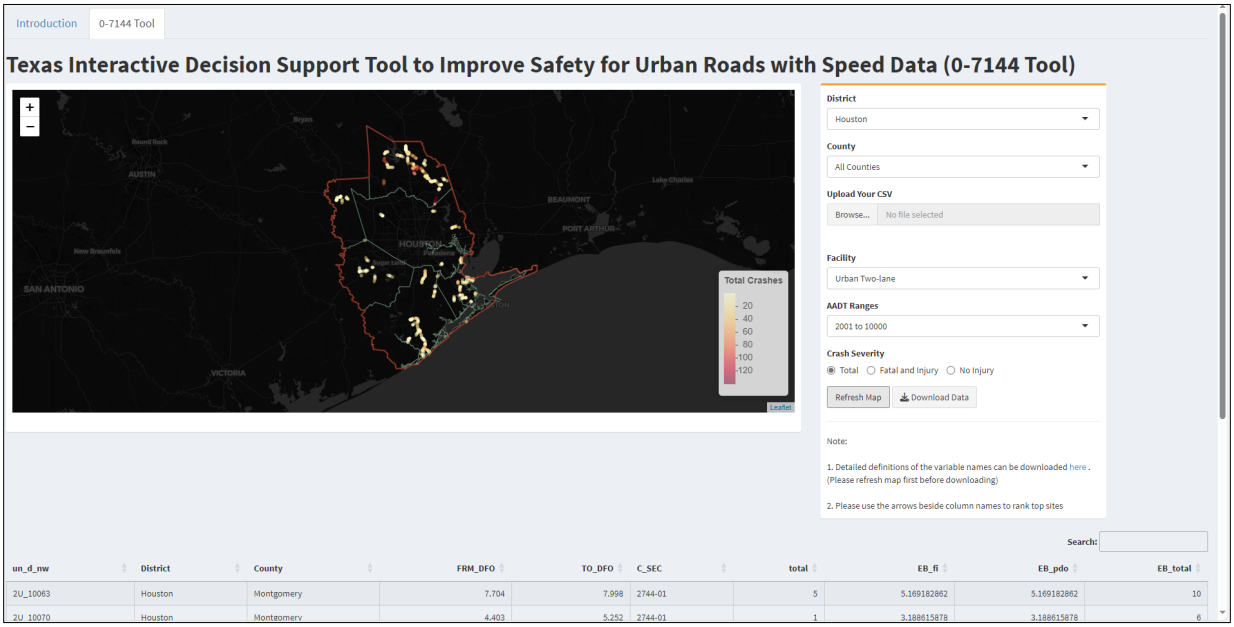


Figure 91. District Specific Map Details.

The user also have options to explore a specific county. For example, if the user selects Harris County, the map will display only urban two-lane median volume roadways in Harris County (see Figure 92). The user can also download the data after finalizing the selection by clicking the 'Download Data' button. To get more details on the segment, the user 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 it has a hovering option to get more details on a particular segment.

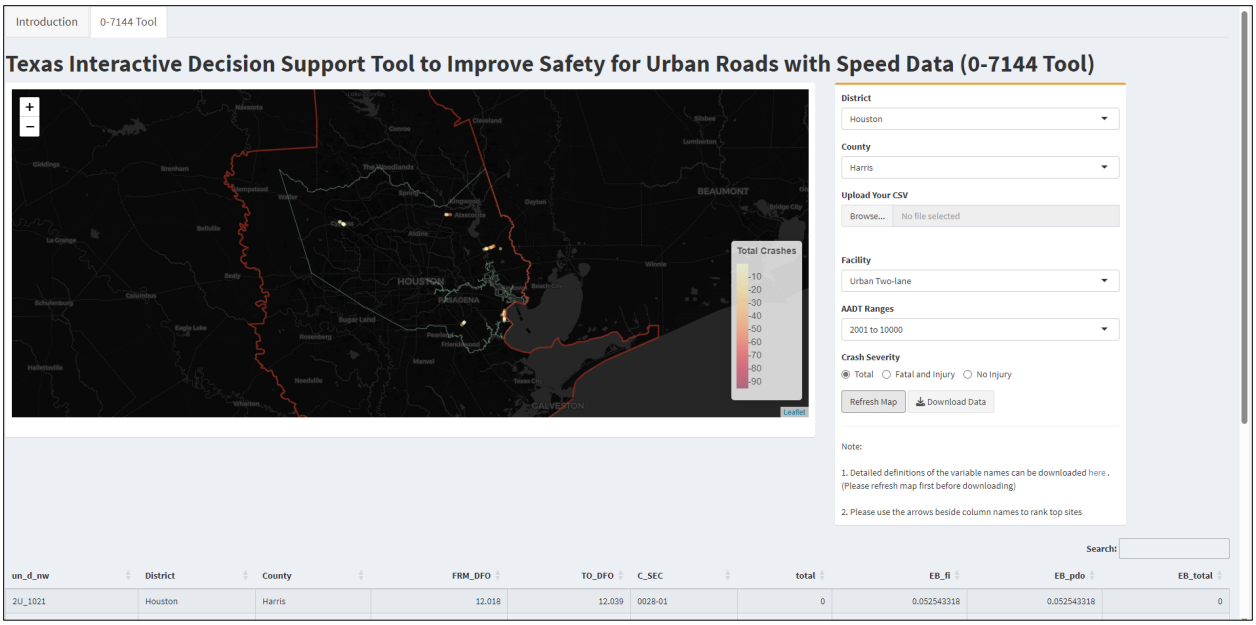


Figure 92. County Specific Map Details.

Crash Estimation Update with New Data: The tool provides users with the option to upload data. For example, if the user wants to update historical data and predict crashes in a particular county in the Houston District, the data can be uploaded by "Upload Your CSV" (see Figure 93). To begin, users must select a district and county from the dropdown menus provided. This initial step is crucial as it narrows the dataset to specific geographic segments, ensuring the results are relevant to the selected area. The dropdown menus allow for easy navigation through the various districts and counties available in the dataset.

When a user needs to update the crash data for road segments in a county, the application offers the option to upload custom CSV files. These files may contain specific roadway or crash data, enabling users to enhance the analysis with modified information. However, to ensure successful integration, the uploaded data must include specific columns, such as County, ADT, Control Section, FromDFO, ToDFO, MedType (Median Type), NumLane (Number of Lanes), and SpdAve (Speed Average). A built-in validation mechanism automatically checks the presence of these columns and notifies users if any required fields are missing. This ensures that only properly formatted data is processed. Once the custom data is uploaded, the application proceeds to verify its consistency with the district and county selected earlier. This validation step includes checking whether the number of rows in the uploaded dataset corresponds to the selected geographic parameters. If any discrepancies are detected, the system promptly alerts users, ensuring that the analysis remains accurate and aligned with the selected district and county.

With the data validated, the application begins processing the crash data. During this phase, it calculates the expected number of crashes, including Fatal/Injury crashes (FICrE), Property Damage Only crashes (PDOCrE), and Total crashes (TotalCrE). While these calculations are being performed, a spinning 'busy' icon appears on the screen, signaling to users that the

processing is underway. This visual feedback helps manage user expectations during potentially time-consuming operations. Upon completing the calculations, the application generates a detailed hotspot map. This map highlights areas within the selected district and county that are prone to crashes, based on the uploaded data and calculated metrics. Users can refine the visualization by applying filters for factors such as Facility Type, AADT Range, and Crash Severity. This flexibility allows for a more targeted analysis, enabling users to focus on specific safety concerns or roadway conditions. After the analysis is done, the application provides options to download the results. The processed data can be saved as a CSV file, along with a detailed map of crash hotspots. Additionally, a downloadable spreadsheet containing definitions for all variable names is available, ensuring users have a comprehensive understanding of the data and its context.

The user also have options to upload the data they have. For example, if the user wants to update historical data and predict crashes in Harris County, the data can be uploaded by "Upload Your CSV" (see Figure 93).

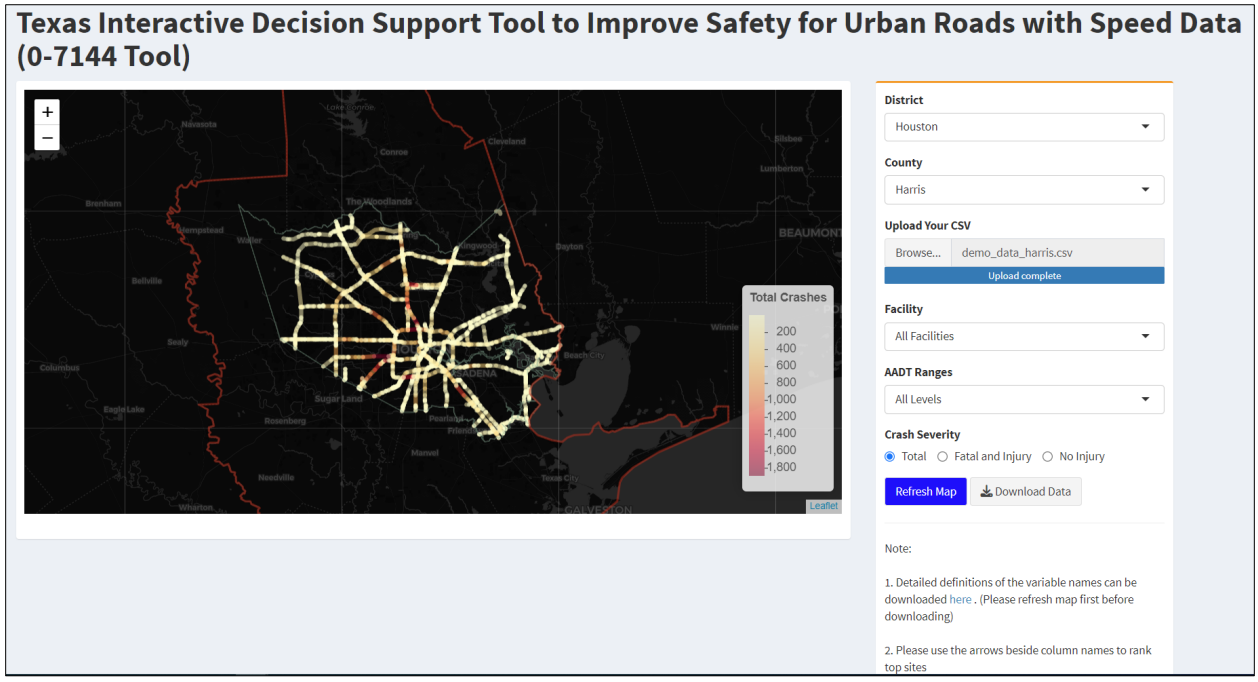


Figure 93. County Specific Map with Uploaded Data.

Figure 94 illustrates the attributes of a specific segment in the default database before and after new data is uploaded. Upon uploading new data, the number of crashes is updated based on historical records.

a) Before New Data is Loaded

b) After New Data is Loaded

Figure 94. Comparison of Segment Attributes Before and After Data Uploaded.

Example #2

A Safety Engineer from the Wichita Falls District wants to explore the tool to understand the safety condition of urban interstate highways.

The safety engineer needs to select the following options:

- *District:* Wichita Falls
- *County:* All Counties
- *Facility:* Urban Interstate Highways
- *AADT Ranges:* All levels

The user can select the 'Crash Severity' option as needed. For example, if 'Fatal and injury' is selected from 'Crash Severity,' a map will be generated after clicking 'Refresh Map.' Figure 95 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.

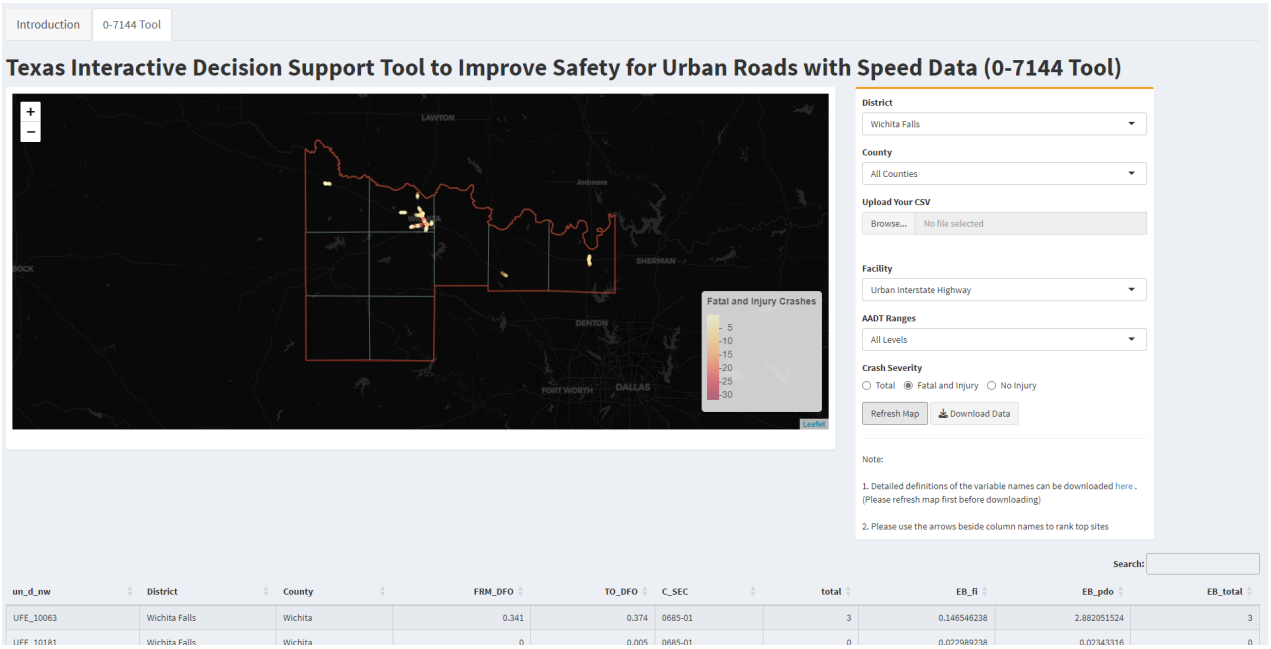


Figure 95. Screenshot Showing Wichita Falls District.

Example #3

A Safety Engineer from the Austin District wants to explore the tool to understand the safety condition of high-volume urban roadways.

The safety engineer needs to select the following options:

- *District:* Austin
- *County:* All Counties
- *Facility:* All Facilities
- *AADT Ranges:* Greater than 10,000

The user can select the ‘Crash Severity’ option as needed. For example, if ‘Fatal and injury’ is selected from ‘Crash Severity,’ a map will be generated after clicking ‘Refresh Map.’ Figure 96 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.

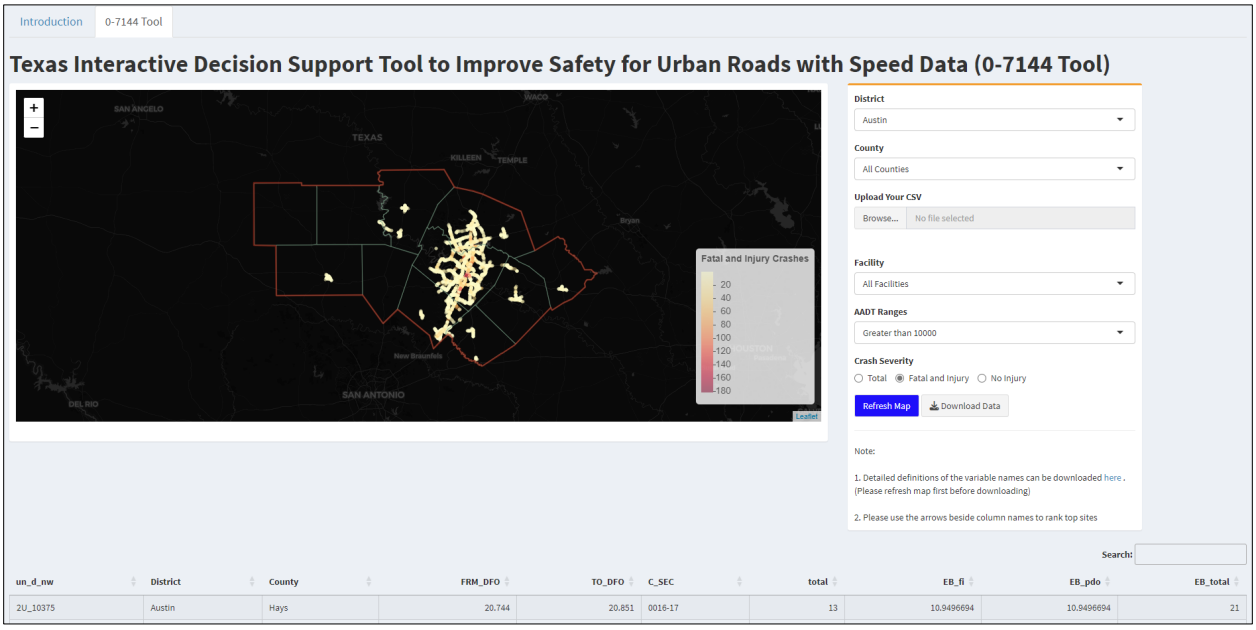


Figure 96. Screenshot Showing Austin District.

5.3 SUMMARY

This chapter presents the development of the Decision Support Tool for assessing safety risks on urban roadways. This tool was developed to assist transportation professionals visualize and analyze crash risks across different roadway segments. The tool integrates data from various sources, including crash data, speed data, and weather conditions, and uses these inputs to calculate and display SPFs for each roadway segment. It includes an interactive, map-based interface built using GIS technology, which allows users to view risk assessments for urban roadways, prioritize high-risk areas, and make data-driven decisions about where to implement safety improvements. The decision support tool is dynamic, with the capability to update risk scores as new data is integrated into the system. The platform includes three key components: a cloud-based data warehouse, a computational platform, and a web server. This chapter describes how the tool was developed, its functionality, and its potential impact on urban roadway safety planning.

CHAPTER 6
CONCLUSIONS AND RECOMMENDATIONS

6.1 INTRODUCTION

Texas Strategic Highway Safety Plan (SHSP) has identified speeding related crashes as one of the seven research emphasis areas for 2018-2022. The conventional crash risk analysis method typically omits real-time speed, real-time volume, and weather data. This can significantly limit their predictive performances. To address this gap, in this study, the Project Team utilized data from several sources: (1) NPMRDS and INRIX XD (real-time speed data); (2) TMAS (traffic volume data), CRIS (Crash data), RHiNO (roadway inventory data), and NOAA (precipitation data). For annual level data analysis, the Project Team developed SPFs for five distinct urban facility types: freeways, multi-lane divided highways, multi-lane undivided highways, multi-lane undivided highways with continuous left-turn lanes (CLTL), and two-lane highways. 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, the Project Team applied the NB-Lindley model to develop a monthly level model by year as well. In the final step, the Project Team developed an interactive decision support tool using the open-source software platform Shiny.

6.2 RESEARCH PRODUCTS

Annual SPFs for Urban Facilities: The annual-level SPFs were generated for five distinct urban facility types: freeways, multi-lane divided highways, multi-lane undivided highways, multi-lane undivided highways with CLTL, and two-lane highways. These models incorporated factors such as lane and shoulder width, average speed, speed variability, and precipitation to analyze their influence on crash frequency.

Short-duration SPFs for Urban Facilities: In contrast, short-duration SPFs were developed for the same five facility types using year-specific data from 2019 to 2022, allowing for an assessment of temporal variability and seasonal effects.

Decision Support Tool: The decision support tool developed in TxDOT Project 0-7144 represents a significant advancement in crash risk analysis by integrating traditionally overlooked but critical variables such as operating speed, traffic volume, and weather data. Developed using the open-source Shiny platform, the GIS-based tool estimates and visualizes expected crash frequencies across various urban roadway types in Texas. Users can interactively explore roadway safety conditions using filters like district, county, facility type, AADT range, and crash severity. One version of the tool operates on built-in datasets and offers visualization, filtering, and downloading capabilities. The inclusion of multiple national datasets (NPMRDS/INRIX XD) for speed data, TMAS for traffic volume, and NOAA for real-time weather) enhances the analytical power of the tool, addressing the complex interdependencies

among roadway geometry, speed, and crash frequency, which are often ignored in conventional models and even in HSM's SPFs.

What sets this tool apart is its ability to consider user-upload data for customized crash estimation, making it a highly adaptable and dynamic platform for transportation agencies. Users can upload CSV files containing local or updated traffic and roadway characteristics to produce refined crash predictions specific to their area of interest. The tool validates the structure of the uploaded data, aligns it with the selected geographic scope, and updates crash predictions in real time, distinguishing between fatal/injury crashes, property damage only crashes, and total crashes. The platform also provides downloadable outputs and documentation to support decision-making and facilitate transparency in safety planning efforts.

6.3 FINDINGS AND CONCLUSIONS

The Project Team developed SPFs using the crash data with an annual aggregation interval. The SPFs are developed for FI and PDO crashes, and both types together. The findings from the annual level model are as follows:

- The freeway SPFs showed that higher AADT and excess speed significantly increase crash frequency, while wider lanes, shoulders, and medians reduce it. Truck-heavy routes showed lower crash rates, likely due to higher roadway standards. The developed CMFs were more sensitive than those in national models, offering improved calibration for Texas-specific conditions.
- Key findings for multi-lane divided highways include increased AADT, a higher number of lanes, excess speed, and driveway density significantly increase crash risk, while wider shoulders, medians, and lane widths generally reduce it. CMFs for factors such as truck proportion, shoulder widths, and excess speed were shown to be more sensitive and context-specific than those in the HSM and Texas WB, indicating the models' improved calibration for Texas conditions.
- The SPFs for multi-lane undivided highways focused on four-lane segments and incorporated CMFs for factors like shoulder width, speed, truck proportion, and precipitation. AADT, excess speed, driveway density, and precipitation were significant predictors of increased crash frequency, while wider shoulders helped reduce crashes. Compared to HSM and Texas WB models, the proposed CMFs offer enhanced sensitivity to local conditions, particularly for precipitation and excess speed, which showed a 30–40% crash increase per 10 mph over the limit.
- The SPFs for multi-lane undivided highways with continuous left turn lanes showed that AADT, excess speed, and driveway density significantly increase crash risk, while wider shoulders and lanes help reduce it. A 10 mph increase over the speed limit raises crash risk by 45%, highlighting the sensitivity of these segments to speeding. Compared to Texas WB and HSM models, the proposed CMFs offer updated, locally calibrated insights that better reflect the safety impacts of geometric and operational features on these roadway types.

- The SPFs for two-lane undivided highways with continuous left turn lanes showed that AADT and driveway density significantly increased crash risk, while wider shoulders reduced it. Excess speed was not statistically significant for FI crashes and was excluded from the PDO model. Compared to standard two-lane highways, the influence of driveways was more pronounced, and the models emphasized the critical role of access management and cross-sectional design in improving safety.

The key findings from short duration analysis are below:

- Crash risk is consistently increased by higher MADT, greater speed variability, and precipitation, while wider shoulders and higher average speeds (in well-designed roads) tend to reduce crash counts. Surface width has a mixed impact, often increasing crashes on narrow roads (e.g., 2U, 3T) but reducing them on divided roadways.
- While core predictors remain stable (MADT, speed variability), their magnitudes and significance vary by year. For instance, shoulder width had a stronger negative impact on crash counts in 2019–2021 than in 2022 for certain facilities (e.g., 2U roads), and the effect of precipitation fluctuated across years, sometimes becoming non-significant (e.g., in 2020).
- 2U roads are more sensitive to shoulder width and commercial driveways, while 4U and 4D roadways show pronounced effects from speed variability. Three- and five-lane roads (3T, 5T) exhibit the highest crash increases from travel time variability, suggesting a need for better signal timing and access control.
- The study used Negative Binomial-Lindley models to better handle excess zeros in crash data, outperforming traditional NB models. This hierarchical modeling approach improves fit and interpretability for short-duration monthly datasets.
- Unlike traditional SPFs based on yearly averages, these short-duration SPFs capture seasonal dynamics, monthly variability in traffic, speed, and weather, enabling more targeted, time-sensitive safety interventions at a granular level for each road type.

6.4 RECOMMENDATIONS

6.4.1 Recommendations for TxDOT Implementation

- Update urban roadway SPFs for Texas by incorporating key operational parameters such as operating speed, speed variability, travel time variability, monthly traffic volumes, and precipitation. These variables demonstrated a strong influence on crash frequency across urban road types, including 2U, 4U, 4D, 3T, and 5T facilities.
- Refine safety analysis guidelines to emphasize not just average speed but also speed consistency (standard deviation of speed) and travel time variability, especially for urban multilane and turn-lane facilities where crash risk increases significantly with inconsistent flow patterns.
- Leverage the decision support tool to enable countermeasures for urban corridors, such as adjusting signal timing, access management, and cross-section design. The tool should

provide roadway-specific risk profiles at both annual and short-duration (e.g., monthly or seasonal) levels.

6.4.2 Updated Recommendations for Future Tool Enhancements

- Enhance the decision support tool's data integration features, allowing TxDOT engineers and planners to upload short-duration datasets (e.g., monthly or daily MADT, weather, speed profiles) for real-time or retrospective evaluations.
- Integrate short-duration SPFs into the tool, enabling evaluation of monthly/seasonal crash patterns and temporal safety diagnostics (e.g., AASHTOWare Safety Analyst does not currently support at such a fine resolution).
- Build modules for daily-level modeling by incorporating emerging data sources (e.g., probe-based traffic counts, connected vehicle data) to evaluate rapidly changing urban conditions like peak-period safety performance.
- Develop a dedicated urban safety evaluation module within the web-based decision support tool, enabling facility-type specific recommendations (e.g., 3T roads needing consistent speed management vs. 2U roads sensitive to commercial access). This module should offer interactive visualization and intervention simulation capabilities.

6.4.3 Comparison with AASHTOWare Safety Analyst

- Unlike AASHTOWare, which primarily relies on annualized SPFs and lacks weather or short-duration modeling integration, the TxDOT Decision Support Tool incorporates both annual and short-duration SPFs, flexible covariate structures (e.g., MADT, speed variability), and support for operational data fusion (weather, driveway density, etc.).
- 0-7144 project's use of Negative Binomial-Lindley (NBL) models provides superior handling of overdispersion and excess zeros, offering more robust predictions than the standard GLM-based AASHTOWare approach.
- The urban-specific modules in the TxDOT tool address key facility design variations (e.g., continuous left-turn lanes, multilane undivided/divided configurations), providing TxDOT with a more granular, localized, and adaptable solution for urban safety planning.
- Conduct an implementation project on the comparison between AASHTOWare safety scores with safety scores generated from 0-7144 project.

REFERENCES

Aarts, L., Schagen, I., 2006. Driving speed and the risk of road crashes: A review. Accident Analysis & Prevention 38, 215–224.

AASHTO, 2010. Highway Safety Manual. American Association of State Highway and Transportation Officials, Washington, DC.

AASHTO, 2014. Highway Safety Manual. Chapter 18 and Chapter 19. American Association of State Highway and Transportation Officials, Washington, DC.

AASHTO, 2004. A Policy on Geometric Design of Highways and Streets. American Association State Highway, Washington, DC.

Abdel-Aty, M.A., Radwan, A.E., 2000. Modeling traffic accident occurrence and involvement. Accident Analysis & Prevention 32, 633–642.

Abdelwahab, H.T., Abdel-Aty, M.A., 2002. Artificial neural networks and logit models for traffic safety analysis of toll plazas. Transportation Research Record: Journal of the Transportation Research Board 1784, 115–125. https://doi.org/10.3141/1784-15

Abdelwahab, H.T., Abdel-Aty, M.A., 2001. Development of artificial neural network models to predict driver injury severity in traffic accidents at signalized intersections. Transportation Research Record: Journal of the Transportation Research Board 1746, 6–13. https://doi.org/10.3141/1746-02

Abdulhafedh, A., 2017. Road Crash Prediction Models: Different Statistical Modeling Approaches. Journal of Transportation Technologies 7, 190–205. https://doi.org/10.4236/jtts.2017.72014

Abohassan, A., El-Basyouny, K., Kwon, T.J., 2022. Effects of Inclement Weather Events on Road Surface Conditions and Traffic Safety: An Event-Based Empirical Analysis Framework. Transportation Research Record: Journal of the Transportation Research Board 2676, pp 51-62. https://doi.org/10.1177/03611981221088588

Alhomaidat, F., Kwigizile, V., Oh, J.-S., 2021. Impacts of freeway speed limit on operation speed of adjacent arterial roads. IATSS Research 45, pp 161-168. https://doi.org/10.1016/j.iatssr.2020.08.007

Alhomaidat, F., Kwigizile, V., Oh, J.-S., Houten, R.V., 2020. How Does an Increased Freeway Speed Limit Influence the Frequency of Crashes on Adjacent Roads? Accident Analysis & Prevention 136, 105433. https://doi.org/10.1016/j.aap.2020.105433

Alkheder, S., Taamneh, M., Taamneh, S., 2017. Severity Prediction of Traffic Accident Using an Artificial Neural Network. Journal of Forecasting 36, 100–108. https://doi.org/10.1002/FOR.2425

Anderson, C.E., Zimmerman, A., Lewis, S., Marmion, J., Gustat, J., 2019. Patterns of cyclist and pedestrian street crossing behavior and safety on an urban greenway. International Journal of Environmental Research and Public Health 16, Article ID 201. https://doi.org/10.3390/ijerph16020201

Andrey, J., 2010. Long-term trends in weather-related crash risks. Journal of Transport Geography 18, 247–258. https://doi.org/10.1016/j.jtrangeo.2009.05.002

Arefkhani, H., Besharati, M. mehdi, Azizi Bondarabadi, M., Tavakoli Kashani, A., 2021. How does the incompatibility of different vehicle types affect the odds of driver injury? Journal of Transportation Safety & Security 13, 860–876. https://doi.org/10.1080/19439962.2019.1691101

Asadi-Shekari, Z., Moeinaddini, M., Zaly Shah, M., 2015. Pedestrian Safety Index for Evaluating Street Facilities in Urban Areas. Safety Science 74, pp 1-14. https://doi.org/10.1016/j.ssci.2014.11.014

Asil, M.R., Bargegol, I., 2022. Investigating the Factors Affecting the Severity of Single-Vehicle Crashes on Urban Roads using Bayesian Binary Probit Regression. Computational Research Progress in Applied Science & Engineering (CRPASE) 8, 6p. https://doi.org/10.52547/crpase.8.2224

Asli, H.H., 2022. Investigation of the Factors Affecting Pedestrian Accidents in Urban Roundabouts. Computational Research Progress in Applied Science & Engineering (CRPASE) 8, 4p. https://doi.org/10.52547/crpase.8.1.2255

Bambrick, N., 2016. Support Vector Machines: A Simple Explanation. URL https://www.kdnuggets.com/2016/07/support-vector-machines-simple-explanation.html (accessed 7.26.18).

Beer, K., 2011. Striving for safe system speeds. p. 8p.

Bergel-Hayat, R., Debbarh, M., Antoniou, C., Yannis, G., 2013. Explaining the road accident risk: weather effects. Accid Anal Prev 60, 456–465. https://doi.org/10.1016/j.aap.2013.03.006

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

Blackburn, L., Zegeer, C.V., Brookshire, K., 2018. Guide for improving pedestrian safety at uncontrolled crossing locations (No. FHWA-SA-17-072). FHWA, Washington, DC.

Bonneson, J., Lord, D., Zimmerman, K., Fitzpatrick, K., Pratt, M., 2007. Development of tools for evaluating the safety implications of highway design decisions. FHWA, Washington, DC.

Brijs, T., Karlis, D., Wets, G., 2008. Studying the effect of weather conditions on daily crash counts using a discrete time-series model. Accid Anal Prev 40, 1180–1190. https://doi.org/10.1016/j.aap.2008.01.001

Buddhavarapu, P., Banerjee, A., Prozzi, J.A., 2013. Influence of pavement condition on horizontal curve safety. Accident Analysis & Prevention 52, 9–18. https://doi.org/10.1016/j.aap.2012.12.010

Cafiso, S., Montella, A., D'Agostino, C., Mauriello, F., Galante, F., 2021. Crash modification functions for pavement surface condition and geometric design indicators. Accident Analysis & Prevention 149, 105887. https://doi.org/10.1016/j.aap.2020.105887

Cai, Q., Abdel-Aty, M., Mahmoud, N., Ugan, J., Al-Omari, M.M.A., 2021. Developing a grouped random parameter beta model to analyze drivers' speeding behavior on urban and suburban arterials with probe speed data. Accident Analysis & Prevention 161. https://doi.org/10.1016/j.aap.2021.106386

Candappa, N., Colobong, R., Bui, B., 2016. The use of a safety platform at an intersection to align approach speeds with safe system: an Australian-first innovative trial. p. 2p.

Candappa, N., Corben, B., van Nes, N., Logan, D., 2013. Innovative intersection designs better aligning with safe system. Swedish National Road and Transport Research Institute (VTI), Sweden.

Carson, J., Mannering, F., 2001. The effect of ice warning signs on ice-accident frequencies and severities. Accident Analysis & Prevention 33, 99–109. https://doi.org/10.1016/S0001-4575(00)00020-8

Cerný, J., Vanžura, M., Konšelová, T., Skokan, A., 2023. Influence of Speed and Other External Factors on the Functioning of Automatic Emergency Braking During Unexpected Interactions With Pedestrians. IEEE Intelligent Transportation Systems Magazine 15, pp 126-135. https://doi.org/10.1109/MITS.2022.3164538

Chang, I., Park, H., Hong, E., Lee, J., Kwon, N., 2022. Predicting effects of built environment on fatal pedestrian accidents at location-specific level: Application of XGBoost and SHAP. Accident Analysis & Prevention 166, 106545. https://doi.org/10.1016/j.aap.2021.106545

Chang, L.Y., Chien, J.T., 2013. Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. Safety Science 51, 17–22. https://doi.org/10.1016/j.ssci.2012.06.017

Chen, Z., Fan, W. (David), 2019. Modeling Pedestrian Injury Severity in Pedestrian-Vehicle Crashes in Rural and Urban Areas: Mixed Logit Model Approach. Transportation Research Record: Journal of the Transportation Research Board 2673, pp 1023-1034. https://doi.org/10.1177/0361198119842825

Cloutier, M.-S., Lachapelle, U., 2021. The effect of speed reductions on collisions: A controlled before-and-after study in Quebec, Canada. Journal of Transport & Health 22. https://doi.org/10.1016/j.jth.2021.101137

Dai, D., Taquechel, E.P., Steward, J., Strasser, S., 2010. The Impact of Built Environment on Pedestrian Crashes and the Identification of Crash Clusters on an Urban University Campus. Western Journal of Emergency Medicine 11, pp 294-301.

Daniels, S., Brijs, T., Nuyts, E., Wets, G., 2010. Explaining variation in safety performance of roundabouts. Accident Analysis & Prevention 42, 393–402. https://doi.org/10.1016/j.aap.2009.08.019

Das, A., Khan, M.N., Ahmed, M.M., 2022. A Deep Learning Approach for Detecting Lane Change Maneuvers Using SHRP2 Naturalistic Driving Data. Transportation Research Record: Journal of the Transportation Research Board.

Das, A., Murthy, S.S., Suddamalla, U., 2017. Enhanced Algorithm of Automated Ground Truth Generation and Validation for Lane Detection System by M2BMT. IEEE Transactions on Intelligent Transportation Systems 18, pp 996-1005. https://doi.org/10.1109/TITS.2016.2594055

Das, S., Dixon, K., Sun, X., Dutta, A., Zupancich, M., 2017. Trends in Transportation Research: Exploring Content Analysis in Topics. Transportation Research Record 2614, 27–38. https://doi.org/10.3141/2614-04

Das, S., Dutta, A., Avelar, R., Dixon, K., Sun, X., 2019. Supervised association rules mining on pedestrian crashes in urban areas : identifying patterns for appropriate countermeasures. International Journal of Urban Sciences 23, 30–48.

Das, S., Dutta, A., Dixon, K., Minjares-Kyle, L., Gillette, G., 2018. Using Deep Learning in Severity Analysis of At-Fault Motorcycle Rider Crashes. Transportation Research Record 2672, 122–134. https://doi.org/10.1177/0361198118797212

Das, S., Khodadadi, A., Liu, J., 2023. Short-Duration Crash Modeling to Understand the Impact of Operating Speed on Freeway Crashes During COVID-19. Transportation Research Record 03611981231169283.

Das, S., Le, M., Fitzpatrick, K., Wu, D., 2022. Did operating speeds during covid-19 result in more fatal and injury crashes on urban freeways? Transportation Research Record: Journal of the Transportation Research Board. https://doi.org/10.1177/03611981221109597

Das, S., Sun, X., Dutta, A., 2016. Text Mining and Topic Modeling of Compendiums of Papers from Transportation Research Board Annual Meetings. Transportation Research Record 2552, 48–56. https://doi.org/10.3141/2552-07

Dash, I., Abkowitz, M., Philip, C., 2022. Factors Impacting Bike Crash Severity in Urban Areas. Journal of Safety Research 83, pp 128-138. https://doi.org/10.1016/j.jsr.2022.08.010

Davidse, R.J., Van Duijvenvoorde, K., Boele-Vos, M.J., Louwerse, W.J.R., Stelling-Konczak, A., Duivenvoorden, C.W.A.E., Algera, A.J., 2019. Scenarios of Crashes Involving Light Mopeds on Urban Bicycle Paths. Accident Analysis & Prevention 129, pp 334-341. https://doi.org/10.1016/j.aap.2019.05.016

Ding, H., Sze, N.N., Li, H., Guo, Y., 2020. Roles of Infrastructure and Land Use in Bicycle Crash Exposure and Frequency: A Case Study Using Greater London Bike Sharing Data. Accident Analysis & Prevention 144. https://doi.org/10.1016/j.aap.2020.105652

Donnell, E.T., Hines, S.C., Mahoney, K.M., Porter, R.J., McGee, H., 2009. Speed concepts: Informational guide. Office of Safety. FHWA, Washington, DC.

Donnell, E.T., Kersavage, K., Tierney, L.F., 2018. Self-Enforcing Roadways: A Guidance Report. FHWA, Washington, DC.

Doustmohammadi, M., 2019. Urban freight crash analysis using ordinal logit and ordinal probit regression in the state of Alabama. International Journal for Traffic and Transport Engineering 9, pp 127-144. https://doi.org/10.7708/ijtte.2019.9(1).10

Dumbaugh, E., Li, W., 2011. Designing for the Safety of Pedestrians, Cyclists, and Motorists in Urban Environments. Journal of the American Planning Association 77, pp 69-88.

Dutta, N., 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., Islam, Md.T., 2014a. Investigation of time and weather effects on crash types using full Bayesian multivariate Poisson lognormal models. Accident Analysis & Prevention 73C, 91–99. https://doi.org/10.1016/j.aap.2014.08.014

El-Basyouny, K., Barua, S., Islam, Md.T., 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. https://doi.org/10.3141/2432-08

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., 2010. A restatement of the case for speed limits. Transport Policy 17, 196–204.

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

Farid, A., Abdel-Aty, M., Lee, J., 2018. A new approach for calibrating safety performance functions. Accident Analysis & Prevention 119, 188–194. https://doi.org/10.1016/j.aap.2018.07.023

Ferenchak, N.N., Marshall, W., 2017. Redefining the child pedestrian safety paradigm: identifying high fatality concentrations in urban areas. Injury Prevention 23, pp 364-369. https://doi.org/10.1136/injuryprev-2016-042115

Ferreira, M.C., Costa, P.D., Abrantes, D., Hora, J., Felício, S., Coimbra, M., Dias, T.G., 2022. Identifying the determinants and understanding their effect on the perception of safety, security, and comfort by pedestrians and cyclists: A systematic review. Transportation Research Part F: Traffic Psychology and Behaviour 91, pp 136-163. https://doi.org/10.1016/j.trf.2022.10.004

FHWA, 2022. Pedestrian & Bicycle Safety. URL https://highways.dot.gov/safety/pedestrian-bicyclist (accessed 5.12.23).

FHWA, 2021. Safe Transportation for Every Pedestrian (STEP). URL https://highways.dot.gov/safety/pedestrian-bicyclist/step (accessed 5.12.23).

FHWA, 2016. Speed Limit Basics (No. FHWA-SA-16-076). FHWA, Washington DC.

Fildes, B., Langford, J., Szwed, N., 2006. Setting speed limits for a safe system. Presented at the Saferoads: Victorian Local Government Road Safety Conference, 2006, Melbourne, Victoria, Australia, p. 11P.

Fildes, B.N., Rumbold, G., Leening, A., 1991. Speed behaviour and drivers' attitude to speeding. Monash University Accident Research Centre, Report 16.

Finkel, E., McCormick, C., Mitman, M., Abel, S., Clark, J., 2020. Integrating the Safe System Approach with the Highway Safety Improvement Program: An Informational Report.

Garber, N.J., Srinivasan, S., 1998. Influence of exposure duration on the effectiveness of changeable-message signs in controlling vehicle speeds at work zones. Transportation Research Record 1650, 62–70.

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

Geedipally, S., Brewer, M., Wunderlich, R., Pratt, M., Wu, L., Das, S., and Florence, D., 2021. Examine Trade-Offs between Center Separation and Shoulder Width Allotment for a Given Roadway Width. Report: FHWA/TX-21/0-7035-R1.

Geedipally, S., Dixon, K., Pratt, M., Wu, L., Avelar, R., Das, S., Tsapakis, I., Lord, D., and Saini, G., 2022. Calibrating the Highway Safety Manual Predictive Methods for Texas Highways. Report: FHWA/TX-23/0-7083-R1.

Gim, T.-H.T., 2022. Generalized ordered logit analysis on the injury severity of traffic crashes by older drivers: the case of South Korea. International Journal of Urban Sciences 26, pp 499-516. https://doi.org/10.1080/12265934.2022.2042363

Goel, R., 2021. A new model to estimate pedestrian deaths from speed-related interventions. Traffic Injury Prevention 22, 330–335. https://doi.org/10.1080/15389588.2021.1908544

Gonzalo-Orden, H., Rojo Arce, M., Linares Unamunzaga, A., Aponte, N., Pérez-Acebo, H., 2021. Why is necessary to reduce the speed in urban areas to 30 Km/h? Presented at the Transportation Research Procedia, pp 209-216. https://doi.org/10.1016/j.trpro.2021.11.029

Goswamy, A., Abdel-Aty, M., Islam, Z., 2023. Factors affecting injury severity at pedestrian crossing locations with Rectangular RAPID Flashing Beacons (RRFB) using XGBoost and random parameters discrete outcome models. Accident Analysis & Prevention 181, 106937. https://doi.org/10.1016/j.aap.2022.106937

Grabowski, D.C., Morrisey, M.A., 2007. Systemwide implications of the repeal of the national maximum speed limit. Accident Analysis & Prevention 39, 180–189.

Guo, Y., Liu, P., Liang, Q., Wang, W., 2016. Effects of parallelogram-shaped pavement markings on vehicle speed and safety of pedestrian crosswalks on urban roads in China. Accident Analysis & Prevention 95, Part B, pp 438-447. https://doi.org/10.1016/j.aap.2015.07.001

Haddon Jr, W., 1972. A logical framework for categorizing highway safety phenomena and activity. Journal of Trauma and Acute Care Surgery 12, 193–207.

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

Haleem, K., Gan, A., 2013. Effect of Driver's Age and Side of Impact on Crash Severity along Urban Freeways: A Mixed Logit Approach. Journal of Safety Research 46, pp 67-76. https://doi.org/10.1016/j.jsr.2013.04.002

Hall, C., Beer, K., Robertson, J., Nguyen, T., Zafar, F., Tan, T., Mani, A., Beer, T., 2021. Guide to road safety part 3: safe speed. AustRoad, Australia.

Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. The Mathematical Intelligencer. https://doi.org/10.1198/jasa.2004.s339

Hauer, E., 2009. Speed and safety. Transportation Research Record 2103, 10–17.

Hauer, E., 1971. Accidents, overtaking and speed control. Accident Analysis & Prevention 3, 1–13.

Heydari, S., Miranda-Moreno, L.F., Liping, F., 2014. Speed limit reduction in urban areas: A before–after study using Bayesian generalized mixed linear models. Accident Analysis & Prevention 73, pp 252-261. https://doi.org/10.1016/j.aap.2014.09.013

Hu, W., Cicchino, J.B., 2020. Lowering the Speed Limit from 30 mph to 25 mph in Boston: Effects on Vehicle Speeds. Injury Prevention 26, pp 99-102. https://doi.org/10.1136/injuryprev-2018-043025

Hutton, J.M., Cook, D.J., Grotheer, J., Conn, M., 2020. Research Utilizing SHRP2 Data to Improve Highway Safety: Development of Speed—Safety Relationships. United States. Federal Highway Administration. Office of Safety Research and Development, McLean, Virginia.

Imprialou, M.I.M., Quddus, M., Pitfield, D.E., Lord, D., 2016. Re-visiting crash–speed relationships: A new perspective in crash modelling. Accident Analysis & Prevention 86, 173–185.

Intini, P., Berloco, N., Fonzone, A., Fountas, G., Ranieri, V., 2020. The Influence of Traffic, Geometric and Context Variables on Urban Crash Types: A Grouped Random Parameter Multinomial Logit Approach. Analytic Methods in Accident Research 28. https://doi.org/10.1016/j.amar.2020.100141

Iranitalab, A., Khattak, A., 2017. Comparison of four statistical and machine learning methods for crash severity prediction. Accident Analysis & Prevention 108, 27–36. https://doi.org/10.1016/j.aap.2017.08.008

Isaksson-Hellman, I., Töreki, J., 2019. The Effect of Speed Limit Reductions in Urban Areas on Cyclists' Injuries in Collisions with Cars. Traffic Injury Prevention 20, pp 39-44. https://doi.org/10.1080/15389588.2019.1680836

Islam, M.T., El-Basyouny, K., 2015. Full Bayesian evaluation of the safety effects of reducing the posted speed limit in urban residential area. Accident Analysis & Prevention 80, pp 18-25. https://doi.org/10.1016/j.aap.2015.02.026

Jackson, T.L., Sharif, H.O., 2016. Rainfall impacts on traffic safety: rain-related fatal crashes in Texas. Geomatics, Natural Hazards and Risk 7, 843–860. https://doi.org/10.1080/19475705.2014.984246

Jiang, F., Ma, J., 2021. A comprehensive study of macro factors related to traffic fatality rates by XGBoost-based model and GIS techniques. Accident Analysis & Prevention 163, 106431. https://doi.org/10.1016/j.aap.2021.106431

Jin, P., Walker, A., Cebelak, M., Walton, C., 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. https://doi.org/10.3141/2440-05

Jung, S., Qin, X., Noyce, D., 2012. Injury Severity of Multivehicle Crash in Rainy Weather. Journal of Transportation Engineering 138, 50–59. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000300

Jurewicz, C., 2010. Speed limits in the safe system concept. Journal of the Australasian College of Road Safety 21, 14–7.

Jurewicz, C., Hall, K., 2009. Speed limit setting principles in the safe system context. Presented at the Australasian Road Safety Research Policing Education Conference, Sydney, New South Wales, Australia, p. 15p.

Jurewicz, C., Sobhani, A., Woolley, J., Dutschke, J., Corben, B., 2015. Proposed vehicle impact speed - severe injury probability relationships for selected crash types. Presented at the Australasian Road Safety Conference.

Jurewicz, C., Turner, B., 2011. Risk-based approach to speed limits: a step towards Safe System. p. 17p.

Kabir, R., Remias, S.M., Lavrenz, S.M., Waddell, J., 2021. Assessing the impact of traffic signal performance on crash frequency for signalized intersections along urban arterials: A random parameter modeling approach. Accident Analysis & Prevention 149. https://doi.org/10.1016/j.aap.2020.105868

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

Khan, G., Qin, X., Noyce, D., 2008. Spatial Analysis of Weather Crash Patterns. Journal of Transportation Engineering-asce - J TRANSP ENG-ASCE 134. https://doi.org/10.1061/(ASCE)0733-947X(2008)134:5(191)

Khan, M.N., Das, A., Ahmed, M.M., 2020. Non-Parametric Association Rules Mining and Parametric Ordinal Logistic Regression for an In-Depth Investigation of Driver Speed Selection Behavior in Adverse Weather using SHRP2 Naturalistic Driving Study Data. Transportation Research Record: Journal of the Transportation Research Board 2674, 101–119. https://doi.org/10.1177/0361198120941509

Khattak, M.W., Pirdavani, A., De Winne, P., Brijs, T., De Backer, H., 2021. Estimation of safety performance functions for urban intersections using various functional forms of the negative binomial regression model and a generalized Poisson regression model. Accident Analysis & Prevention 151, 105964. https://doi.org/10.1016/j.aap.2020.105964

Kim, D.-G., Washington, S., 2006. The significance of endogeneity problems in crash models: An examination of left-turn lanes in intersection crash models. Accident Analysis & Prevention 38, 1094–1100. https://doi.org/10.1016/j.aap.2006.04.017

Kloeden, C., Woolley, J., McLean, J., 2007. A follow up evaluation of the 50km/h default urban speed limit in South Australia. p. 12p.

Kloeden, C., Woolley, J., McLean, J., 2004. Evaluation of the 50km/h default urban speed limit in South Australia. Road Safety Research, Policing and Education Conference, 2004, Perth, Western Australia, Australia 1, 11p.

Kloeden, C.N., McLean, J., Glonek, G.F.V., 2002. Reanalysis of travelling speed and the risk of crash involvement in Adelaide South Australia (No. CR 207). Australian Transport Safety Bureau.

Kloeden, C.N., Woolley, J.E., University of Adelaide. Centre for Automotive Safety Research (CASR), 2017. Vehicle speeds in South Australia 2016 (No. 9781921645822).

Koloushani, M., Karaer, A., Erman Ozguven, E., Sando, T., Dulebenets, M.A., Moses, R., 2022. Investigating Spatial Correlations Between Land Use and Pedestrian Injury Severity in Crashes Occurring Away From Intersections in Northwest Florida. Transportation Research Record:

Journal of the Transportation Research Board 2676, pp 599-614. https://doi.org/10.1177/03611981221096433

Koopmans, J.M., Friedman, L., Kwon, S., Sheehan, K., 2015. Urban crash-related child pedestrian injury incidence and characteristics associated with injury severity. Accident Analysis & Prevention 77, pp 127-136. https://doi.org/10.1016/j.aap.2015.02.005

Kraidi, R., Evdorides, H., 2020. Pedestrian Safety Models for Urban Environments with High Roadside Activities. Safety Science 130. https://doi.org/10.1016/j.ssci.2020.104847

Kristen Brookshire, Sandt, L., Sundstrom, C., Thomas, L., Blomberg, R., 2016. Advancing Pedestrian and Bicyclist Safety : A Primer for Highway Safety Professionals. NHTSA. Washington, DC.

Kuhnert, P.M., Do, K.A., McClure, R., 2000. Combining non-parametric models with logistic regression: An application to motor vehicle injury data. Computational Statistics & Data Analysis 34, 371–386. https://doi.org/10.1016/S0167-9473(99)00099-7

Kumara, S.S.P., Chin, H.C., 2003. Modeling Accident Occurrence at Signalized Tee Intersections with Special Emphasis on Excess Zeros. Traffic Injury Prevention 4, 53–57. https://doi.org/10.1080/15389580309852

Kuşkapan, E., Çodur, M.Y., Atalay, A., 2021. Speed violation analysis of heavy vehicles on highways using spatial analysis and machine learning algorithms. Accident Analysis & Prevention 155, 106098. https://doi.org/10.1016/j.aap.2021.106098

Kwayu, K.M., Kwigizile, V., Oh, J.-S., Cho, H., Bott, M., 2018. Evaluating the Impact of Raising Speed Limit on Urban Freeways Using Mixed Effects Negative Binomial Regression. p. 15p.

Lagerwey, P.A., Hintze, M.J., Elliott, J.B., Toole, J.L., Schneider, R.J., 2015. Pedestrian and Bicycle Transportation Along Existing Roads - ActiveTrans Priority Tool Guidebook. Transportation Research Board, Washington, D.C. https://doi.org/10.17226/22163

Lee, H., Seo, J., Kassas, Z.Z.M., 2022. Urban Road Safety Prediction: A Satellite Navigation Perspective. IEEE Intelligent Transportation Systems Magazine 14, pp 94-106. https://doi.org/10.1109/MITS.2022.3181557

Lee, J., Chae, J., Yoon, T., Yang, H., 2018. Traffic accident severity analysis with rain-related factors using structural equation modeling - A case study of Seoul City. Accid Anal Prev 112, 1–10. https://doi.org/10.1016/j.aap.2017.12.013

Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. Accident Analysis & Prevention 34, 149–161. https://doi.org/10.1016/S0001-4575(01)00009-4

Levin, M., Chen, R.C., Liao, C.F.L., Zhang, T., 2019. Improving intersection safety through variable speed limits for connected vehicles. Center for Transportation Studies, University of Minnesota, Minneapolis, Minnesota.

Liu, C., Zhao, M., Li, W., Sharma, A., 2018. Multivariate random parameters zero-inflated negative binomial regression for analyzing urban midblock crashes. Analytic Methods in Accident Research 17, pp 32-46. https://doi.org/10.1016/j.amar.2018.03.001

Lord, D., Guikema, S.D., Geedipally, S.R., 2008. Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes. Accident Analysis & Prevention 40, 1123–1134. https://doi.org/10.1016/j.aap.2007.12.003

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation Research Part A: Policy and Practice 44, 291–305. https://doi.org/10.1016/j.tra.2010.02.001

Malin, F., Norros, I., Innamaa, S., 2019. Accident risk of road and weather conditions on different road types. Accident Analysis & Prevention 122, 181–188. https://doi.org/10.1016/j.aap.2018.10.014

Masello, L., Castignani, G., Sheehan, B., Murphy, F., McDonnell, K., 2022. On the road safety benefits of advanced driver assistance systems in different driving contexts. Transportation Research Interdisciplinary Perspectives 15, 100670. https://doi.org/10.1016/j.trip.2022.100670

Mathew, S., Pulugurtha, S.S., Duvvuri, S., 2022. Exploring the effect of road network, demographic, and land use characteristics on teen crash frequency using geographically weighted negative binomial regression. Accident Analysis & Prevention 168, 106615. https://doi.org/10.1016/j.aap.2022.106615

Maycock, G., Brocklebank, P.J., Hall, R.D., 1998. Road layout design standards and driver behaviour (TRL REPORT 332,). Transportation Research Laboratory, Berkshire, England.

Mclean, A.J., Anderson, R.W., 2008. Metrication of the urban speed limit and pedestrian fatalities. Australasian Road Safety Research Policing Education Conference, 2008, Adelaide, South Australia, Australia 7p.

Mitra, S., Job, S., Han, S., Eom, K., 2021. Do Speed Limit Reductions Help Road Safety? Lessons from the Republic of Korea's Recent Move to Lower Speed Limit on Urban Roads, Mobility and Transport Connectivity Series. World Bank Group.

Mohammadnazar, A., Arvin, R., Khattak, A.J., 2021. Classifying travelers' driving style using basic safety messages generated by connected vehicles: Application of unsupervised machine learning. Transportation Research Part C: Emerging Technologies 122, 102917. https://doi.org/10.1016/j.trc.2020.102917

Mooren, L., Grzebieta, R., Job, S., 2014. Speed limit setting and the Safe System principle. p. 11p.

Mukoko, K.K., Pulugurtha, S.S., 2020. Examining the influence of network, land use, and demographic characteristics to estimate the number of bicycle-vehicle crashes on urban roads. IATSS Research 44, pp 8-16. https://doi.org/10.1016/j.iatssr.2019.04.001

Najafi, S., Flintsch, G., Medina, A., 2015. Linking roadway crashes and tire–pavement friction: a case study. International Journal of Pavement Engineering 18. https://doi.org/10.1080/10298436.2015.1039005

Ngo, C., Milton, J., Reynolds, L., Carpenter, R., Veka, C., 2022. The Safe System Approach: How States and Cities Are Saving Lives. Public Roads 85.

NHTSA, 2020. Speeding and Speed Management. URL https://www.nhtsa.gov/book/countermeasures/countermeasures-work/speeding-and-speed-management (accessed 5.16.23).

Nilsson, G., 2004. Traffic safety dimensions and the power model to describe the effect of speed on safety. Lund Institute of Technology.

Norrman, J., Eriksson, M., Lindqvist, S., 2000. Relationships between road slipperiness, traffic accident risk and winter road maintenance activity. Clim. Res. 15, 185–193. https://doi.org/10.3354/cr015185

Olowosegun, A., Babajide, N., Akintola, A., Fountas, G., Fonzone, A., 2022. Analysis of pedestrian accident injury-severities at road junctions and crossings using an advanced random parameter modelling framework: The case of Scotland. Accident Analysis & Prevention 169. https://doi.org/10.1016/j.aap.2022.106610

Park, E.S., Fitzpatrick, K., Das, S., Avelar, R., 2021. Exploration of the relationship among roadway characteristics, operating speed, and crashes for city streets using path analysis. Accident Analysis & Prevention 150, 105896. https://doi.org/10.1016/j.aap.2020.105896

Park, E.S., Lord, D., 2007. Multivariate Poisson-Lognormal Models for Jointly Modeling Crash Frequency by Severity. Transportation Research Record 2019, 1–6. https://doi.org/10.3141/2019-01

Parsa, A.B., Movahedi, A., Taghipour, H., Derrible, S., Mohammadian, A. (Kouros), 2020. Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. Accident Analysis & Prevention 136, 105405. https://doi.org/10.1016/j.aap.2019.105405

Pauw, E.D., Daniels, S., Franckx, L., Mayeres, I., 2018. Safety effects of dynamic speed limits on motorways. Accident Analysis & Prevention 114, 83–89.

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

Pobudzei, M., Tießler, M., Sellaouti, A., Hoffmann, S., Transportation Research Board, 2023. E-Scooter and Bicycle Accidents: Spatial, Temporal, and Demographic Characteristics in Munich, Germany. p. 14p.

Prati, G., Pietrantoni, L., Fraboni, F., 2017. Using data mining techniques to predict the severity of bicycle crashes. Accident Analysis & Prevention 101, pp 44-54. https://doi.org/10.1016/j.aap.2017.01.008

Pulugurtha, S.S., Duddu, V.R., Kotagiri, Y., 2013. Traffic analysis zone level crash estimation models based on land use characteristics. Accident Analysis & Prevention 50, pp 678-687. https://doi.org/10.1016/j.aap.2012.06.016

Qin, X., Noyce, D., Lee, C., Kinar, J., 2006. Snowstorm Event-Based Crash Analysis. Transportation Research Record 1948, 135–141. https://doi.org/10.3141/1948-15

Qiu, B., Fan, W.D., 2022. Mixed logit models for examining pedestrian injury severities at intersection and non-intersection locations. Journal of Transportation Safety & Security 14, pp 1333-1357. https://doi.org/10.1080/19439962.2021.1923101

Rahim, M.A., Hassan, H.M., 2021. A deep learning based traffic crash severity prediction framework. Accident Analysis and Prevention 154, 106090. https://doi.org/10.1016/j.aap.2021.106090

Raihan, M.A., Alluri, P., Wu, W., Gan, A., 2019. Estimation of bicycle crash modification factors (CMFs) on urban facilities using zero inflated negative binomial models. Accident Analysis & Prevention 123, pp 303-313. https://doi.org/10.1016/j.aap.2018.12.009

Rakotonirainy, A., Chen, S., Scott-Parker, B., Loke, S.W., Krishnaswamy, S., 2015. A Novel Approach to Assessing Road-Curve Crash Severity. Journal of Transportation Safety and Security 7, 358–375. https://doi.org/10.1080/19439962.2014.959585

Richard, C., Campbell, J.L., Brown, J.L., Lichty, M.G., Chrysler, S.T., Atkins, R., 2013a. Investigating Speeding Behavior with Naturalistic Approaches: Methodological Lessons Learned. Transportation Research Record 2365, 58–65. https://doi.org/10.3141/2365-08

Richard, C., Campbell, J.L., Lichty, M.G., Brown, J.L., Chrysler, S., Lee, J.D., Boyle, L., Reagle, G., 2013b. Motivations for Speeding, Volume II: Findings Report (No. DOT HS 811 818). United States. National Highway Traffic Safety Administration. Office of ….

Richard, C., Divekar, G., Brown, J.L., 2016. Motivations for speeding - Additional data analysis.

Rista, E., Goswamy, A., Wang, B., Barrette, T., Hamzeie, R., Russo, B., Bou-Saab, G., Savolainen, P.T., 2018. Examining the Safety Impacts of Narrow Lane Widths on Urban/Suburban Arterials: Estimation of a Panel Data Random Parameters Negative Binomial Model. Journal of Transportation Safety & Security 10, pp 213-228. https://doi.org/10.1080/19439962.2016.1273291

Rosén, E., Sander, U., 2009. Pedestrian fatality risk as a function of car impact speed. Accident Analysis & Prevention 41, 536–542. https://doi.org/10.1016/j.aap.2009.02.002

Rossetti, M.A., Johnsen, M., 2011. Weather and Climate Impacts on Commercial Motor Vehicle Safety.

Saha, P., Ahmed, M.M., Young, R.K., 2015. Safety effectiveness of variable speed limit system in adverse weather conditions on challenging roadway geometry. Transportation Research Record 2521, 45–53.

Salazar-Miranda, A., Heine, C., Duarte, F., Schechtner, K., Ratti, C., 2022. Measuring the impact of slow zones on street life using social media. Cities 131, 104010. https://doi.org/10.1016/j.cities.2022.104010

Se, C., Champahom, T., Jomnonkwao, S., Chaimuang, P., Ratanavaraha, V., 2021. Empirical comparison of the effects of urban and rural crashes on motorcyclist injury severities: A correlated random parameters ordered probit approach with heterogeneity in means. Accident Analysis & Prevention 161. https://doi.org/10.1016/j.aap.2021.106352

Sellers, K.F., Shmueli, G., 2010. A Flexible Regression Model for Count Data. The Annals of Applied Statistics 4, 943–961.

Shankar, V.N., Ulfarsson, G.F., Pendyala, R.M., Nebergall, M.B., 2003. Modeling crashes involving pedestrians and motorized traffic. Safety Science 41, 627–640. https://doi.org/10.1016/S0925-7535(02)00017-6

Siddiqui, S., Al-Kaisy, A., Transportation Research Board, 2017. Assessing Potential Safety Benefits of an Advisory Variable Speed Limit System along an Urban Freeway Corridor. p. 21p.

Silvano, A.P., Bang, K.L., 2016. Impact of Speed Limits and Road Characteristics on Free-Flow Speed in Urban Areas. Journal of Transportation Engineering 142, 04015039. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000800

Sohn, S.Y., Lee, S.H., 2003. Data fusion, ensemble and clustering to improve the classification accuracy for the severity of road traffic accidents in Korea. Safety Science 41, 1–14. https://doi.org/10.1016/S0925-7535(01)00032-7

Sohn, S.Y., Shin, H., 2001. Pattern recognition for road traffic accident severity in Korea. Ergonomics 44, 107–117. https://doi.org/10.1080/00140130120928

Solomon, D.H., 1964. Accidents on Main Rural Highways: Related to Speed, Driver, and Vehicle. U.S. Department of Transportation, FHWA.

Son, S., Park, J., Park, S., Cho, J., Choi, S., 2022. Evaluation of Direct and Indirect Safety Effects of Speed-Limit Reduction on Urban Networks. Journal of Transportation Engineering, Part A: Systems 148, 04022071. https://doi.org/10.1061/JTEPBS.0000724

Stamatiadis, N., Kirk, A., Hartman, D., Jasper, J., Wright, S., King, M., Chellman, R., 2018. An expanded functional classification system for highways and streets (No. NCHRP Reserach Report 855). Washington. DC.

Steinmetz, L., Jurewicz, C., Davis, C., Beer, K., Hall, C., 2018. Delivering safe system outcomes in Mildura. p. 4p.

Stiles, J., Kar, A., Lee, J., Miller, H.J., 2023. Lower Volumes, Higher Speeds: Changes to Crash Type, Timing, and Severity on Urban Roads from COVID-19 Stay-at-Home Policies. Transportation Research Record: Journal of the Transportation Research Board 2677, pp 15-27. https://doi.org/10.1177/03611981211044454

Strong, C.K., Ye, Z., Shi, X., 2010. Safety Effects of Winter Weather: The State of Knowledge and Remaining Challenges. Transport Reviews 30, 677–699. https://doi.org/10.1080/01441640903414470

Sultana, S., 2018. Trends in driver speed behaviours on Perth metropolitan road network 2000 to 2018.

Sun, Jian, Sun, Jie, Chen, P., 2014. Use of Support Vector Machine Models for Real-Time Prediction of Crash Risk on Urban Expressways. Transportation Research Record: Journal of the Transportation Research Board pp 91-98. https://doi.org/10.3141/2432-11

Sun, Z., Xing, Y., Gu, X., Chen, Y., 2022. Influence factors on injury severity of bicycle-motor vehicle crashes: A two-stage comparative analysis of urban and suburban areas in Beijing. Traffic Injury Prevention 23, pp 118-124. https://doi.org/10.1080/15389588.2021.2024523

Susilawati, S., Wong, W.J., Pang, Z.J., 2023. Safety Effectiveness of Autonomous Vehicles and Connected Autonomous Vehicles in Reducing Pedestrian Crashes. Transportation Research Record: Journal of the Transportation Research Board 2677, pp 1605-1618. https://doi.org/10.1177/03611981221108984

Tankasem, P., Satiennam, T., Satiennam, W., Jaensirisak, S., Rujopakarn, W., 2022. Effects of automated speed control on speeding intention and behavior on mixed-traffic urban arterial roads. IATSS Research 46, 492–498. https://doi.org/10.1016/j.iatssr.2022.08.002

Tarko, A.P., Pineda-Mendez, R., Guo, Q., Purdue University, Indiana Department of Transportation, Federal Highway Administration, 2019. Predicting the Impact of Changing Speed Limits on Traffic Safety and Mobility on Indiana Freeways (Digital/other). https://doi.org/10.5703/1288284316922

Tay, R., 2015. A random parameters probit model of urban and rural intersection crashes. Accident Analysis & Prevention 84, pp 38-40. https://doi.org/10.1016/j.aap.2015.07.013

Taylor, M.C., Lynam, D.A., Baruya, A., 2000. The effects of drivers' speed on the frequency of road accidents. Transport Research Laboratory Berkshire, Crowthorne.

Tefft, B.C., 2013. Impact speed and a pedestrian's risk of severe injury or death. Accident Analysis & Prevention 50, 871–878. https://doi.org/10.1016/j.aap.2012.07.022

Theofilatos, A., 2019. Utilizing Real-time Traffic and Weather Data to Explore Crash Frequency on Urban Motorways: A Cusp Catastrophe Approach. Transportation Research Procedia, Urban Mobility – Shaping the Future Together mobil.TUM 2018 – International Scientific Conference on Mobility and Transport Conference Proceedings 41, 471–479. https://doi.org/10.1016/j.trpro.2019.09.078

Thomas, L., Sandt, L., Zegeer, C., Kumfer, W., Lang, K., Lan, B., Horowitz, Z., Butsick, A., Toole, J., Schneider, R.J., 2018. Systemic Pedestrian Safety Analysis (No. NCHRP Report 893).

Tingvall, C., Haworth, N., 1999. Vision Zero-An ethical approach to safety and mobility, in: 6th ITE International Conference Road Safety & Traffic Enforcement: Beyond 2000.

Ullman, G.L., Rose, E.R., 2005. Evaluation of dynamic speed display signs. Transportation research record 1918, 92–97.

USDOT, 2022a. What Is a Safe System Approach?. URL https://www.transportation.gov/NRSS/SafeSystem (accessed 5.15.23).

USDOT, 2022b. Safer Speeds. URL https://www.transportation.gov/NRSS/SaferSpeeds (accessed 5.17.23).

Wahab, L., Jiang, H., 2019. A comparative study on machine learning based algorithms for prediction of motorcycle crash severity. PLOS ONE 14, e0214966. https://doi.org/10.1371/journal.pone.0214966

Wang, J., Cicchino, J.B., 2023. Changes in speeding on Virginia roads during the beginning of the COVID-19 pandemic. Traffic Injury Prevention 24, pp 38-43. https://doi.org/10.1080/15389588.2022.2127322

Wang, X., Kim, S.H., 2019. Prediction and factor identification for crash severity: Comparison of discrete choice and tree-based models. Transportation Research Record: Journal of the Transportation Research Board 2673, 640–653. https://doi.org/10.1177/0361198119844456

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

Weibull, K., Lidestam, B., Prytz, E., 2023. Potential of Cooperative Intelligent Transport System Services to Mitigate Risk Factors Associated with Emergency Vehicle Accidents. Transportation Research Record: Journal of the Transportation Research Board 2677, pp 999-1015. https://doi.org/10.1177/03611981221119459

Wen, H., Zhang, X., Zeng, Q., Sze, N.N., 2019. Bayesian spatial-temporal model for the main and interaction effects of roadway and weather characteristics on freeway crash incidence. Accid Anal Prev 132, 105249. https://doi.org/10.1016/j.aap.2019.07.025

WHO, 2017. Managing speed. World Health Organization.

Wramborg, P., 2005. A New Approach to a Safe and Sustainable Road Structure and Street Design for Urban Areas. Presented at the Road Safety on Four Continents: 13th International Conference.

Xie, B., An, Z., Zheng, Y., Li, Z., 2019. Incorporating transportation safety into land use planning: Pre-assessment of land use conversion effects on severe crashes in urban China. Applied Geography 103, pp 1-11. https://doi.org/10.1016/j.apgeog.2018.12.003

Xing, L., He, J., Li, Y., Wu, Y., Yuan, J., Gu, X., 2020. Comparison of different models for evaluating vehicle collision risks at upstream diverging area of toll plaza. Accident Analysis & Prevention 135, 105343. https://doi.org/10.1016/j.aap.2019.105343

Xu, C., Wang, C., Liu, P., 2018. Evaluating the Combined Effects of Weather and Real-Time Traffic Conditions on Freeway Crash Risks. Weather, Climate, and Society 10, 837–850. https://doi.org/10.1175/WCAS-D-17-0124.1

Xu, C., Wang, X., Yang, H., Xie, K., 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.

Xu, G., Zineddin, A., Atkins, R., Abel, S., 2022. Speed Management is Key to Road Safety. Public Roads 85.

Yadav, A.K., Velaga, N.R., 2020. Alcohol-impaired driving in rural and urban road environments: Effect on speeding behaviour and crash probabilities. Accident Analysis & Prevention 140, 105512. https://doi.org/10.1016/j.aap.2020.105512

Yang, C., Chen, M., Yuan, Q., 2021. The application of XGBoost and SHAP to examining the factors in freight truck-related crashes: An exploratory analysis. Accident Analysis and Prevention 158, 106153. https://doi.org/10.1016/j.aap.2021.106153

Yu, R., Abdel-Aty, M., 2014a. Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. Safety Science 63, 50–56. https://doi.org/10.1016/j.ssci.2013.10.012

Yu, R., Abdel-Aty, M., 2014b. Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. Safety Science 63, 50–56. https://doi.org/10.1016/j.ssci.2013.10.012

Yu, R., Abdel-Aty, M., 2013. Utilizing support vector machine in real-time crash risk evaluation. Accident Analysis & Prevention 51, 252–259. https://doi.org/10.1016/j.aap.2012.11.027

Yu, R., Abdel-Aty, M., 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, Rongjie, Abdel-Aty, M., Ahmed, M., 2013. Bayesian random effect models incorporating real-time weather and traffic data to investigate mountainous freeway hazardous factors. Accid Anal Prev 50, 371–376. https://doi.org/10.1016/j.aap.2012.05.011

Yu, R., Quddus, M., Wang, X., 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., 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.

Yu, R., Zheng, Y., Qin, Y., Peng, Y., 2019. Utilizing Partial Least-Squares Path Modeling to Analyze Crash Risk Contributing Factors for Shanghai Urban Expressway System. ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering 5, 05019001. https://doi.org/10.1061/AJRUA6.0001022

Yu, W., Guan, M., Chen, Z., 2019. Analyzing Spatial Community Pattern of Network Traffic Flow and Its Variations across Time Based on Taxi GPS Trajectories. Applied Sciences 9, 2054. https://doi.org/10.3390/app9102054

Zegeer, C., Lyon, C., Srinivasan, R., Persaud, B., Lan, B., Smith, S., Carter, D., Thirsk, N.J., Zegeer, J., Ferguson, E., 2017. Development of crash modification factors for uncontrolled pedestrian crossing treatments (No. NCHRP Research Report 841). SAGE Publications Sage CA: Los Angeles, CA.

Zhang, J., Li, Z., Pu, Z., Xu, C., 2018. Comparing prediction performance for crash injury severity among various machine learning and statistical methods. IEEE Access 6, 60079–60087. https://doi.org/10.1109/ACCESS.2018.2874979

Zhao, M., Liu, C., Li, W., Sharma, A., 2018. Multivariate Poisson-Lognormal Model for Analysis of Crashes on Urban Signalized Intersections Approach. Journal of Transportation Safety & Security 10, pp 251-265. https://doi.org/10.1080/19439962.2017.1323059

Zheng, J., Suzuki, K., Fujita, M., 2014. Predicting driver's lane-changing decisions using a neural network model. Simulation Modelling Practice and Theory 42, 73–83. https://doi.org/10.1016/j.simpat.2013.12.007

Zheng, M., Li, T., Zhu, R., Chen, J., Ma, Z., Tang, M., Cui, Z., Wang, Z., 2019. Traffic accident's severity prediction: A deep-learning approach-based CNN network. IEEE Access 7, 39897–39910. https://doi.org/10.1109/ACCESS.2019.2903319

APPENDIX A: HSM SAFETY PERFORMANCE FUNCTIONS

Predictive models can be used to estimate total average crashes (i.e., all crash severities and collision types) or can be used to predict the average frequency of specific crash severity types or specific collision types. The predictive model for an individual roadway segment or intersection combines the SPF, CMFs, and a calibration factor. Chapter 12 of the HSM contains separate predictive models for roadway segments and for intersections. The predictive models for roadway segments estimate the predicted average crash frequency of non-intersection-related crashes. Non-intersection-related crashes may include crashes that occur within the limits of an intersection but are not related to the intersection. The roadway segment predictive models estimate crashes that would occur regardless of the presence of the intersection.

In the predictive method, the appropriate SPFs are used to predict crash frequencies for specific base conditions. SPFs are regression models that estimate the predicted average crash frequency of individual roadway segments or intersections. Each SPF in the predictive method was developed with observed crash data for a set of similar sites. The SPFs, like all regression models, estimate the value of a dependent variable as a function of a set of independent variables. In the SPFs developed for the HSM, the dependent variable estimated is the predicted average crash frequency for a roadway segment or intersection under base conditions, and the independent variables are the AADTs of the roadway segment or intersection legs (and, for roadway segments, the length of the roadway segment). The effect of traffic volume (AADT) on crash frequency is incorporated through the SPF, while the effects of geometric design and traffic control features are incorporated through the CMFs.
SPFs and adjustment factors are provided for five types of roadway segments on urban and suburban arterials:

- Two-lane undivided arterials (2U)
- Three-lane arterials including a center two-way left-turn lane (TWLTL) (3T)
- Four-lane undivided arterials (4U)
- Four-lane divided arterials (i.e., including a raised or depressed median) (4D)
- Five-lane arterials including a center TWLTL (5T)

The SPFs for roadway segments on urban and suburban arterials are applicable to the following AADT ranges:

- 2U: 0 to 32,600 vehicles per day
- 3T: 0 to 32,900 vehicles per day
- 4U: 0 to 40,100 vehicles per day
- 4D: 0 to 66,000 vehicles per day
- 5T: 0 to 53,800 vehicles per day

Application to sites with AADTs substantially outside these ranges may not provide reliable results. Other types of roadway segments may be found on urban and suburban arterials but are not addressed by the predictive model in Chapter 12 of HSM. The procedure addresses five types of collisions.

- Multiple-vehicle non-driveway collisions
- Single-vehicle crashes
- Multiple-vehicle driveway-related collisions
- Vehicle-pedestrian collisions
- Vehicle-bicycle collisions

The effect of traffic volume on predicted crash frequency is incorporated through the SPFs, while the effects of geometric design and traffic control features are incorporated through the CMFs. SPFs are provided for multiple-vehicle non-driveway collisions and single-vehicle crashes. Adjustment factors are provided for multi-vehicle driveway-related, vehicle-pedestrian, and vehicle-bicycle collisions.

Multiple-Vehicle Non-driveway Collisions

The SPF for multiple-vehicle non-driveway collisions is applied as follows:

$$N_{brmv} = exp\ (a + b \times ln(AADT) + ln(L)) \quad (85)$$

Where,

$(AADT)$ = Average annual daily traffic volume (vehicles/day) on roadway segment,
L = Length of roadway segment (mi), and
a, b = Regression coefficients.

Equation (85) is first applied to determine N_{brmv} using the coefficients for total crashes. $N_{brmv(FI)}$ is then divided into components by severity level, for fatal-and-injury crashes and $N_{brmv(PDO)}$, designated as $N'_{brmv(FI)}$ and $N'_{brmv(PDO)}$ in Equation (86), are determined with Equation (85) using the coefficients for fatal-and-injury and property-damage-only crashes, respectively. The following adjustments are then made to assure that $N_{brmv(FI)}$ and $N_{brmv(PDO)}$ sum to N_{brmv}:

$$N_{brmv(FI)} = N_{brmv(total)}\left(\frac{N'_{brmv(FI)}}{N'_{brmv(FI)} + N'_{brmv(PDO)}}\right) \quad (86)$$

$$N_{brmv(PDO)} = N_{brmv(total)} - N_{brmv(FI)} \quad (87)$$

Single-Vehicle Crashes

SPFs for single-vehicle crashes for roadway segments are applied as follows:

$$N_{brsv} = exp\ (a + b \times ln(AADT) + ln(L)) \quad (88)$$

Equation (88) is first applied to determine N_{brsv} using the coefficients for total crashes. N_{brmv} is then divided into components by severity level; $N_{brsv(FI)}$ for fatal-and-injury crashes and $N_{brsv(PDO)}$ for property-damage-only crashes. Preliminary values of $N_{brsv(FI)}$ and $N_{brsv(PDO)}$, designated as $N'_{brsv(FI)}$ and $N'_{brsv(PDO)}$ in Equation (89), are determined with Equation (88) using the coefficients for fatal-and-injury and property-damage-only crashes, respectively. The following adjustments are then made to assure that $N_{brsv(FI)}$ and $N_{brsv(PDO)}$ sum to N_{brsv}:

$$N_{brsv(FI)} = N_{brsv(total)}\left(\frac{N'_{brsv(FI)}}{N'_{brsv(FI)} + N'_{brsv(PDO)}}\right) \quad (89)$$

$$N_{brsv(PDO)} = N_{brsv(total)} - N_{brsv(FI)} \quad (90)$$

Multiple-Vehicle Driveway-Related Collisions

The model presented above for multiple-vehicle collisions addressed only collisions that are not related to driveways. Driveway-related collisions also generally involve multiple vehicles, but are addressed separately because the frequency of driveway-related collisions on a roadway segment depends on the number and type of driveways. Only unsignalized driveways are considered; signalized driveways are analyzed as signalized intersections. The total number of multiple-vehicle driveway-related collisions within a roadway segment is determined as:

$$N_{brdwy} = \sum_{\substack{all \\ driveway \\ types}} n_j \times N_j \times \left(\frac{AADT}{15{,}000}\right)^{(t)} \quad (91)$$

Where,

N_j = Number of driveway-related collisions per driveway per year for driveway type j,

n_j = Number of driveways within roadway segment of driveway type j including all driveways on both sides of the road, and

t = Coefficient for traffic volume adjustment.

The number of driveways of a specific type, n, is the sum of the number of driveways of that type for both sides of the road combined. The number of driveways is determined separately for each side of the road and then added together.

Seven specific driveway types have been considered in modeling. These are:

- Major commercial driveways
- Minor commercial driveways
- Major industrial/institutional driveways
- Minor industrial/institutional driveways
- Major residential driveways
- Minor residential driveways
- Other driveways

Major driveways are those that serve sites with 50 or more parking spaces. Minor driveways are those that serve sites with less than 50 parking spaces. It is not intended that an exact count of the number of parking spaces be made for each site. Driveways can be readily classified as major or minor from a quick review of aerial photographs that show parking areas or through user judgment based on the character of the establishment served by the driveway. Commercial driveways provide access to establishments that serve retail customers. Residential driveways serve single- and multiple-family dwellings. Industrial/institutional driveways serve factories, warehouses, schools, hospitals, churches, offices, public facilities, and other places of employment. Commercial sites with no restriction on access along an entire property frontage are generally counted as two driveways.

Driveway-related collisions can be separated into components by severity level as follows:

$$N_{brdwv(FI)} = N_{brdwv(total)} \times f_{dwv} \quad (92)$$

$$N_{brdwv(PDO)} = N_{brdwv(total)} \times N_{brdwv(FI)} \quad (93)$$

Where,

f_{dwv} = Proportion of driveway-related collisions that involve fatalities or injuries,

Vehicle-Pedestrian Collisions

The number of vehicle-pedestrian collisions per year for a roadway segment is estimated as:

$$N_{pedr} = N_{br} \times f_{pedr} \quad (94)$$

Where,

f_{pedr} = Pedestrian crash adjustment factor.

Vehicle-Bicycle Collisions

The number of vehicle-bicycle collisions per year for a roadway segment is estimated as:

$$N_{biker} = N_{br} \times f_{biker} \quad (95)$$

Where,

f_{biker} = Bicycle crash adjustment factor.

APPENDIX B: DATA DICTIONARY

Column Name	Details	Source
UnqID	Unique ID	TXST generated
GID	RDBD_GMTRY_LN_ID	RHiNO
Frm_Dfo	FROM-DFO	RHiNO
To_Dfo	TO-DFO	RHiNO
C_Sec	CONTROL-SECTION	RHiNO
Hwy	SIGNED-HIGHWAY	RHiNO
Ste_Nam	STREET-NAME	RHiNO
District_Code	DISTRICT-ID	RHiNO
District	DISTRICT-NAME	RHiNO
County_Code	COUNTY-NUMBER	RHiNO
County	COUNTY-NAME	RHiNO
City	CITY-NUMBER	RHiNO
Ru	RURAL-URBAN-CODE	RHiNO
F_Syste	FUNCTIONAL-CLASSIFICATION	RHiNO
Ru_F_Sy	FUNCTIONAL-CLASSIFICATION	RHiNO
Spd_Max	SPEED-LIMIT-MAXIMUM	RHiNO
Med_Typ	MEDIAN-TYPE	RHiNO
Med_Wid	MEDIAN-WIDTH	RHiNO
Num_Lan	NUMBER-OF-THROUGH-LANES	RHiNO
Hov_Lan	HOV-LANES	RHiNO
Hov_Typ	HOV-TYPE	RHiNO
Rb_Wid	ROADBED-WIDTH	RHiNO
Sur_W	SURFACE-WIDTH	RHiNO
S_Type_I	SHOULDER-TYPE-INSIDE	RHiNO
S_Wid_I	SHOULDER-WIDTH-INSIDE	RHiNO
S_Use_I	SHOULDER-USE-INSIDE	RHiNO
S_Type_O	SHOULDER-TYPE-OUTSIDE	RHiNO
S_Wid_O	SHOULDER-WIDTH-OUTSIDE	RHiNO
S_Use_O	SHOULDER-USE-OUTSIDE	RHiNO
Curb_L	CURB-TYPE-LEFT	RHiNO
Srf_Typ	SURFACE-TYPE	RHiNO
Adt_Yea	YEAR-OF-ANNUAL-AVERAGE-DAILYTRAFFIC	RHiNO
Adt_Cur	AADT-CURRENT	RHiNO
Adt_Adj	AADT-ADJUST-CURRENT	RHiNO
K_Fac	PEAK-FACTOR	RHiNO
D_Fac	DIRECTIONAL-DISTRIBUTION-FACTOR	RHiNO
Trk_Aad	TRUCK-AADT-PCT	RHiNO
Hy_1	ADT-HISTORY-YEAR-1	RHiNO
Hy_2	ADT-HISTORY-YEAR-2	RHiNO
Hy_3	ADT-HISTORY-YEAR-3	RHiNO
Hy_4	ADT-HISTORY-YEAR-4	RHiNO
Hy_5	ADT-HISTORY-YEAR-5	RHiNO
Hy_6	ADT-HISTORY-YEAR-6	RHiNO
Hy_7	ADT-HISTORY-YEAR-7	RHiNO
Hy_8	ADT-HISTORY-YEAR-8	RHiNO
Hy_9	ADT-HISTORY-YEAR-9	RHiNO

Desgn_Y	DESIGN-YEAR	RHiNO
Len_Sec	LENGTH-OF-SECTION	RHiNO
Ln_Mile	LANE-MILES	RHiNO
Dvmt	DAILY-VEHICLE-MILES-OF-TRAVEL	RHiNO
Fclty_C	Facility type of each RHiNO segement	RHiNO
Spd18	Average speed in 2018 (mph)	INRIX XD
Ss18	Speed standard deviation in 2018 (mph)	INRIX XD
Spd19	Average speed in 2019 (mph)	INRIX XD
Ss19	Speed standard deviation in 2019 (mph)	INRIX XD
Spd20	Average speed in 2020 (mph)	INRIX XD
Ss20	Speed standard deviation in 2020 (mph)	INRIX XD
Spd21	Average speed in 2021 (mph)	INRIX XD
Ss21	Speed standard deviation in 2021 (mph)	INRIX XD
Spd22	Average speed in 2022(mph)	INRIX XD
Ss22	Speed standard deviation in 2022 (mph)	INRIX XD
Sef22	85 percentile speed in 2022 (mph)	INRIX XD
PrcSum19	Sum of precipitation in 2019 (in)	NOAA (CDS)
PrcAvg19	Average of precipitation in 2019 (in)	NOAA (CDS)
PrcSum20	Sum of precipitation in 2020 (in)	NOAA (CDS)
PrcAvg20	Average of precipitation in 2020 (in)	NOAA (CDS)
PrcSum21	Sum of precipitation in 2021 (in)	NOAA (CDS)
PrcAvg21	Average of precipitation in 2021 (in)	NOAA (CDS)
PrcSum22	Sum of precipitation in 2022 (in)	NOAA (CDS)
PrcAvg22	Average of precipitation in 2022 (in)	NOAA (CDS)
A_18	Number of Incapacitating injury crashes in 2018	CRIS
B_18	Number of non-incapacitating injury crashes in 2018	CRIS
C_18	Number of possible injury crashes in 2018	CRIS
K_18	Number of fatal crashes in 2018	CRIS
O_18	Number of not injured crashes in 2018	CRIS
U_18	Number of unknown crashes in 2018	CRIS
A_19	Number of Incapacitating injury crashes in 2019	CRIS
B_19	Number of non-incapacitating injury crashes in 2019	CRIS
C_19	Number of possible injury crashes in 2019	CRIS
K_19	Number of fatal crashes in 2019	CRIS
O_19	Number of not injured crashes in 2019	CRIS
U_19	Number of unknown crashes in 2019	CRIS
A_20	Number of Incapacitating injury crashes in 2020	CRIS
B_20	Number of non-incapacitating injury crashes in 2020	CRIS
C_20	Number of possible injury crashes in 2020	CRIS
K_20	Number of fatal crashes in 2020	CRIS
O_20	Number of not injured crashes in 2020	CRIS
U_20	Number of unknown crashes in 2020	CRIS
A_21	Number of Incapacitating injury crashes in 2021	CRIS
B_21	Number of non-incapacitating injury crashes in 2021	CRIS
C_21	Number of possible injury crashes in 2021	CRIS
K_21	Number of fatal crashes in 2021	CRIS
O_21	Number of not injured crashes in 2021	CRIS
U_21	Number of unknown crashes in 2021	CRIS
A_22	Number of Incapacitating injury crashes in 2022	CRIS

B_22	Number of non-incapacitating injury crashes in 2022	CRIS
C_22	Number of possible injury crashes in 2022	CRIS
K_22	Number of fatal crashes in 2022	CRIS
O_22	Number of not injured crashes in 2022	CRIS
U_22	Number of unknown crashes in 2022	CRIS
MnrCmmr	Count of Minor Commercial Driveway	TXST generated
MjrCmmr	Count of Major Commercial Driveway	TXST generated
MnrInds	Count of Minor Industrial Driveway	TXST generated
MnrRsdn	Count of Minor Residential Driveway	TXST generated
MjrInds	Count of Major Industrial Driveway	TXST generated
Other	Count of Other Driveway	TXST generated
MjrRsdn	Count of Major Residential Driveway	TXST generated
Fclty_C1	Facility Codes	TXST generated
Facility	Facility Names	TXST generated
Cr18_22_FI	Count of Fatal and Injury Crashes (2018 - 2022)	TXST generated
Cr18_22_PDO	Count of PDO Crashes (2018 - 2022)	TXST generated
TotalCrash	Count of Total Crashes (2018 - 2022)	TXST generated
Pav_S_Wid	Paved Shoulder Width (ft)	TXST generated
Sw*	Weight Factor	TXST generated
Equi_Dwy	Count of Equivalent Driveway	TXST generated
Equi_Dwy_M*	Weight Factor	TXST generated
Spd*	Weight Factor	TXST generated
Prc_Avg*	Weight Factor	TXST generated
Site_No*	Weight Factor	TXST generated
Predicted_FI	SPF-based Predicted Fatal and Injury Crashes	TXST generated
Estimated_FI	EB based Estimated Fatal and Injury Crashes	TXST generated
Predicted_PDO	SPF based Predicted PDO Crashes	TXST generated
Estimated_PDO	EB based Estimated PDO Crashes	TXST generated
Estimated_Total	EB based Estimated Total Crashes	TXST generated
UnqID	Unique Segment ID2	TXST generated

***** *indicates weight factor used for normalization of variable measures.*

APPENDIX C: DECISION SUPPORT TOOL SCRIPTS

```
# Developed by Subasish Das and David Mills
# With Data Upload Option.
# 1. After selecting District and County, user has option to upload
custom data.
# 2. App checks for required Columns for CSV file upload to make sure
we can match data to shapefile.
#    County - ADT, ControlS - ToDFO, MedType, NumLane, SpdAve <- Look
at yellow columns in 7144.
# 3. App checks # of rows for district/county selection to make sure
data matches.
# 4. App provides a notice to user if the selection # is different.
# 5. App processes function to calculate Expected Crashes based on
user uploaded data.
# 6. Busy icon spins while processing data.
# 7. Hot spot map loads for user to view and data can be downloaded.
# 8. Output only shows FICrE, PDOCrE, TotalCrE.

library(shiny)
library(shinydashboard)
library(shinyjs)
library(sf)
library(leaflet)
library(leaflet.extras)
library(dplyr)
library(DT)
library(htmltools)
library(shinybusy)
library(data.table)
library(openxlsx)

options(shiny.maxRequestSize = 105*1024^2)  # Sets the limit to 105MB

# 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")])
)
```

```
# Load custom shapefile and CSV
shapefile <- st_transform(st_read("0_data/tmc_shapefile.shp"), 4326)
demo_data <- read.csv("0_data/demo_data.csv")

# Save original column names for restoration
original_demo_colnames <- names(demo_data)

# Merge without unnecessary suffixes
TX_shp <- shapefile %>%
  left_join(demo_data, by = "un_d_nw")

# Join displaydata with StateCountyData using district and county IDs
TX_shp <- left_join(TX_shp, StateCountyData, by = c("DI" =
"DistrictID", "CO" = "CountyID"))

# Function to update crashes by adding new values
update_crashes <- function(existing_data, new_data) {
  tryCatch({
    # Ensure both datasets are data.tables
    existing_data <- as.data.table(existing_data)
    new_data <- as.data.table(new_data)

    print(head(new_data))

    # Required columns for crash data
    required_columns <- c("GID", "FRM_DFO", "TO_DFO", "C_SEC", "FI",
"PDO")

    # Check for required columns
    if (!all(required_columns %in% names(new_data))) {
      stop("New data is missing required columns.")
    }

    # Replace NA values with 0 for new crash data
    new_data[, `:=`(
      FI = fifelse(is.na(FI), 0, FI),
      PDO = fifelse(is.na(PDO), 0, PDO)
    )]

    print("This Worked.")

    # Merge with existing data, ensuring FI maps to EB_fi and PDO maps
to EB_pdo
    updated_data <- merge(existing_data, new_data, by = c("GID",
"FRM_DFO", "TO_DFO", "C_SEC"), all.x = TRUE, suffixes = c("", ".new"))

    # Update EB_fi and EB_pdo
    updated_data[, `:=`(
      EB_fi = as.integer(fifelse(is.na(FI), EB_fi, (EB_fi + FI) * 0.5
+ pred_fi * 0.5)),
      EB_pdo = as.integer(fifelse(is.na(PDO), EB_pdo, (EB_pdo + PDO) *
0.5 + pred_pdo * 0.5)),
```

```
      EB_total = EB_fi + EB_pdo  # Ensure EB_total is recalculated
properly
    )]

    # Remove unnecessary columns
    updated_data[, c("FI", "PDO") := NULL]

    return(updated_data)
  }, error = function(e) {
    print("Error updating crashes:")
    print(e$message)
    stop(e)
  })
}

update_demo_data_and_shapefile <- function(uploaded_file) {
  if (!is.null(uploaded_file)) {
    uploaded_data <- fread(uploaded_file$datapath)

    # Ensure required columns exist
    required_columns <- c("GID", "FRM_DFO", "TO_DFO", "C_SEC", "FI",
"PDO")
    if (!all(required_columns %in% names(uploaded_data))) {
      showNotification("Uploaded data is missing required columns.",
type = "error")
      return()
    }

    # Create a copy of the original data before modifying
    modified_demo_data <- demo_data

    # Update the copied data instead of overwriting the original
    modified_demo_data <- update_crashes(modified_demo_data,
uploaded_data)

    # Merge updated demo_data back into the shapefile
    modified_TX_shp <- shapefile %>%
      left_join(as.data.frame(modified_demo_data), by = "un_d_nw") %>%
      left_join(StateCountyData, by = c("DI" = "DistrictID", "CO" =
"CountyID"))

    # Assign modified data
    demo_data <<- modified_demo_data
    TX_shp <<- modified_TX_shp

    return(TX_shp)
  }
}

# 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')))

# UI Code
body <- dashboardBody(
  useShinyjs(),
  add_busy_spinner("fulfilling-bouncing-circle"),
  tabsetPanel(
    tabPanel(HTML(paste(tags$span(style="font-size: 18px",
"Introduction"))),
             tags$br(),
             tags$h2(tags$b("Texas Interactive Decision Support Tool
to Improve Safety for Urban Roads with Speed Data")),
             h2(),
             div(style = "font-size: 18px;", HTML("This project aims
to provide an interactive tool to identify crash hotspots on Texas
urban state-maintained roadways. Six different roadway facility types
(Urban Two-lane, Urban Two-lane with a TWLTL, Urban Interstate, Urban
Multilane Divided, Urban Multilane Undivided, Urban Multilane with a
TWLTL) are included in this tool. The tool can provide 5 years
observed number of crashes and 5 years expected number of crashes for
each roadway segment. Moreover, geometric and speed distribution
information are also available in this tool to help support safety
decisions.")),
             h2(),
             div(style = "font-size: 18px;", HTML("The current tool
has one tab. Users need to follow some steps to make the tool
interactive:", "<br>")),
             tags$span(style = "font-size: 18px", tags$ul(
               tags$li("Select District"),
               tags$li("Select County"),
               tags$li("Select Facility Type"),
               tags$li("Select AADT Range"),
               tags$li("Select Crash Severity Level"),
               tags$li("Click 'Refresh Map' (will take some time to
load the map)"),
               tags$li("Detailed Data can be downloaded by clicking
the download button"),
```

```
              tags$li("Detailed definitions of the variable names can
be downloaded after refreshing the map")
            )),
            tags$h3(tags$b("Acknowledgments")),
            h2(),
            div(style = "font-size: 18px;", HTML("The project was
funded by Texas Department of Transportation (TXDOT).")),
            h2(),
            div(style = "font-size: 18px;", HTML("The project was
conducted by Texas State University (TXST) and Texas A&M
Transportation Institute (TTI). The interactive online tool was
developed by the TXST project team members Dr. Subasish Das and Mr.
David Mills. Questions about the tool can be sent to the Principal
Investigator, Dr. Subasish Das at <a
href='mailto:subasish@txstate.edu'>subasish@txstate.edu</a>")),
            hr(),
            tags$img(src='txdot_Logo.png', height=120), HTML(" 
 "),
            tags$img(src='txstLogo.png', height=80),
            tags$img(src='TTI_Logo.png', height=80),
            hr(),
            div(style = "font-size: 18px;", HTML("Last updated:
January 28, 2025."))
    ),
    tabPanel(HTML(paste(tags$span(style="font-size: 18px", "0-7144
Tool"))), id="UrbanSpeedTool",
            tags$h1(tags$b("Texas Interactive Decision Support Tool
to Improve Safety for Urban Roads with Speed Data (0-7144 Tool)")),
            fluidRow(
              column(width = 8,
                     box(width = NULL, solidHeader = TRUE,
                         div(
                           id = "noDataNotice",
                           style = "position: absolute; top: 50%;
left: 50%; transform: translate(-50%, -50%);
                                    text-align: center; background-
color: rgba(255, 255, 255, 0.8);
                                    padding: 20px; border: 2px solid
#d9534f; border-radius: 8px;
                                    font-size: 18px; color: #d9534f;
z-index: 1000; display: none;",
                                   "No applicable state maintained
roadways."
                         ),
                         leafletOutput("MapOut", height = 500),
                         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"),

                          # Conditional file input except when All
Districts is selected.
                          conditionalPanel(
                            condition = "input.DistrictInput !== 'All
Districts' && (input.CountyInput === 'All Counties' ||
input.CountyInput !== 'All Counties')",
                            fileInput("user_file", "Upload Your CSV",
accept = ".csv")
                          ),

                          selectInput("FacilityInput", "Facility",
                                      choices = c("All Facilities",
                                                  "Urban Two-lane",
                                                  "Urban Two-lane with
a TWLTL",
                                                  "Urban Interstate
Highway",
                                                  "Urban Multilane
Divided",
                                                  "Urban Multilane
Undivided",
                                                  "Urban Multilane
with a TWLTL"),
                                      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),
                          actionButton("resetData", "Reset to Original
Data", class = "butt"),
                          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("1. Detailed definitions of the
variable names can be downloaded"),
                          downloadLink("downloadDefination",label
="here"),
```

```
                          HTML(". (Please refresh map first before
downloading)"),
                          h2(),
                          HTML("2. Please use the arrows beside column
names to rank top sites")
                        )
                )
              ),
              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) {
  # Store original copies of data
  original_demo_data <<- demo_data
  original_TX_shp <<- TX_shp

  # Watch for file upload event and call update function
  observeEvent(input$user_file, {
    tryCatch({
      uploaded_data <- read.csv(input$user_file$datapath)

      required_columns <- c("di", "co", "frm_dfo", "to_dfo", "c_sec")

      colnames(uploaded_data) <- tolower(colnames(uploaded_data))
      colnames(demo_data) <- tolower(colnames(demo_data))

      missing_columns <- setdiff(required_columns,
colnames(uploaded_data))
      if (length(missing_columns) > 0) {
        showNotification(paste("Missing columns:",
paste(missing_columns, collapse = ", ")), type = "error")
        return()
      }

      if (input$CountyInput == "All Counties") {
        # When 'All Counties' is selected, filter by District
        district_id <- unique(filter(StateCountyData, District ==
input$DistrictInput)$DistrictID)[1]
        uploaded_data <- uploaded_data %>% filter(di == district_id)
      }
```

```
      TX_shp <<- update_demo_data_and_shapefile(input$user_file)

      showNotification("Data successfully uploaded and processed.",
type = "message")

    }, error = function(e) {
      showNotification("Error processing the uploaded file.", type =
"error")
    })
  })

  observeEvent(input$resetData, {
    # Restore original data
    demo_data <<- original_demo_data
    TX_shp <<- original_TX_shp

    # Clear the file input field using JavaScript
    runjs("document.getElementById('user_file').value = '';")

    # Refresh the map automatically
    shinyjs::click("RefreshMap")

    showNotification("Data reset to original state. Map refreshed.",
type = "message")
  })


  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, {
```

```
    # Get District Code based on District name input
    Districtin <- input$DistrictInput
    if (Districtin != "All Districts") {
      # Retrieve the District ID corresponding to the selected
District name
      Districtin_Code <- unique(filter(StateCountyData, District ==
Districtin)$DistrictID)[1]
    } else {
      Districtin_Code <- 0
    }

    # Get County Code based on County name input
    COUNTYin <- input$CountyInput
    if (COUNTYin != "All Counties") {
      # Retrieve the County ID corresponding to the selected County
name and District ID
      Countyin_Code <- filter(StateCountyData, DistrictID ==
Districtin_Code, County == COUNTYin)$CountyID[1]
    } else {
      Countyin_Code <- 0
    }

    # Map input Facility selection to the updated facility codes
    Facilityin_Code <- switch(input$FacilityInput,
                              "All Facilities" = 0,
                              "Urban Two-lane" = "2U",
                              "Urban Two-lane with a TWLTL" = "2T",
                              "Urban Interstate Highway" = "IH",
                              "Urban Multilane Divided" = "MD",
                              "Urban Multilane Undivided" = "MU",
                              "Urban Multilane with a TWLTL" = "MT"
    )

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

    MapOutputData <- TX_shp

    # Filter MapOutputData based on Districtin_Code using `di` column
in MapOutputData
    if (Districtin_Code == 0) {
      MapOutputDataTempDistrict <- MapOutputData
      TXdistricts_shp_selected <-
st_as_sf(as.data.frame(TXDistricts_shp))
    } else {
      # Filter for the selected district code in both MapOutputData
and TXDistricts_shp
```

```
      MapOutputDataTempDistrict <-
st_as_sf(filter(as.data.frame(MapOutputData), DI == Districtin_Code))
      TXdistricts_shp_selected <-
st_as_sf(filter(as.data.frame(TXDistricts_shp), DIST_NBR ==
Districtin_Code))
    }

    # Filter MapOutputDataTempDistrict based on Countyin_Code using
`co` column in MapOutputData
    if (Countyin_Code == 0) {
      MapOutputDataTempCounty <- MapOutputDataTempDistrict

      # Further filter TXcounties_shp based on Districtin_Code if
District is specified
      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 {
      # Filter MapOutputDataTempDistrict for the selected county code
in `co` column
      MapOutputDataTempCounty <-
st_as_sf(filter(as.data.frame(MapOutputDataTempDistrict), CO ==
Countyin_Code))
      TXcounties_shp_selected <-
st_as_sf(filter(as.data.frame(TXcounties_shp), CNTY_NBR ==
Countyin_Code))

      # Get coordinates for centering map based on the selected county
      corr <-
as.data.frame(st_coordinates(st_centroid(TXcounties_shp_selected)))
      LATzoom <- corr$Y
      LONzoom <- corr$X
      zoomLevel <- 9
    }

    # Filter MapOutputDataTempCounty based on Facilityin_Code using
the `facility` column
```

```
    if (Facilityin_Code == 0) {
      MapOutputDataTempFacility <- MapOutputDataTempCounty
    } else {
      # Filter for the selected facility code in `facility` column
      MapOutputDataTempFacility <-
st_as_sf(filter(as.data.frame(MapOutputDataTempCounty), facility ==
Facilityin_Code))
    }


    # Filter MapOutputDataTempFacility based on AADTin_Code using the
`ADT_CUR` column
    if (AADTin_Code == 0) {
      MapOutputDataTempAADT <- MapOutputDataTempFacility
    } else if (AADTin_Code == 1) {
      # Filter for AADT <= 2000
      MapOutputDataTempAADT <-
st_as_sf(filter(as.data.frame(MapOutputDataTempFacility), ADT_CUR <=
2000))
    } else if (AADTin_Code == 2) {
      # Filter for 2000 < AADT <= 10000
      MapOutputDataTempAADT <-
st_as_sf(filter(as.data.frame(MapOutputDataTempFacility), ADT_CUR >
2000 & ADT_CUR <= 10000))
    } else {
      # Filter for AADT > 10000
      MapOutputDataTempAADT <-
st_as_sf(filter(as.data.frame(MapOutputDataTempFacility), ADT_CUR >
10000))
    }

    # Final assignment for MapOutputDataFinal after all filtering
steps
    MapOutputDataFinal <- MapOutputDataTempAADT

    # Convert MapOutputDataFinal to a data frame
    displaydata <- as.data.frame(MapOutputDataFinal)

    # Observe the displaydata row count and toggle the noDataNotice
div visibility
    observe({
      if (nrow(displaydata) == 0) {
        shinyjs::show("noDataNotice")  # Show the notification if no
data
      } else {
        shinyjs::hide("noDataNotice")  # Hide the notification if data
is available
      }
    })

    # Map facility codes to descriptive names in the `facility` column
```

```
    displaydata <- within(displaydata, facility[facility == '2U'] <-
'Urban Two-lane')
    displaydata <- within(displaydata, facility[facility == '2T'] <-
'Urban Two-lane with a TWLTL')
    displaydata <- within(displaydata, facility[facility == 'IH'] <-
'Urban Interstate Highway')
    displaydata <- within(displaydata, facility[facility == 'MD'] <-
'Urban Multilane Divided')
    displaydata <- within(displaydata, facility[facility == 'MU'] <-
'Urban Multilane Undivided')
    displaydata <- within(displaydata, facility[facility == 'MT'] <-
'Urban Multilane with a TWLTL')

    # Set a minimum value of 0.01 for EB_total, EB_pdo, and EB_fi if
they are less than 0.01
    displaydata <- within(displaydata, EB_total[EB_total < 0.01] <-
0.01)
    displaydata <- within(displaydata, EB_pdo[EB_pdo < 0.01] <- 0.01)
    displaydata <- within(displaydata, EB_fi[EB_fi < 0.01] <- 0.01)

    #print(head(displaydata))
    displaydata$EB_total <- as.integer(displaydata$EB_fi +
displaydata$EB_pdo)

    # Ensure column names in displaydata are clear and consistent
    displaydata <- displaydata %>%
      rename(
        District_Code = DI,    # District code from `di`
        County_Code = CO,      # County code from `co`
        District = District,   # Descriptive district name
        County = County        # Descriptive county name
      )

    # Select the appropriate column in MapOutputDataFinal based on the
selected severity
    DataForPal <- switch(input$Severity,
                         "Total" = MapOutputDataFinal$EB_total,
                         "Fatal and Injury" =
MapOutputDataFinal$EB_fi,
                         "No Injury" = MapOutputDataFinal$EB_pdo
    )

    pal_Total <- colorNumeric("YlOrRd", DataForPal)

    # Define a named vector mapping each column to its display name
    column_names <- c(
      "un_d_nw" = "Segment ID",
      "HWY" = "Highway Number",
      "District" = "District",
      "County" = "County",
      "MED_WID" = "Median Width",
      "NUM_LAN" = "Number of Lanes",
```

```
      "SUR_W" = "Surface Width",
      "DHV" = "Daily Hourly Volume",
      "LEN_SEC" = "Segment Length",
      "MnrCmmr" = "Minor Commercial Driveway",
      "MjrCmmr" = "Major Commercial Driveway",
      "MnrInds" = "Minor Industrial Driveway",
      "MnrRsdn" = "Minor Residential Driveway",
      "MjrInds" = "Major Industrial Driveway",
      "MjrRsdn" = "Major Residential Driveway",
      "Other" = "Other Driveway",
      "facility" = "Facility Type",
      "equi_dwy" = "Equivalent Driveway",
      "EB_fi" = "Expected FI Crashes",
      "EB_pdo" = "Expected PDO Crashes",
      "EB_total" = "Expected Total Crashes"
    )

    # Generate labelOut dynamically based on columns present in
displaydata
    labelOut <- lapply(names(column_names), function(col) {
      if (col %in% names(displaydata)) {
        paste0(column_names[col], ": ", displaydata[[col]], "<br>")
      } else {
        NULL
      }
    })

    # Combine all elements of labelOut into a single list with HTML
    labelOut <- as.list(do.call(paste0, labelOut))

    # Create popup content with expected and observed crashes
information
    popupOut <- paste0(
      "<div style='max-height: 300px; overflow-y: auto;'>",  # Start
scrollable div
      'Segment ID: ', displaydata$un_d_nw, "<br>",
      'Highway Number: ', displaydata$HWY, "<br>",
      'District: ', displaydata$District, "<br>",
      'County: ', displaydata$County, "<br>",
      'Median Width: ', displaydata$MED_WID, "<br>",
      'Number of Lanes: ', displaydata$NUM_LAN, "<br>",
      'Surface Width: ', displaydata$SUR_W, "<br>",
      'Daily Hourly Volume: ', displaydata$DHV, "<br>",
      'Segment Length: ', format(round(displaydata$LEN_SEC, 3), nsmall
= 3), "<br>",
      'Minor Commercial Driveway: ', displaydata$MnrCmmr, "<br>",
      'Major Commercial Driveway: ', displaydata$MjrCmmr, "<br>",
      'Minor Industrial Driveway: ', displaydata$MnrInds, "<br>",
      'Minor Residential Driveway: ', displaydata$MnrRsdn, "<br>",
      'Major Industrial Driveway: ', displaydata$MjrInds, "<br>",
      'Major Residential Driveway: ', displaydata$MjrRsdn, "<br>",
      'Other Driveway: ', displaydata$Other, "<br>",
```

```
      'Facility Type: ', displaydata$facility, "<br>",
      'Equivalent Driveway: ', displaydata$equi_dwy, "<br>",
      'Expected FI Crashes: ', format(round(displaydata$EB_fi, 2),
nsmall = 2), "<br>",
      'Expected PDO Crashes: ', format(round(displaydata$EB_pdo, 2),
nsmall = 2), "<br>",
      'Expected Total Crashes: ', format(round(displaydata$EB_total,
2), nsmall = 2), "<br>",
      "</div>"  # End scrollable div
    )

    leafletProxy("MapOut") %>%
      clearPopups() %>%
      clearGroup("Total/Fata/Injury") %>%
      clearGroup("CountiesSHP") %>%
      clearGroup("DistrictsSHP") %>%
      clearControls() %>%
      setView(lng = LONzoom, lat = LATzoom, zoom = zoomLevel) %>%

      # Add polylines for MapOutputDataFinal with updated columns for
crash data
      addPolylines(
        data = MapOutputDataFinal,
        color = ~pal_Total(
          switch(input$Severity,
                 "Total" = MapOutputDataFinal$EB_total,
                 "Fatal and Injury" = MapOutputDataFinal$EB_fi,
                 "No Injury" = MapOutputDataFinal$EB_pdo)
        ),
        group = "Total/Fata/Injury",
        popup = popupOut,
        label = lapply(labelOut, HTML)
      ) %>%

      # Add county boundaries
      addPolylines(
        data = TXcounties_shp_selected,
        color = '#81A88D',
        group = "CountiesSHP",
        weight = 1
      ) %>%

      # Add district boundaries
      addPolylines(
        data = TXdistricts_shp_selected,
        color = '#C93312',
        group = "DistrictsSHP",
        weight = 2.5
      ) %>%

      # Add legend for crash severity based on updated columns
      addLegend(
```

```
        "bottomright",
        pal = pal_Total,
        values = switch(input$Severity,
                        "Total" = MapOutputDataFinal$EB_total,
                        "Fatal and Injury" = MapOutputDataFinal$EB_fi,
                        "No Injury" = MapOutputDataFinal$EB_pdo),
        title = paste0(input$Severity, " Crashes")
      )

    if (nrow(displaydata) > 0) {
      MapOutputDataFinalDTtemp <- cbind(
        select(displaydata,
               c(
                 'un_d_nw',          # Segment ID
                 'District',         # District
                 'County',           # County
                 'HWY',              # Highway Number
                 'facility',         # Facility Type
                 'LEN_SEC',          # Segment Length
                 'ADT_CUR',          # AADT
                 'Spd18',            # Average Operating Speed
                 'SS18',             # Standard Deviation of Operating
Speed
                 'PSL18',            # PSL Equivalent from INRIX
                 'pre_avg',          # Average Precipitation
                 'pred_fi',          # Predicted FI Crashes
                 'pred_pdo',         # Predicted PDO Crashes
                 'EB_total',         # Expected Total Crashes
                 'EB_pdo',           # Expected PDO Crashes
                 'EB_fi',            # Expected FI Crashes
                 'total'             # Observed Total Crashes
                 # 'Crash_18_22_pdo',  # Observed PDO Crashes
                 # 'Crash_18_22_fi'    # Observed FI Crashes
               )
        )
      )

      # Remove duplicates based on the Segment ID (un_d_nw) and keep
all other columns
      MapOutputDataFinalDTtemp <- distinct(MapOutputDataFinalDTtemp,
un_d_nw, .keep_all = TRUE)

      # Format specific columns to ensure consistent numeric
formatting
      MapOutputDataFinalDTtemp$spd18 <-
format(round(MapOutputDataFinalDTtemp$Spd18, 2), nsmall = 2)  #
Average Operating Speed
      MapOutputDataFinalDTtemp$ss18 <-
format(round(MapOutputDataFinalDTtemp$SS18, 2), nsmall = 2)    #
Standard Deviation of Operating Speed
```

```
      MapOutputDataFinalDTtemp$pre_avg <-
format(round(MapOutputDataFinalDTtemp$pre_avg, 2), nsmall = 2)  #
Average Precipitation
      MapOutputDataFinalDTtemp$EB_total <-
format(round(MapOutputDataFinalDTtemp$EB_total, 2), nsmall = 2) #
Expected Total Crashes
      MapOutputDataFinalDTtemp$EB_pdo <-
format(round(MapOutputDataFinalDTtemp$EB_pdo, 2), nsmall = 2)     #
Expected PDO Crashes
      MapOutputDataFinalDTtemp$EB_fi <-
format(round(MapOutputDataFinalDTtemp$EB_fi, 2), nsmall = 2)       #
Expected FI Crashes
      MapOutputDataFinalDTtemp$total <-
format(round(MapOutputDataFinalDTtemp$total, 2), nsmall = 2)       #
Observed Total Crashes
      # MapOutputDataFinalDTtemp$Crash_18_22_pdo <-
format(round(MapOutputDataFinalDTtemp$Crash_18_22_pdo, 2), nsmall = 2)
# Observed PDO Crashes
      # MapOutputDataFinalDTtemp$Crash_18_22_fi <-
format(round(MapOutputDataFinalDTtemp$Crash_18_22_fi, 2), nsmall = 2)
# Observed FI Crashes
      MapOutputDataFinalDTtemp$len_sec <-
format(round(MapOutputDataFinalDTtemp$LEN_SEC, 3), nsmall = 3)  #
Segment Length

    } else {
      # Define the column names expected in MapOutputDataFinalDTtemp
      column_names <- c(
        'un_d_nw',          # Segment ID
        'District',         # District
        'County',           # County
        'HWY',              # Highway Number
        'facility',         # Facility Type
        'LEN_SEC',          # Segment Length
        'ADT_CUR',          # AADT
        'Spd18',            # Average Operating Speed
        'SS18',             # Standard Deviation of Operating Speed
        'PSL18',            # PSL Equivalent from INRIX
        'pre_avg',          # Average Precipitation
        'pred_fi',          # Predicted FI Crashes
        'pred_pdo',         # Predicted PDO Crashes
        'EB_total',         # Expected Total Crashes
        'EB_pdo',           # Expected PDO Crashes
        'EB_fi',            # Expected FI Crashes
        'total'             # Observed Total Crashes
        # 'Crash_18_22_pdo',  # Observed PDO Crashes
        # 'Crash_18_22_fi'    # Observed FI Crashes
      )

      # Create an empty data frame with "None" for each column
      MapOutputDataFinalDTtemp <- data.frame(matrix("None", nrow = 1,
ncol = length(column_names)))
```

```
      names(MapOutputDataFinalDTtemp) <- column_names

    }

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

    # Render the data table with the expected and observed crashes
columns and enable single-row selection
    output$outputDT <- DT::renderDT({
      displaydata_selected <- displaydata %>%
        select(
          un_d_nw,           # Segment ID
          District,          # District
          County,            # County
          FRM_DFO,           # From DFO
          TO_DFO,            # To DFO
          C_SEC,             # Control Section
          total,             # Observed Total Crashes
          EB_fi,             # Expected FI Crashes
          EB_pdo,            # Expected PDO Crashes
          EB_total           # Expected Total Crashes
        ) %>%
        rename(
          Segment_ID = un_d_nw,
          From_DFO = FRM_DFO,
          To_DFO = TO_DFO,
          Control_Section = C_SEC,
          Observed_Total_Crashes = total,
          Expected_FI_Crashes = EB_fi,
          Expected_PDO_Crashes = EB_pdo,
          Expected_Total_Crashes = EB_total
        )

      datatable(
        displaydata_selected,
        class = 'cell-border stripe',
        rownames = FALSE,
        options = list(lengthChange = FALSE),
        selection = 'single'  # Enable single-row selection
      )
    })

    # Observe the selected row in the data table and show the popup
    observeEvent(input$outputDT_rows_selected, {
      selected_row <- input$outputDT_rows_selected

      if (length(selected_row) > 0) {
        # Get the Segment ID of the selected row
        selected_id <- displaydata[selected_row, "un_d_nw"]
```

```
        # Filter TX_shp to get the corresponding feature
        selected_feature <- TX_shp %>% filter(un_d_nw == selected_id)

        # Get the coordinates for centering the map
        selected_coords <- st_centroid(st_geometry(selected_feature))
%>% st_coordinates()
        lng <- selected_coords[1]
        lat <- selected_coords[2]

        # Create popup content with expected and observed crashes
information
        popupOut <- paste0(
          "<div style='max-height: 300px; overflow-y: auto;'>",  #
Start scrollable div
          'Segment ID: ', displaydata$un_d_nw, "<br>",
          'Highway Number: ', displaydata$HWY, "<br>",
          'District: ', displaydata$District, "<br>",
          'County: ', displaydata$County, "<br>",
          'Median Width: ', displaydata$MED_WID, "<br>",
          'Number of Lanes: ', displaydata$NUM_LAN, "<br>",
          'Surface Width: ', displaydata$SUR_W, "<br>",
          'AADT: ', displaydata$ADT_CUR, "<br>",
          'K-Factor: ', displaydata$K_FAC, "<br>",
          'D-Factor: ', displaydata$D_FAC, "<br>",
          'Daily Hourly Volume: ', displaydata$DHV, "<br>",
          'Segment Length: ', format(round(displaydata$LEN_SEC, 3),
nsmall = 3), "<br>",
          'Average Operating Speed: ', format(round(displaydata$Spd18,
2), nsmall = 2), "<br>",
          'Standard Deviation of Operating Speed: ',
format(round(displaydata$SS18, 2), nsmall = 2), "<br>",
          'PSL Equivalent from INRIX: ',
format(round(displaydata$PSL18, 2), nsmall = 2), "<br>",
          'Free Flow Operating Speed: ',
format(round(displaydata$SFF18, 2), nsmall = 2), "<br>",
          'Major Commercial Driveway: ', displaydata$MnrCmmr, "<br>",
          'Minor Commercial Driveway: ', displaydata$MjrCmmr, "<br>",
          'Minor Industrial Driveway: ', displaydata$MnrInds, "<br>",
          'Minor Residential Driveway: ', displaydata$MnrRsdn, "<br>",
          'Major Industrial Driveway: ', displaydata$MjrInds, "<br>",
          'Other Driveway: ', displaydata$Other, "<br>",
          'Major Residential Driveway: ', displaydata$MjrRsdn, "<br>",
          'Facility Type: ', displaydata$facility, "<br>",
          # 'FI Crashes: ', displaydata$Crash_18_22_fi, "<br>",
          # 'PDO Crashes: ', displaydata$Crash_18_22_pdo, "<br>",
          'Total Crashes: ', displaydata$total, "<br>",
          'Equivalent Driveway: ', displaydata$equi_dwy, "<br>",
          'Average Precipitation: ', displaydata$pre_avg, "<br>",
          'Predicted FI Crashes: ', format(round(displaydata$pred_fi,
2), nsmall = 2), "<br>",
```

```
          'Expected FI Crashes: ', format(round(displaydata$EB_fi, 2),
nsmall = 2), "<br>",
          'Predicted PDO Crashes: ',
format(round(displaydata$pred_pdo, 2), nsmall = 2), "<br>",
          'Expected PDO Crashes: ', format(round(displaydata$EB_pdo,
2), nsmall = 2), "<br>",
          'Expected Total Crashes: ',
format(round(displaydata$EB_total, 2), nsmall = 2), "<br>",
          "</div>"  # End scrollable div
        )

        # Update the map: clear previous popups and add a new one
        leafletProxy("MapOut") %>%
          clearPopups() %>%
          setView(lng = lng, lat = lat, zoom = 12) %>%
          addPopups(lng, lat, popupOut, options =
popupOptions(closeButton = TRUE))
      }
    })


    if (nrow(displaydata) > 0) {
      # Remove the geometry column and ensure distinct rows based on
`un_d_nw`
      outputDTdownload <- displaydata %>%
        select(-geometry) %>%
        distinct(un_d_nw, .keep_all = TRUE) %>%

        # Relocate District and County names (assuming names are
`District_Name` and `County_Name`)
        relocate(District, .after = District_Code) %>%
        relocate(County, .after = County_Code)

    } else {
      # Create a single-row data frame with "None" for each column if
there is no data
      outputDTdownload <- data.frame(matrix("None", nrow = 1, ncol =
ncol(displaydata) - 1))
      colnames(outputDTdownload) <- setdiff(names(displaydata),
"geometry")
    }


    # Define download handler for output data with updated facility
codes
    output$downloadData <- downloadHandler(
      filename = function() {
        paste0(
          gsub(" ", "", paste(
            input$DistrictInput, "_",
            input$CountyInput, "_",
            switch(input$FacilityInput,
```

```
                    "All Facilities" = "All",
                    "Urban Two-lane" = "2U",
                    "Urban Two-lane with a TWLTL" = "2T",
                    "Urban Interstate Highway" = "IH",
                    "Urban Multilane Divided" = "MD",
                    "Urban Multilane Undivided" = "MU",
                    "Urban Multilane with a TWLTL" = "MT"
            ),
            input$YearInput
          )),
          ".csv"
        )
      },
      content = function(file) {
        required_columns <- c(
          "un_d_nw", "GID", "FRM_DFO", "TO_DFO", "C_SEC", "HWY",
"STE_NAM", "District_Code",
          "District", "County_Code", "County", "CITY", "RU",
"F_SYSTE", "RU_F_SY", "SPD_MAX",
          "MED_TYP", "MED_WID", "NUM_LAN", "HOV_LAN", "HOV_TYP",
"RB_WID", "SUR_W", "S_TYPE_I",
          "S_WID_I", "S_USE_I", "S_TYPE_O", "S_WID_O", "S_USE_O",
"CURB_L", "SRF_TYP",
          "ADT_YEA", "ADT_CUR", "ADT_ADJ", "K_FAC", "D_FAC",
"TRK_AAD", "HY_1", "HY_2", "HY_3",
          "HY_4", "HY_5", "HY_6", "HY_7", "HY_8", "HY_9", "DESGN_Y",
"LEN_SEC", "LN_MILE",
          "DVMT", "Fclty_C", "Spd18", "SS18", "Spd19", "SS19",
"Spd20", "SS20", "Spd21", "SS21",
          "Spd22", "SS22", "X2019pS", "X2019pA", "X2020pS", "X2020pA",
"X2021pS", "X2021pA",
          "X2022pS", "X2022pA", "X2018_A", "X2018_B", "X2018_C",
"X2018_K", "X2018_O", "X2018_U",
          "X2019_A", "X2019_B", "X2019_C", "X2019_K", "X2019_O",
"X2019_U", "X2020_A", "X2020_B",
          "X2020_C", "X2020_K", "X2020_O", "X2020_U", "X2021_A",
"X2021_B", "X2021_C", "X2021_K",
          "X2021_O", "X2021_U", "X2022_A", "X2022_B", "X2022_C",
"X2022_K", "X2022_O", "X2022_U",
          "MnrCmmr", "MjrCmmr", "MnrInds", "MnrRsdn", "MjrInds",
"Other", "MjrRsdn", "fclty_c1",
          "facility", "Crash_18_2", "Crash_18_1", "total",
"pav_s_wid_", "sw",
          "equi_dwy", "equi_dwy_m", "spd", "pre_avg", "Site_no",
"pred_fi", "EB_fi", "pred_pdo",
          "EB_pdo", "EB_total"
        )

        # Corresponding new column names
        new_column_names <- c(
          "UnqID", "GID", "Frm_Dfo", "To_Dfo", "C_Sec", "Hwy",
"Ste_Nam", "District_Code",
```

```
          "District", "County_Code", "County", "City", "Ru",
"F_Syste", "Ru_F_Sy", "Spd_Max",
          "Med_Typ", "Med_Wid", "Num_Lan", "Hov_Lan", "Hov_Typ",
"Rb_Wid", "Sur_W", "S_Type_I",
          "S_Wid_I", "S_Use_I", "S_Type_O", "S_Wid_O", "S_Use_O",
"Curb_L", "Srf_Typ",
          "Adt_Yea", "Adt_Cur", "Adt_Adj", "K_Fac", "D_Fac",
"Trk_Aad", "Hy_1", "Hy_2", "Hy_3",
          "Hy_4", "Hy_5", "Hy_6", "Hy_7", "Hy_8", "Hy_9", "Desgn_Y",
"Len_Sec", "Ln_Mile",
          "Dvmt", "Fclty_C", "Spd18", "Ss18", "Spd19", "Ss19",
"Spd20", "Ss20", "Spd21", "Ss21",
          "Spd22", "Ss22", "PrcSum19", "PrcAvg19", "PrcSum20",
"PrcAvg20", "PrcSum21",
          "PrcAvg21", "PrcSum22", "PrcAvg22", "A_18", "B_18", "C_18",
"K_18", "O_18", "U_18",
          "A_19", "B_19", "C_19", "K_19", "O_19", "U_19", "A_20",
"B_20", "C_20", "K_20",
          "O_20", "U_20", "A_21", "B_21", "C_21", "K_21", "O_21",
"U_21", "A_22", "B_22",
          "C_22", "K_22", "O_22", "U_22", "MnrCmmr", "MjrCmmr",
"MnrInds", "MnrRsdn", "MjrInds",
          "Other", "MjrRsdn", "Fclty_C1", "Facility", "Cr18_22_FI",
"Cr18_22_PDO", "TotalCrash",
          "Pav_S_Wid", "Sw", "Equi_Dwy", "Equi_Dwy_M", "Spd",
"Prc_Avg", "Site_No", "Predicted_FI",
          "Estimated_FI", "Predicted_PDO", "Estimated_PDO",
"Estimated_Total"
        )

        # Subset and rename the columns
        outputDTdownload <- outputDTdownload[, required_columns, drop
= FALSE]
        colnames(outputDTdownload) <- new_column_names

        write.csv(outputDTdownload, file, row.names = FALSE)
      }
    )

    # Define download handler for variable definitions
    output$downloadDefination <- downloadHandler(
      filename = function() {
        paste0("VariableCodes_7144.xlsx")
      },
      content = function(file) {
        file.copy("ShapeFiles/VariableCodes_7144.xlsx", file)
      }
    )
  })
}

shinyApp(ui, server)
```