Safety Performance on Rural Multilane Roadways in Tennessee

Research Report from University of Tennessee Knoxville | Asad J. Khattak, Numan Ahmad, Amin Mohammadnazar, and Ramin Arvin | November 30, 2021

Sponsored by Tennessee Department of Transportation Long Range Planning

Research Office & Federal Highway Administration

DISCLAIMER

This research was funded through the State Planning and Research (SPR) Program by the Tennessee Department of Transportation and the Federal Highway Administration under ***RES #: 2020-04 Title: A Localized Safety Performance Functions Approach Accounting for "Within" Tennessee Variations on Freeways & Interchanges.*** The research relates to work that commenced during an earlier project**,** RES#: RES2016-27, Research Project Title: Highway Safety Manual, Safety Performance Functions (SPF) & Roadway Calibration Factors

This document is disseminated under the sponsorship of the Tennessee Department of Transportation and the United States Department of Transportation in the interest of information exchange. The State of Tennessee and the United States Government assume no liability of its contents or use thereof.

The contents of this report reflect the views of the author(s) who are solely responsible for the facts and accuracy of the material presented. The contents do not necessarily reflect the official views of the Tennessee Department of Transportation or the United States Department of Transportation.

Technical Report Documentation Page

1. Report No. RES 2020-04	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle *Safety Performance on Rural Multilane Roadways in Tennessee*		5. Report Date November 2021 6. Performing Organization Code
7. Author(s) Asad J. Khattak, Numan Ahmad, Amin Mohammadnazar, and Ramin Arvin		8. Performing Organization Report No. RES 2016-27 & RES 2020-04
9. Performing Organization Name and Address The University of Tennessee-Center for Transportation Research 309 Conference Center Building Knoxville, Tennessee 37996-4133		10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Tennessee Department of Transportation 505 Deaderick Street, Suite 900 Nashville, TN 37243		13. Type of Report and Period Covered Final Report: November 2019 to February 2022 14. Sponsoring Agency Code RES 2020-04

15. Supplementary Notes: Ms. Meredith King helped with editing the report.

Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.

16. Abstract

The Tennessee Department of Transportation (TDOT) is in the process of adopting the Highway Safety Manual (HSM) by developing jurisdiction-specific crash prediction models (termed as Safety Performance Functions or SPFs). This report documents the development and implementation of SPFs for three types of rural multilane roadways. It aims to 1) develop Tennessee-specific SPFs for total and fatal/injury (FI) crashes on rural five-lane undivided-with a two-way left-turn lane (5T) segments, 4-lane divided (4D), and 4-lane undivided (4U) roadway segments, 2) compute TN-specific crash modification factors (CMFs) for the key explanatory variables, 3) explore the role of new variables such as speed limit and surrounding land use in SPFs, and 4) prepare new TN-specific prediction tools (spreadsheets) for future use by TDOT. The project required extensive data extraction, integration, and analysis. Specifically, crash, roadway, and traffic data were extracted using various sources and websites maintained by TDOT. Based on the results, this report recommends strategies that can advance the analysis of safety data.

17. Key Words	18. Distribution Statement
Highway Safety Manual, Calibration Procedure, Calibration Spreadsheets, Crash Modification Factors, Tennessee Department of Transportation	No restriction. This document is available to the public from the sponsoring agency at the website http://www.tn.gov/.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	90	N/A

Executive Summary

To enhance safety on Tennessee roadways, this report documents the estimation and implementation of Safety Performance Functions (SPFs) for three types of rural roadways. The study 1) develops Tennessee-specific SPFs for total and fatal/injury (FI) crashes on rural five-lane undivided segments with a two-way left-turn lane (5T), as the 2010 Highway Safety Manual (HSM) does not provide crash prediction models for such roadways. Also included are rural 4-lane divided (4D), and 4-lane undivided (4U) roadway segments, 2) computes TN-specific crash modification factors (CMFs) for the key explanatory variables, 3) explores the role of new variables such as speed limit and surrounding land use (e.g., commercial, mixed, and residential) in SPFs, and 4) prepares new TN-specific calibration spreadsheets for rural 5T, 4D and 4U roadway segments for future use by Tennessee Department of Transportation (TDOT). The project required extensive data extraction, integration, and analysis. Specifically, crash, roadway, and traffic data were extracted using various sources and websites maintained by TDOT. Based on the analysis, this report recommends strategies that can assist TDOT in advancing safety analysis in Tennessee.

Key Findings

To estimate crashes, the research team collected and used Tennessee crash, traffic, and road inventory data (2013-2017). A separate analysis was conducted for total and FI crashes for the three rural (4D, 4U, and 5T) roadway types. The study applied both Poisson and negative binomial models using the HSM functional form. For consistency with HSM and to capture any over-dispersion in the crash data, the negative binomial models are preferred. In Table E-1, results of the models are shown. Note that "a" and "b" indicate the constant and parameter values of average annual daily traffic (AADT), whereas segment length is used as an exposure variable consistent with the HSM (2010) functional form. The results in the table indicate that the base models for Tennessee give estimates of parameters that are consistent with the HSM (2010) default values for total and FI crashes.

Table E-1 RESULTS OF MODELING USING TENNESSEE DATA FOR RURAL ROADS

Roadway Type	***TN-Specific parameters***		***Default HSM (2010) parameters***	
	a	***b***	***a***	***b***
1. Rural 5T Roadway Segments (N = 205)				
Total Crashes	-11.660	1.410	None	None
FI Crashes	-9.816	1.055	None	None
2. Rural 4D Roadway Segments (N = 271)				
Total Crashes	-3.078	0.430	-9.025	1.049
FI Crashes	-4.614	0.456	-8.505	0.874
3. Rural 4U Roadway Segments (N = 81)				
Total Crashes	-7.950	1.092	-9.653	1.176

The key findings for each of the three rural roadway types are listed below.

- Rural 5T roadways. The average total crashes per year are expected to be lower on rural 5T roadway segments that have wider center lanes (2WLTL), wider travel lanes, and higher speed limits (miler per hour (MPH)). Higher frequencies are observed on 5T roadway segments with more driveways. The CMFs for the key variables were computed using their corresponding parameter estimates in the enhanced TN-SPFs for total crashes. For instance, the CMF for speed limit of 45 MPH or higher is 1.000 and for speed limits of 40 MPH and 35 MPH, it is 1.100 and 1.201, respectively. Similarly, for 12 ft or higher lane widths, the CMF is 1.000, increasing to 1.287 and 1.573 if the lane widths are 11 ft and 10 ft, respectively.
- Rural 4D roadways. For such roadway segments, the average total crashes per year are lower with wider inner (left) shoulders, higher speed limits (MPH), and with the presence of rumble strips along the inner shoulder. Notably, land use can have a substantial impact—a segment that passes through commercial or mixed land use development (compared with residential developments) the crash frequencies are higher. The computed CMFs show that if the inner (left) shoulder width (in ft) is 4 ft or higher, the CMF is 1.000, which increases to 1.114, 1.228, and 1.342 for the inner shoulder widths of 3 ft, 2 ft, and 1 foot, respectively.
- Rural 4U roadways. Lower crash frequencies are observed on 4U roadways when the segments have wider lanes, higher speed limits (MPH), wider outer (right) shoulders, and streetlights. When such roadways pass through areas with commercial or mixed-use developments (compared to residential developments), the yearly crash frequencies are higher. For CMFs, the results indicate that if the outer (right) shoulder width is 5 ft or higher, the CMF for outer shoulder width is 1.000, and it increases to 1.060, 1.121, and 1.182 for widths of 4 ft, 3 ft, and 2 ft, respectively.

A unique aspect of the study is the use of new variables in predicting crashes on rural 5T, 4D, and 4U roadways, i.e., speed limit and surrounding land use (commercial, mixed, and residential). This study has helped us understand the correlations of these new variables with crash frequencies. Finally, the regression coefficients for the constant and the AADT from the base-case TN-SPF and CMFs for roadway and other important variables from the enhanced TN-SPFs are used to prepare the calibration spreadsheets which predict crashes for TN-specific base and enhanced (after applying TN-CMFs) conditions. Given that they are calibrated using Tennessee data, the spreadsheets can provide relatively more accurate predictions compared with the HSM (2010) and Federal Highway Administration (FHWA) default values.

Key Recommendations

For improving the safety analysis in Tennessee, this report provides a list of recommendations below.

- Use of HSM procedures and tools. The Tennessee Department of Transportation's adoption of the HSM procedures and corresponding investments in calibration procedures for safety improvements are ready for implementation. This study provides the data and tools for the analysis of segments on three rural roadway types. Specifically, TDOT can use these tools for rural 5T, 4D, and 4U roadway segments to predict crashes and any reductions associated with safety improvements in crash modification factors.

- Countermeasure selection. Based on the findings of this study, TDOT's Strategic Transportation Investments Division can use spreadsheet tools to explore countermeasures that can substantially improve safety on rural roads in Tennessee.
- Periodic updating of calibration factors and safety performance functions. The results of this study are based on crash data collected between 2013-2017. Given the substantial spatial and temporal variability in conditions, updating the calibration factors will help in applying the procedures embedded in the spreadsheets.
- For future research, TDOT can consider using emerging methods for crash prediction. These include artificial intelligence (AI) and machine learning (ML) methods to enhance safety outcome predictions. Notably, AI/ML methods can enhance prediction accuracy (though they are weaker in generating inferential knowledge). In this regard, the more robust heterogenous ensemble methods such as stacking can combine both traditional statistical models and AI/ML methods. Stacking is relevant for rural 5T roadway segments as the HSM (2010) does not provide crash prediction models for these roadway types. Note that stacking has shown promising performance for 5T urban and suburban arterials in Tennessee. For details, please see a relevant research paper in the Appendix.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

To improve transportation safety, this study uses historical crash and road inventory data. The study develops provides crash prediction procedures that are based on Safety Performance Functions (SPFs) for different types of rural roadways [1]. The study focuses on rural four-lanes undivided (4U), four-lanes divided (4D), and five-lanes undivided (5T) that have a two-way left-turn lane (2WLTL) [2, 3]. An earlier report (for Phase II) has documented the development of Tennessee-specific SPFs for total crashes on rural 4D and 4U rural roadway segments [4]. This report adds to the previous analysis by:

- Developing Tennessee-specific SPFs for total and fatal/injury (FI) crashes on rural five-lane undivided-with a two-way left-turn lane (5T) segments as (HSM (2010) does not provide crash prediction models for such roadways), as well as rural 4-lane divided (4D), and 4-lane undivided (4U) roadway segments.
- Computing TN-specific crash modification factors (CMFs) for the key explanatory variables.
- Exploring the role of new variables such as speed limit and surrounding land use (e.g., commercial, mixed, and residential) in SPFs.
- Preparing new TN-specific calibration spreadsheets for rural 5T, 4D, and 4U roadway segments for future use by the Tennessee Department of Transportation (TDOT).

This project required extensive data extraction, integration, and analysis, e.g., for rural 5T roadway segments, the crash, roadway, and traffic data were extracted using various sources and websites maintained by TDOT.

Chapter 2 Methodology

2.1. Data Description

For the rural 4D and 4U roadway segments, the research team had already collected data for Phase II of the project which included data on 5-years (2013-2017) crashes, average annual daily traffic (AADT), roadway factors, and other important variables including speed limit (MPH) and land use. To develop TN-SPFs for FI crashes on rural 4U and 4D roadway types, additional data was collected on FI crashes from 2013 to 2017 which was not collected during an earlier project, i.e., Phase II of the HSM project. Based on the Phase II project, a cleaned random sample of 271 segments for rural 4D roadway segments was used. For rural 4U roadway segments, the final sample includes a total of 81 segments.

For rural 5T roadway segments, the research team first identified a total of 482 roadway segments across Tennessee by selecting the appropriate query (i.e., functional classification, number of lanes, and presence of specific feature "2WLTL" separating opposite directional traffic flow) in the Enhanced Tennessee Roadway Information Management System (E-TRIMS). After excluding shorter segments and those with incomplete data for AADT, a total of 205 segments of rural 5T roadways were considered for developing TN-SPFs. Next, data was collected on total and FI crashes from 2013-2017, AADT (2013-2017), roadway variables for the 271 rural 5T roadway segments. The data extraction procedure is discussed below.

- The 5-year crash data (2013-2017) were extracted from the police crash reports in E-TRIMS (https://e-trims.tdot.tn.gov).
- The TDOT's Image Viewer Software in E-TRIMS (https://e-trims.tdot.tn.gov) was used to extract roadway variables. A substantial amount of effort went into extracting the number of driveways along rural 5T roadway segments. To the best of the researchers' knowledge, this variable has not been considered in the analysis of rural 5T roadway segments. Although the driveway variable is used in urban and suburban arterials segments in HSM [1].
- For traffic data, year-wise (2013-2017) AADT data for rural 5T roadway segments was collected using TDOT's traffic history application (https://www.tdot.tn.gov/APPLICATIONS/traffichistory).
- Finally, the three data files including crash, traffic, and roadway geometric files were merged for further analysis based on the beginning log mile (BLM), ending log mile (ELM), county, route name, sequence number, and station number; also see [4].

For analysis, the average of 5-year (2013-2017) of total crashes, FI crashes, and AADT on each roadway segment for a specific rural roadway type was computed. After computing the average of 5-years of total and FI crashes, the values were rounded up to the nearest whole numbers for two reasons. The key rationale behind this was that count data models can only be applied to count (integer) data, i.e., 0, 1, 2, and 3, and so on. Furthermore, the average of 5-year (2013-2017) AADT "(AADT 2013 + AADT 2014 + … +AADT 2017) / 5)" was also computed for each roadway segment belonging to the three types of rural roadway segments which was then used as a predictor (explanatory variable) in the TN-SPFs for the three roadway types.

2.2. Methodology/Data Analysis

Given the desire to predict crash frequency, Poisson and negative binomial regressions are appropriate [7-13]. Notably, Poisson regression assumes that the mean is equal to the variance, which is often not the case with crash data. To account for the potential overdispersion in the data, negative binomial regression can be applied which includes an overdispersion parameter. Poisson and negative binomial regression can be shown in Equations 1 and 2 respectively as below [2-4].

$$\lambda_i = EXP(\beta X_i) \tag{1}$$

$$\lambda_i = EXP(\beta X_i + \varepsilon_i) \tag{2}$$

Where, λ_i indicates expected number of crashes on a specific roadway segment i, β is a vector of coefficients of the significantly associated explanatory factors X_i, and ε_i is an error term that indicates the overdispersion parameter in the negative binomial (NB) regression.

2.2.1. TN-Specific Base-Case SPFs

Similar to the HSM (2010) functional form, base-case TN-SPF was estimated below [2, 3].

$$N_{TN-SPF} = e^{(\beta_0)} * \ AADT^{(\beta_1)} * Exposure\ (segment\ length) \tag{3}$$

After applying NB regression using the functional form, the values of β_0 and β_1 are obtained which indicates the values of TN-Specific parameters including "a" and "b" in similar to base-case HSM SPFs [1]. Note that the values of "a" and "b" stands for the coefficients of constant and AADT respectively, For the "c" parameter, the equation for "k" in HSM (2010) is used as below [1].

$$k = \frac{1}{e^{(c+\ln L)}} \tag{4}$$

Where "k" indicates the overdispersion parameter in the NB model which can be used to compute the value of "c" parameter using the equation below.

$$c = -(\ln L + \ln k) \tag{5}$$

Once the TN-Specific values of "a", "b", and "c" are computed for the total crashes, these new values can be used to predict the number of total crashes at the base conditions in TN. A similar procedure could be used to determine the values of "a", "b", and "c" for the FI crashes on the three roadway types.

2.2.2. TN-Specific Enhanced SPFs: Computing TN-Specific CMFs

To estimate TN-Specific enhanced SPFs for total crashes on rural 4D, 4U, and 5T roadway segments, the roadway and other important variables (speed limit and land use) in the SPFs were used while keeping AADT (logarithmic form) and segment length (as exposure) for consistency with the HSM (2010) functional form. By doing so, Equation (3) could be modified as below [2-4].

$$N_{TN-SPF\ (Enhanced)} = e^{(\beta_0)} * \ AADT^{(\beta_1)} * Exposure\ (segment\ length) + X_i^{(\beta_i)} \quad (6)$$

Note that β_i indicates the estimated parameters of X_i which is a set of key roadway variables. To predict crashes for the enhanced conditions on a specific roadway type (rural 4D, 4U, and 5T), similar to the HSM CMFs, the team used the parameters (coefficients) of the key explanatory variables in the final enhanced TN- SPFs to compute the corresponding CMFs. To compute the percentage change (increase or decrease) in the average of 5-year total crashes associated with a unit increase in a specific explanatory variable, the following equation was used:

$$Percentage\ (\%)\ Change\ in\ Crashes = \left(1 \pm e^{(\beta_{Xi})}\right) * 100 \quad (7)$$

Note that for Equation (7) the parameter sign can be negative or positive. If the parameter sign was negative for a specific variable, $Exponent\ (\beta_{Xi})$ was subtracted from 1 to compute the percentage change in the average of total 5-year crashes. The change is associated with a unit increase in the value of the specific variable from its mean value in the sample. Similarly, if the parameter sign of a particular variable was positive, the value of $Exponent\ (\beta_{Xi})$ is added to 1 to compute the percentage change in crash frequency. This change is associated with a unit increase in the value of the specific variable from its mean value in the sample. To compute CMFs for a specific variable, the percentage change (increase or decrease) was considered. The predicted crash frequency per year obtained using "a" and "b" parameters from base-case TN-SPFs can be adjusted with the TN-CMFs with additional explanatory variables related to the roadway geometry, speed limit, and surrounding land use.

2.2.3. Comparing Model Performance

To access how well a statistical model fits the data, several goodness-of-fit measures including loglikelihood (LL) at convergence and Akaike Information criteria (AIC) were used. Note that the lower value of AIC and LL at convergence which maximizes the value of McFadden's Pseudo (R^2) of the model indicates a relatively better model. Furthermore, to understand and compare the prediction performance of models, the mean absolute error (MAE) was used as given below [2-4].

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| \quad (8)$$

The value of n is the total number of roadway segments, and f_i and y_i indicate predicted and observed crash frequency, respectively. Low values of MAE indicate higher prediction accuracy.

Chapter 3 Results and Discussion

The development of new TN-SPFs for rural 5T, 4D, and 4U roadway types and associated statistical analysis is discussed below.

3.1. Rural Five-lane Undivided Roadway Segments

3.1.1. Descriptive Statistics

The mean value of the average of 5-years of total crashes, FI crashes, and AADT (in 1000s) on 5T roadways are found to be 1.68, 0.43, and 8.85, respectively (Table 3.1). The average segment length is 0.51 miles whereas the minimum and maximum values for the segment length are found to be 0.11 and 3.29 miles, respectively. Statistics reveal that the average number of driveways on rural 5T roadways (along both sides) is 12.49, ranging from 0 to 95 driveways. For the indicator variables (dummy variables), the mean for a specific indicator variable can be interpreted as a percentage (when multiplied by 100). For instance, statistics indicate that 52% of the rural 5T roadway segments pass through locations with commercial or mixed commercial land use. For descriptive statistics of other variables, please see Table 3.1.

To have a deeper understanding of the distribution of rural 5T roadway segments based on the average of 5-years total crashes, FI crashes, and AADT, please see Figure 3.1, Figure 3.2, and Figure 3.3, respectively. For the distribution of rural 5T roadway segments based on the number of driveways along both sides of the roadway segment, please see Figure 3.4.

TABLE 3-1 DESCRIPTIVE STATISTICS OF KEY VARIABLES (RURAL 5T ROADWAY SEGMENTS)

Variables	*N*	*Mean*	*S.D.*	*Min*	*Max*
*Average of 5-years of all crashes**	205	1.68	2.55	0.00	19.00
*Average of 5-years of FI crashes**	205	0.43	0.81	0.00	6.00
Average of 5-years AADT (in 1000s)	205	8.85	3.72	2.22	19.98
Segment length (miles)	205	0.51	0.52	0.11	3.29
Number of driveways along segments (both sides)	205	12.49	13.40	0.00	95.00
Density of driveways (number per mile) (both sides)	205	27.19	19.12	0.00	171.43
Width of 2-way left-turn lane (TWLTL) in ft	205	12.08	0.59	10.00	14.00
Indicator for land use (1= commercial/mixed, 0= residential)	205	0.52	0.50	0.00	1.00
Outer shoulder width (ft)	205	7.47	4.20	1.00	12.00
Lane width (ft)	205	11.90	0.33	10.00	12.00
Presence of rumble strip along outer shoulder (1/0)	205	0.26	0.44	0.00	1.00
Presence of lighting (1/0)	205	0.47	0.50	0.00	1.00
Speed limit (MPH)	205	46.44	7.98	25.00	65.00

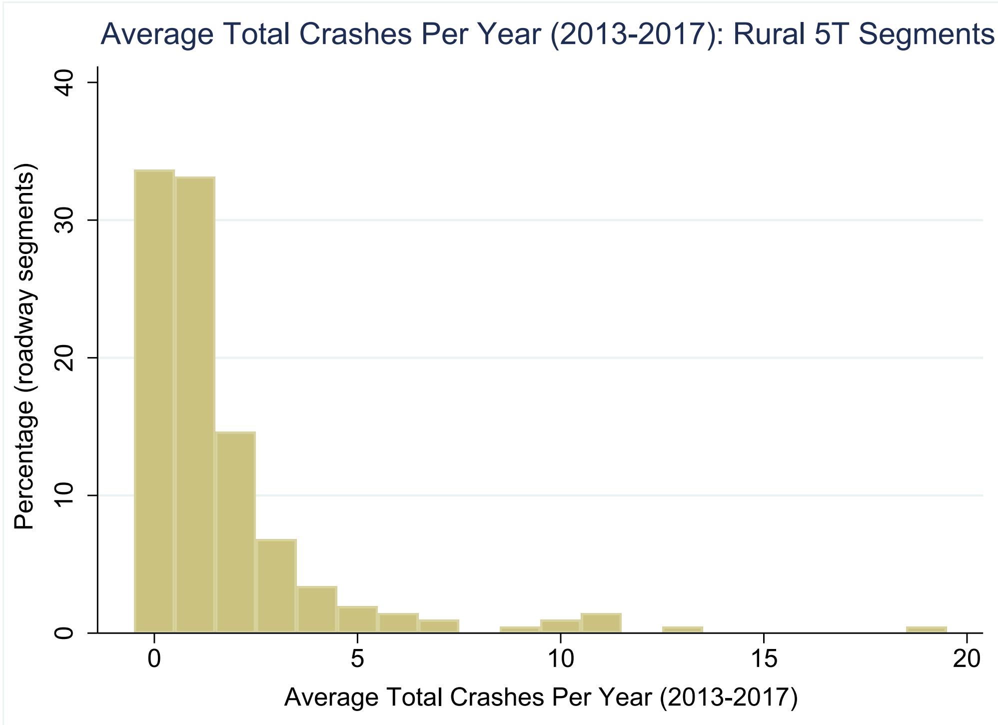


Figure 3.1 Distribution of Rural 5T Roadway Segments based on Average Total Crashes per Year

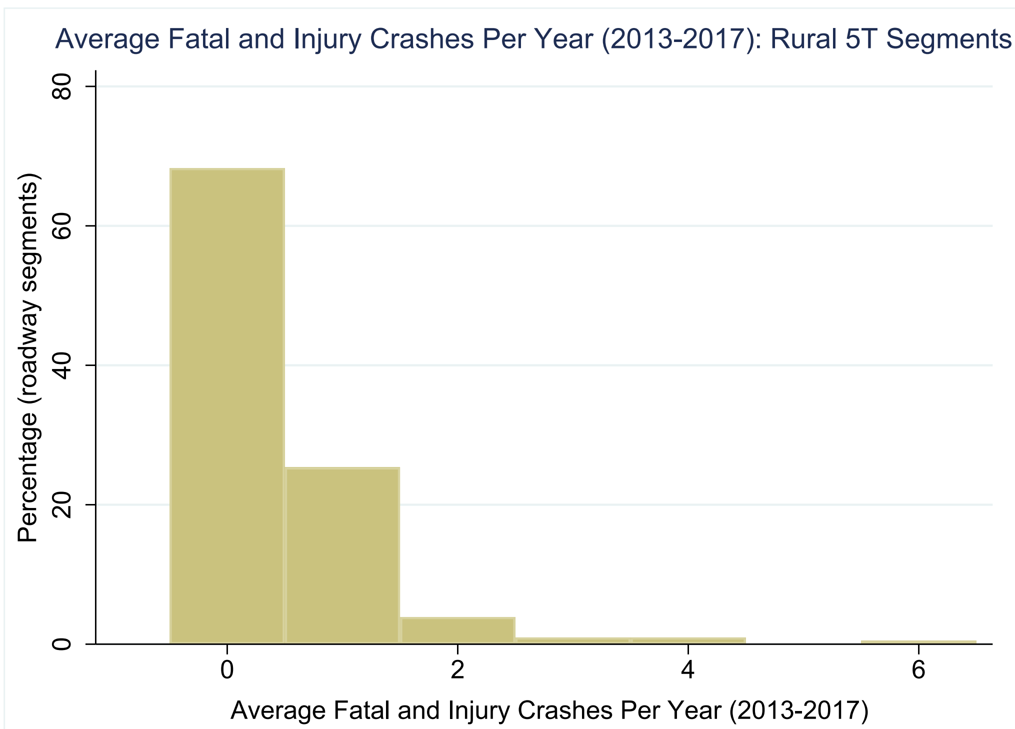


Figure 3.2 Distribution of Rural 5T Roadway Segments based on Average FI Crashes per Year

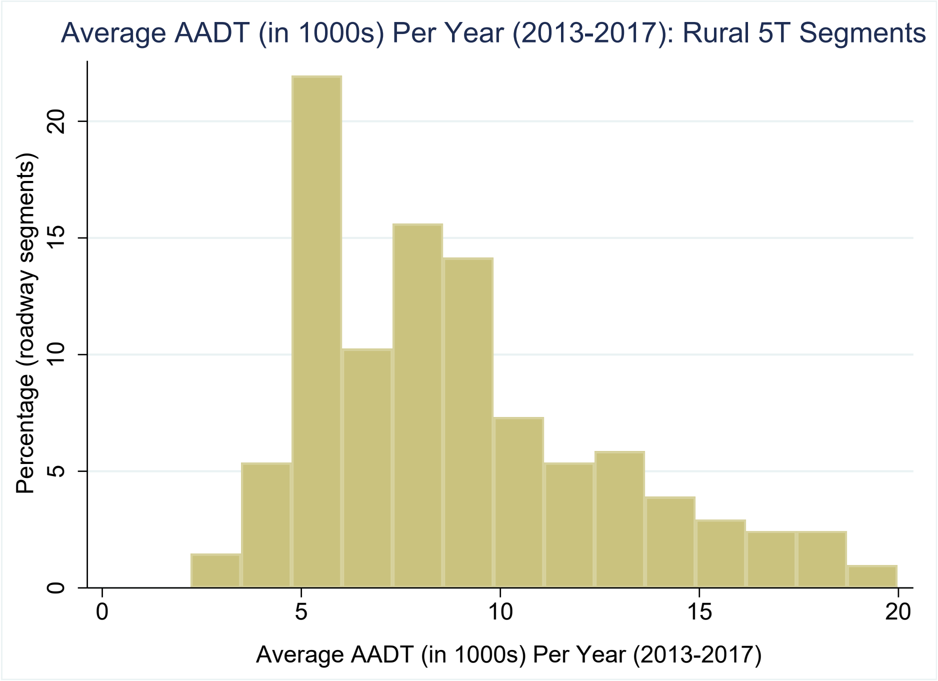


Figure 3.3 Distribution of Rural 5T Roadway Segments based on Average AADT (in 1000s)

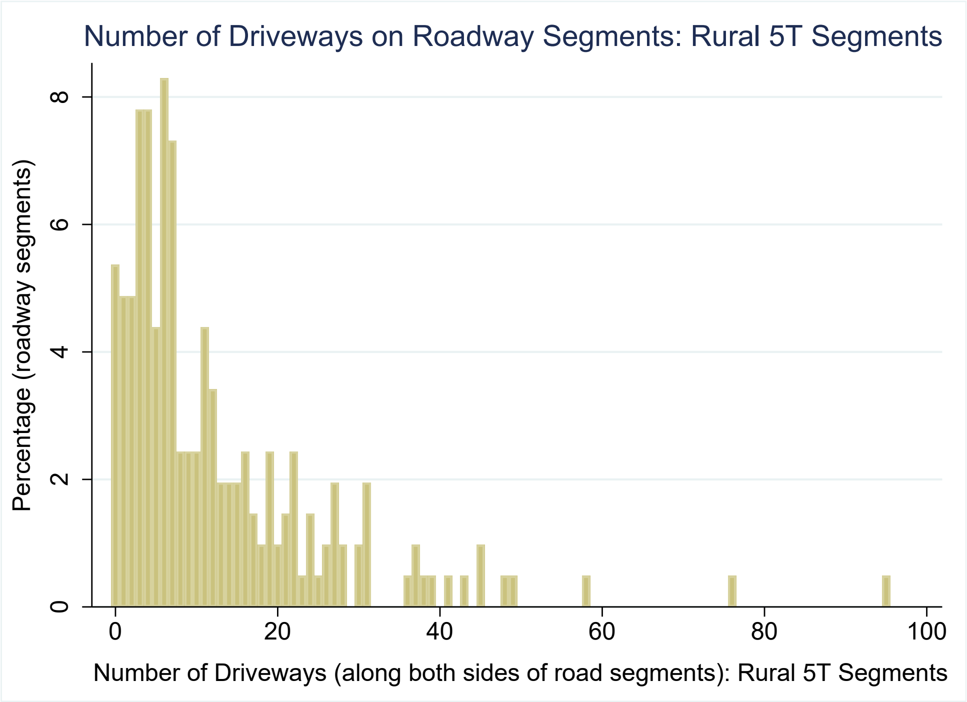


Figure 3.4 Distribution of Rural 5T Roadway Segments based on Number of Driveways

3.1.2. Modeling Results

3.1.2.1. Tennessee-Specific SPFs for Total Crashes

For completeness, both Poisson and negative binomial models for overall crashes using the HSM (2010) functional form (Table 3.2) are reported. For consistency with HSM SPFs and the evidence for the presence of overdispersion, the research team consider the negative binomial as the final model. Significant evidence of overdispersion was found in the data (Table 3.2). Furthermore, based on values of AIC and LL at convergence, the negative binomial model shows better performance compared to the Poisson model. The values of TN-specific regression parameters for total crashes on rural 5T roadway segments including "a" and "b" are found to be -11.660 and 1.410, respectively. The HSM does not include SPFs for rural 5T roadway segments; notably, the default values of "a" and "b" for total crashes on rural 4D roadway segments are -9.025 and 1.049, respectively, whereas the default values of "a" and "b" for rural 4U roadway segments are - 9.653 and 1.176, respectively (HSM 2010). It can be seen that the new TN-specific values of "a" and "b" parameters for rural 5T roadway segments are fairly close to the corresponding default values of these parameters on rural 4D and 4U roadway segments in HSM [1].

TABLE 3-2 BASE-CASE TN-SPECIFIC SPFS FOR TOTAL CRASHES ON RURAL 5T ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 1A)***			***NB (Model 1B)***		
	Coef.	**Std. Err.**	**t-stat**	**Coef.**	**Std. Err.**	**t-stat**
Average 5-years AADT: ln form	1.249	0.134	9.35	1.410	0.180	7.84
Segment length (mile): Exposure	1	---	---	1	---	---
Constant	-10.25	1.24	-8.28	-11.660	1.653	-7.05
Over-dispersion parameter (alpha)	---	---	---	0.242	0.072	3.36
Test for alpha significantly different than 0						
Chi-square test statistics	---			34.95		
Prob. (Chi-square)	---			<0.001		
Summary Statistics						
Log-likelihood (Convergence)	-314.575			-297.098		
AIC	633.150			600.195		
Degrees of freedom	2			3		
Sample Size (N)	205			205		

Using equation (5), the mean value of the "c" parameter for total crashes on rural 5T roadway segments is found to be 2.448. Note that the default HSM values of "c" parameters for total crashes on rural 4D and 4U roadway segments are 1.549 and 1.675, respectively [1].

In the enhanced TN-SPFs for rural 5T roadway segments, the over-dispersion parameter is statistically insignificant. Nevertheless, the negative binomial model (model 1D) is still considered as the final model for consistency with the HSM functional form [1]. The estimation results of Model 1D reveal that total crashes per year increase with an increase in the number of driveways in a non-linear manner as shown in Figure 3.5. Furthermore, according to the estimation results, total crashes per year are expected to be lower on rural 5T roadway segments with wider lanes, wider 2WLTL, and higher speed limits (MPH). The land-use variable (residential/commercial) did not show any statistical significance in the final model; hence it was dropped. Note that the parameter estimates of the key roadway variables in Model 1D are used to compute the TN-specific CMFs on rural 5T roadway segments which are briefly discussed in the subsequent subsection. The incidence rate ratios (IRR) for a specific explanatory variable in the enhanced TN-SPFs indicate the change (increase/decrease) in the average total crashes per year with a unit increase in a specific variable when all other variables are held at constant values in the model. For instance, the results indicate that with a unit increase in width of 2WLTL, the average total crashes per year on rural 5T roadway segments would decrease by 27.11% (= 1-0.7289*100).

TABLE 3-3 ENHANCED TN-SPECIFIC SPFS FOR TOTAL CRASHES ON RURAL 5T ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 1C)***			***NB (Model 1D)***		
	Coef.	***t-stat***	***IRR***	***Coef.***	***t-stat***	***IRR***
Average 5-years AADT (in 1000s): ln form	1.21	8.17	3.37	1.20	7.64	3.33
Width of two-way left-turn lane (ft)	-0.33	-3.07	0.72	-0.32	-2.73	0.73
Speed limit (MPH)	-0.02	-2.51	0.98	-0.02	-2.47	0.98
Lane width (ft)	-0.34	-2.15	0.71	-0.34	-1.94	0.71
Number of driveways (both sides)	0.02	1.90	1.02	0.01	1.73	1.01
*Number of driveways*Number of driveways*	0.00	-2.27	1.00	0.00	-2.02	1.00
Segment length (mile): Exposure	1.00	---	---	1.00	---	---
Constant	-1.05	-0.48	---	-1.19	-0.51	---
Overdispersion parameter (alpha)	---	---	---	0.03	0.77	---
Test for alpha significantly different than 0						
Chi-square test statistics	---			0.77		
Prob. (Chi-square)	---			0.191		
Summary Statistics						
Log-likelihood (Convergence)	-275.26			-274.87		
AIC	564.52			565.75		
Degrees of freedom	7			8		
Sample Size (N)	205			205		

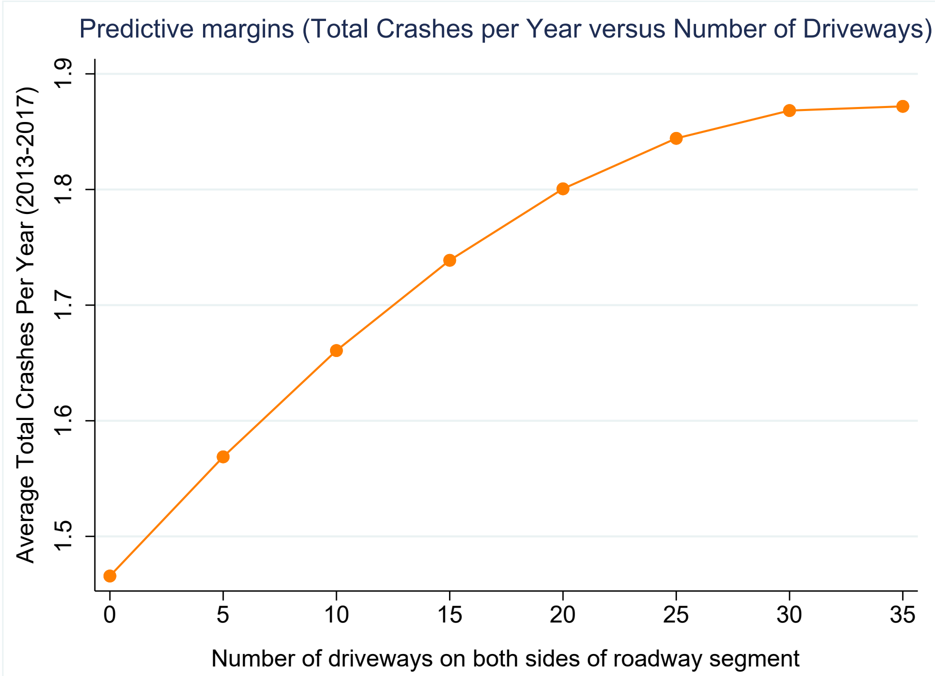


Figure 3.5 Predicted Total Crashes Per Year versus Number of Driveways (Model 1D)

For comparing model prediction performance, the results are shown in Table 3.4. Importantly, while the base-case TN-specific negative binomial model (model 1B) shows better performance compared with the base-case TN-specific Poisson model (model 1A) based on AIC and loglikelihood (LL) at convergence values [5, 6], it shows weaker prediction accuracy based on MAE compared to the base model (Table 3.4). To summarize, model 1D (enhanced TN-specific negative binomial model) is selected as the best-fit model based on AIC, LL, and in-sample MAE.

Table 3-4 MEAN ABSOLUTE ERROR: COMPARING MODELS FOR TOTAL CRASHES ON RURAL 5T ROADWAY SEGMENTS

Model	***Mean Absolute Error***					***MAE (%)***	***Model Summary***		
	Obs.	***Mean***	***S.D.***	***Min.***	***Max.***		***LL (Convergence)***	***AIC***	***Deg. of freedom***
Model 1A	205	1.11	1.84	0.03	15.35	Base	-314.57	633.15	2
Model 1B	205	1.17	2.03	0.00	16.82	4.92	-297.10	600.20	3
Model 1C	205	---	---	---	---	---	---	---	---
Model 1D	205	0.89	1.10	0.01	8.69	-20.10	-274.88	565.76	8

Note: Model 1A and 1C indicate the base-case and Enhanced TN-Specific Poisson models. Since the negative binomial models (Model 1B and Model 1D) were selected, the MAE of only Model 1B and 1D are compared with the base model (Model 1A). Notably, the HSM (2010) does not provide any SPF for rural 5T roadway segments which could be considered as a true base for comparison. For comparing the MAE with the HSM Enhanced-SPF, the formula in equation (9) below was used.

$$\text{MAE } (\%) = \frac{\text{(Mean MAE of Model 1B/1C/1D – Mean MAE of Model 1A)}}{\text{(Mean MAE of Model 1A)}} * 100 \quad (9)$$

(Percentage of MAE equals, open parenthesis, "Mean MAE of Model 1B/1C/1D" minus "Mean MAE of Model 1A", close parenthesis, over, open parenthesis, "Mean MAE of Model 1A", close parenthesis, times 100)

3.1.2.2. TN-Specific SPFs for Fatal and Injury (FI) Crashes: Rural 5T Roadway Segments

Referring to the base-case TN-SPFs for FI crashes on rural 5T roadway, the over-dispersion parameter in the negative binomial model (model 2B) is found to be insignificant. For consistency with HSM, Model 2B was selected as the final model. The TN-specific values of “a” and “b” for FI crashes on rural 5T roadway segments are found to be -9.8163 and 1.0550, respectively (Table 3.5). Note that the HSM (2010) does not provide the default values of “a”, “b”, and “c” for FI crashes on rural 5T segments. Notably, the HSM default values of “a” and “b” parameters for FI crashes on rural 4D roadway segments are -8.505 and 0.874, respectively. On other hand, the values of “a” and “b” for FI crashes on rural 4D roadway segments are found to be -8.577 and 0.938, respectively. The TN-Specific values of the two parameters for FI crashes on rural 5T roadway segments are closer to the corresponding HSM default values on rural 4D and 4U roadway segments. Notably, the TN-Specific value of “c” becomes 0 due to the lack of over-dispersion as the Poisson model shows better performance, and the over-dispersion parameter is found to be 0.

Table 3-5 BASE-CASE TN-SPECIFIC SPF FOR FI CRASHES ON RURAL 5T ROADWAY SEGMENT

Explanatory Variables	***Poisson (Model 2A)***			***NB (Model 2B)***		
	Coef.	***Std. Err.***	***t-stat***	***Coef.***	***Std. Err.***	***t-stat***
Average of 5-years AADT: ln form	1.0553	0.2597	4.06	1.0550	0.2597	4.06
Segment length (miles): Exposure	1	---	---	1	---	---
Constant	-9.8192	2.3991	-4.09	-9.8163	2.3989	-4.09
Over-dispersion parameter (alpha)	---	---	---	$7.68*10^{-6}$	0.0005	0.001
Test for alpha significantly different than 0						
Chi-square test statistics	---			0.00		
Prob. (Chi-square)	---			0.500		
Summary Statistics						
Log-likelihood (Convergence)	-140.0876			-140.0877		
AIC	284.1753			286.1753		
Degrees of freedom	2			3		
Sample Size (N)	205			205		

The results of the enhanced TN-Specific SPFs for FI crashes on rural 5T roadway segments are provided in Table 3.6. The Poisson model (Model 2C) shows slightly better performance compared with the negative binomial (Model 2D). Nevertheless, Model 2D was chosen as the final model for consistency with HSM (2010). The results indicate that the average of 5-years FI crashes on rural 5T roadway segments increases with an increase in the number of driveways (both sides) on the rural 5T roadway segments but decreases with an increase in width of 2WLTL on these segments. For completeness, the speed limit (MPH) variable was kept in the model, but it did not show a statistically significant relationship with the average FI crashes. Note that the land-use variable (residential, commercial, etc.) did not show any significant association with the FI crashes. The squared term for the number of driveways on the rural 5T roadway segments is found to be statistically significant indicating non-linearity in the effects of the number of driveways on the FI crashes (Figure 3.6).

Table 3-6 ENHANCED TN-SPECIFIC SPFS FOR FI CRASHES ON RURAL 5T ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 2C)***			***NB (Model 2D)***		
	Coef.	***t-stat***	***IRR***	***Coef.***	***t-stat***	***IRR***
Av. 5-years AADT (in 1000s): ln form	1.0998	3.88	3.0034	1.0998	3.88	3.0036
Width of two-way left-turn lane (ft)	-0.4363	-2.48	0.6464	-0.4363	-2.48	0.6464
Number of driveways (both sides)	0.0323	1.98	1.0328	0.0324	1.98	1.0328
*Number of driveways*Number of driveways*	-0.0004	-1.96	0.9995	-0.0004	-1.96	0.9995
Segment length (mile): Exposure	1	---	---	1.0000	---	---
Speed limit (MPH)	-0.0103	-0.69	0.9897	-0.0103	-0.69	0.9897
Constant	-4.8611	-1.40	---	-4.8613	-1.40	---
Overdispersion parameter (alpha)	---	---	---	$2.28*10^{-7}$	$1.73*10^{-3}$	---
Test for alpha significantly different than 0						
Chi-square test statistics	*---*			*0.00*		
Prob. (Chi-square)	*---*			*0.500*		
Summary Statistics						
Log-likelihood (Convergence)	*-132.4691*			*-132.4691*		
AIC	*276.9382*			*278.9382*		
Degrees of freedom	*6*			*7*		
Sample Size (N)	205			205		

Note: The incidence rate ratios (IRR) for a specific explanatory variable in the enhanced TN-SPFs indicate the change (increase/decrease) in the average total crashes per year with a unit increase in a specific variable when all other variables are held at constant values in the model. If the value of IRR is less than or greater than 1, it shows the corresponding reduction or increase in total crashes per year by the time it is less than or greater than 1, respectively.

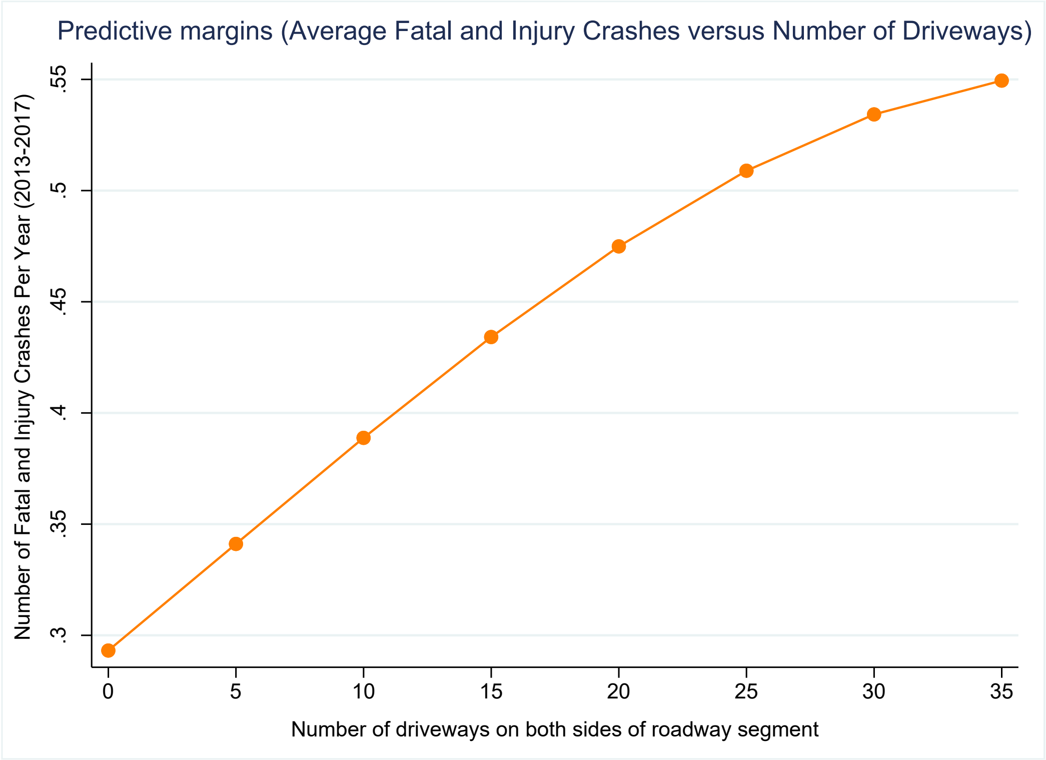


Figure 3.6 Predicted FI Crashes Per Year versus Number of Driveways (Model 2D)

Considering Model 2A (Base-case TN-Specific Poisson model) as a base, the MAE of the other models was computed. The average MAE reduces by 1.57% for Model 2D compared to the base model where Model 2B (Base-case TN-Specific NB) does not show similar prediction accuracy to Model 2A (Base-case TN-Specific Poisson) (Table 3.7).

Table 3-7 MEAN ABSOLUTE ERROR: COMPARING MODELS FOR FI CRASHES ON RURAL 5T ROADWAY SEGMENTS

Model	***Mean Absolute Error***					***Comparison: Mean MAE (%)***	***Model Summary***		
	N	***Mean***	***S.D.***	***Min***	***Max***		***LL (Convergence)***	***AIC***	***Deg. of freedom***
Model 2A	205	1.358	2.16	0.026	18.15	Base	-140.09	284.18	2
Model 2B	205	1.358	2.16	0.026	18.15	0.00	-140.09	286.18	3
Model 2C	205	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Model 2D	205	1.337	2.06	0.017	16.45	-1.57	-132.46	276.92	7

Note: Model 2A and 2C indicate base-case and Enhanced TN-Specific Poisson models. Since the negative binomial counterparts (Model 2B and Model 2D respectively) were selected for consistency with HSM (2010), the MAE of only Model 2B and 2D are compared with the base-case TN-Specific Poisson model (Model 2A). This is because the HSM (2010) does not provide any SPF for rural 5T roadway segments which could be considered as a base for comparison. For comparing the MAE with the HSM Enhanced-SPF, the Mean MAE formula was used, provided above.

3.1.2.3. CMFs for Key Factors Using TN-SPFs for Total Crashes on Rural 5T Segments

The distribution and parameter estimates (based on Model 1D) for the key variables used in the final TN-SPFs for average total crashes per year are shown in Table 3.8. The percentage change including an increase (+) or decrease (-) in average total crashes per year with a unit increase in the specific explanatory variable is shown in Table 3.8. The findings indicate that a unit decrease in width of 2WLTL (ft) below its mean value (~12 ft) may increase the average total crashes per year by 27.10%. Similarly, a unit decrease in lane width (ft) below its mean value (~12 ft) is expected to increase the total crashes per year by 28.65%. On other hand, the findings suggest that a unit decrease in speed limit (MPH) below the mean speed limit (~45 MPH) may increase the average total crashes per year by 2.01%. Having said this, if the speed limit on a particular rural 5T roadway segment is 40 MPH, the CMF for speed limit is found to be 1.100 indicating a total of 10% (=5*2.01) increase in total crashes per year associated with 5 units reduction in speed limit from the mean speed limit in the sample. The CMFs for different values of the number of driveways on rural 5T roadway segments are computed by using its parameter estimates of the original variable and its squared term as shown in Table 3.8 (for detail, please see the corresponding calibration spreadsheet). For a detailed procedure of computing CMFs for different variables used in TN-SPFs (Model 1D) for total crashes on the rural 5T roadway segments, please refer to the corresponding calibration spreadsheet (or Tables 3.9-3.12).

TABLE 3-8 DISTRIBUTION AND PARAMETER ESTIMATES OF KEY VARIABLES USED IN TN-SPFS FOR RURAL 5T ROADWAY SEGMENTS

Variable	***Mean***	***Min***	***Max***	***Coeff.***	***Exponent (Coeff.)***	***Percent change in average yearly crashes***
Width of 2WLTL (ft)	*12.08*	*10*	*14*	*-0.32*	*0.73*	*27.10*
Lane width (ft)	*11.9*	*10*	*12*	*-0.34*	*0.71*	*28.65*
Speed limit (MPH)	*46.44*	*25*	*65*	*-0.02*	*0.98*	*2.01*
Number of driveways	*12.49*	*0*	*95*	*0.01*	*1.01*	*1.48**
Number of driveways (squared term)	---	---	---	0	1	-0.04*

Note: * indicates that the parameter for both number of driveways and its squared term were considered while computing the CMF as documented in the corresponding spreadsheet. Note that if the number of driveways was 13 or lower, the CMF was found to be 1.

Table 3-9 CMF For Speed Limit (mph) on Rural 5T Roadway Segments

Speed limit (MPH)	***CMF***
25	1.402
30	1.301
35	1.201
40	1.100
45 (~46.44 mean value)	1
50	1
55	1
60	1
65	1

Table 3-10 CMF for Number of Driveways on Rural 5T Roadway Segments

Number of Driveways	***CMF***
1	1.000
2	1.000
3	1.000
4	1.000
5	1.000
6	1.000
7	1.000
8	1.000
9	1.000
10	1.000
11	1.000
12	1.000
13(~12.49 mean value)	1.000
14	1.0070
15	1.0132
16	1.0186
17	1.0232
18	1.0270
19	1.0300
20	1.0322
Note: CMF for # of driveways are shown for explication. In the calibration spreadsheet, the formula automatically calculates the CMF for the specific number of driveways (using the same parameters and procedure shown here).	

TABLE 3-11 CMF FOR LANE WIDTH ON RURAL 5T ROADWAY SEGMENTS

Lane Width (ft)	*CMF*
10	*1.573*
11	*1.287*
12 (~11.9 mean value)	*1*
13	*1*
14	*1*

TABLE 3-12 CMF FOR WIDTH OF SHARED LANE (2WLTL) ON RURAL 5T ROADWAY SEGMENTS

Width of 2WLTL (ft)	*CMF*
10	*1.542*
11	*1.271*
12 (~12.08 mean value)	*1*
13	*1*
14	*1*

3.2. Rural Four-lanes Divided (4D) Roadway Segments

3.2.1. Descriptive Statistics

Descriptive statistics of the key variables considered to develop SPFs for rural 4D roadway segments are provided in Table 3.13. Statistics reveal that the mean total crashes per year on rural 4D segments is 1.34. Note that there was at least one segment on which 0 or 18 total crashes per year have occurred (Table 3.13). To understand the distribution of rural 4D roadway segments based on total crashes per year, please see Figure 3.7. Statistics indicate that 0.37 crashes per year per rural 4D roadway segment have occurred. To see the distribution of rural 4D roadway segments based on average FI crashes per year, please see Figure 3.8. The mean AADT (in 1000s) per year is found to be 8.14 whereas, on at least one roadway segment, the AADT (in 1000s) per year is found to be 27.08 (maximum value in the data). The mean segment length in the data used for analysis is found to be 0.67 miles whereas at least one roadway segment was 0.1 miles long. For the descriptive statistics of other key variables including geometric variables, please refer to Table 3.13**.**

TABLE 3-13 DESCRIPTIVE STATISTICS OF KEY VARIABLES (RURAL 4D ROADWAY SEGMENTS)

Variables	***N***	***Mean***	***S.D.***	***Min***	***Max***
*Average of 5-years total crashes**	*271*	*1.34*	*2.00*	*0.00*	*18.00*
*Average of 5-years of FI crashes**	*271*	*0.37*	*0.74*	*0.00*	*6.00*
Average of 5-years AADT (in 1000s)	*271*	*8.14*	*4.59*	*0.49*	*27.08*
Segment length (mile)	*271*	*0.67*	*0.80*	*0.10*	*4.80*
Inner shoulder width (ft)	*271*	*3.78*	*1.24*	*0.00*	*8.00*
Speed limit (MPH)	*271*	*56.61*	*8.35*	*30.00*	*70.00*
Land use					
Indicator for land use (1 if commercial or mixed, 0 if residential)	*271*	*0.09*	*0.30*	*0.00*	*1.00*
Median width (ft)	*271*	*35.07*	*12.23*	*2.00*	*56.00*
Outer shoulder width (ft)	*271*	*10.16*	*1.60*	*2.00*	*12.00*
Presence of rumble strip along inner shoulder (1/0)	*271*	*0.67*	*0.47*	*0.00*	*1.00*
Presence of rumble strip along outer shoulder (1/0)	*271*	*0.75*	*0.43*	*0.00*	*1.00*
Lane width (ft)	*271*	*11.99*	*0.12*	*11.00*	*12.00*

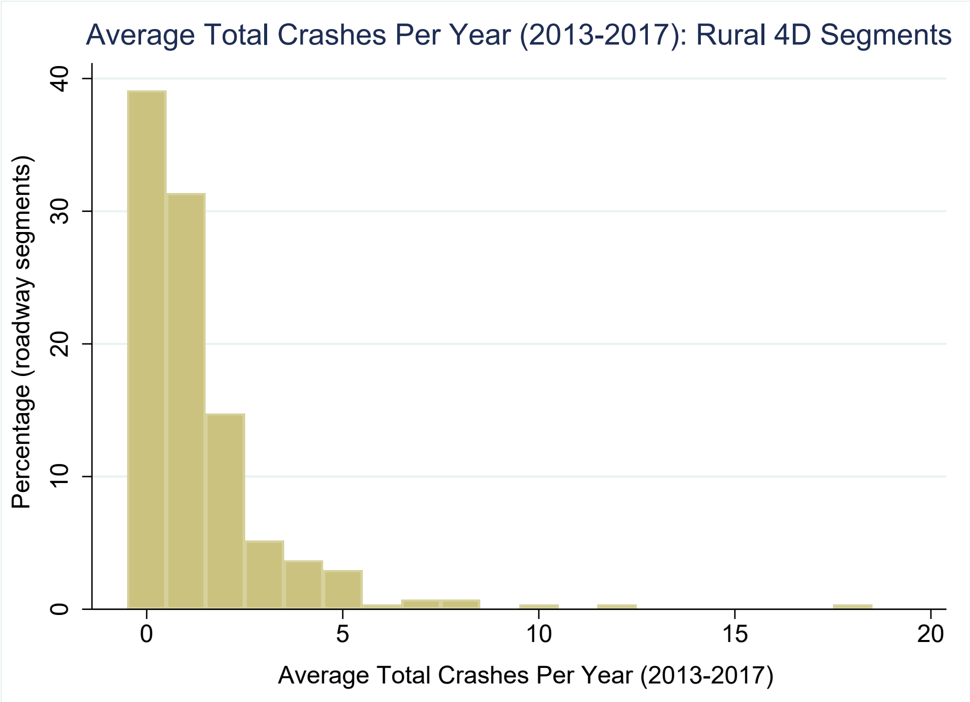


Figure 3.7 Distribution of Rural 4D Roadway Segments based on Average Total Crashes per Year

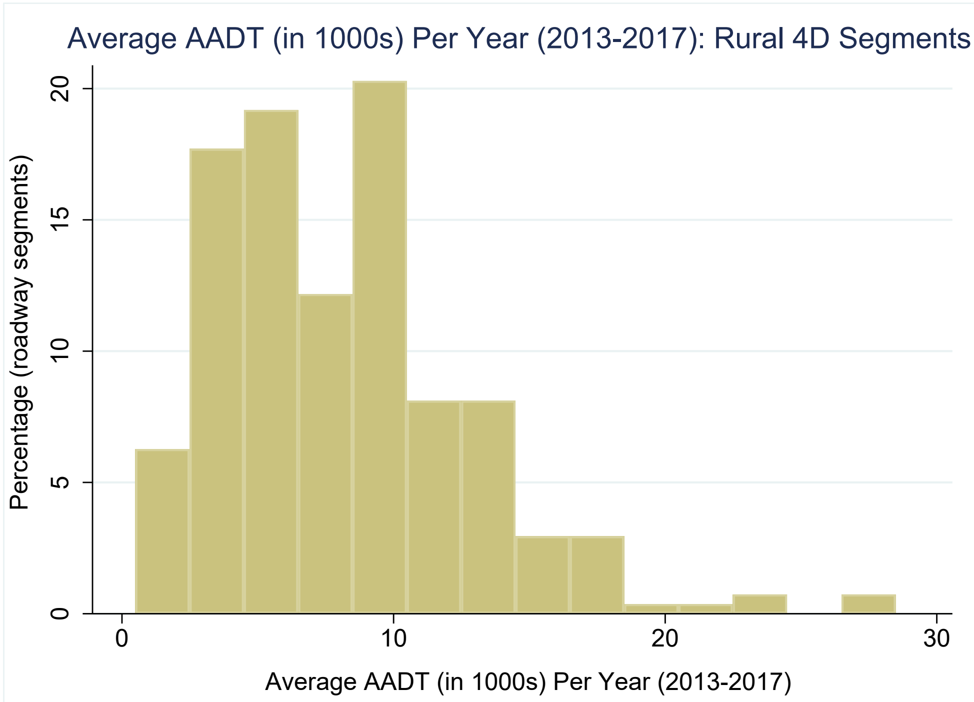


Figure 3.8 Distribution of Rural 4D Roadway Segments based on Average AADT per Year

3.2.2. Modeling Results

3.2.2.1. TN-Specific SPFs for Total Crashes

Following the base-case HSM (2010) functional form, first the base-case TN-SPFs were estimated for the average of 5-years total crashes on rural 4D roadway segments. The estimation results are provided in Table 3.14 which reveal that the NB model (Model 3B) does not show any significant improvement compared to the Poisson model (Model 3A)- 3B was selected as the final model. Based on the results of Model 3B, the values of "a" and "b" are found to be -3.0779 and 0.4295 respectively. The HSM default values of "a" and "b" parameters for total crashes on rural 4D roadway segments are -9.025 and 1.049 [1]. This indicates a substantial difference in the values of the two estimated parameters compare with the HSM default values.

TABLE 3-14 BASE-CASE TN-SPECIFIC SPF FOR TOTAL CRASHES ON RURAL 4D ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 3A)***			***NB (Model 3B)***		
	Coef.	***Std. Err.***	***t-stat***	***Coef.***	***Std. Err.***	***t-stat***
Average 5-years AADT: ln form	0.4514	0.0874	5.16	0.4295	0.0969	4.43
Segment length (mile): Exposure	1	---	---	1	---	---
Constant	-3.2947	0.7819	-4.21	-3.0779	0.8668	-3.55
Over-dispersion parameter (alpha)	---	---	---	0.1094	0.0675	1.62
Test for alpha significantly different than 0						
Chi-square test statistics	---			4.37		
Prob. (Chi-square)	---			0.018		
Summary Statistics						
Log-likelihood (Convergence)	-369.6043			-367.4193		
AIC	743.2086			740.8387		
Degrees of freedom	2			3		
Sample Size (N)	271			271		

The estimation results of the enhanced SPFs for total crashes on rural 4D roadway segments are summarized in Table 3.15. The overdispersion parameter in the NB model is found to be partially significant; for consistency with the HSM (2010) functional form, Model 3D was selected as the final model. Some of the geometric variables (e.g., lane width, median width, and outer (right) shoulder width) did not show statistical significance; however, the research team decided to keep all these important variables in the final model to be consistent with HSM [1]. Referring to the results, it was found that total crashes per year on rural 4D roadway segments reduce if the inner (left) shoulders on rural 4D roadway segments are wider. Similarly, total crashes per year are expected to be lower on the rural 4D segments with a higher speed limit. As expected, the total crash frequency increases on rural 4D roadway segments which pass through commercial or mixed land use compared to residential. Based on the results, the presence of a rumble strip along the inner shoulder may reduce overall crashes on rural 4D roadway segments. While the parameter signs of the median width (ft) and lane width (ft) were found to be negative which were expected, these variables did not show a statistically significant relationship with the crash frequency (Table 3.15).

TABLE 3-15 ENHANCED TN-SPECIFIC SPFS FOR TOTAL CRASHES ON RURAL 4D ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 3C)***				***NB (Model 3D)***			
	Coef.	***t-stat***	***MEs***	***IRR***	***Coef.***	***t-stat***	***MEs***	***IRR***
Average 5-years AADT (in 1000s): ln form	0.3736	4.26	0.5004	1.4529	0.3505	3.61	0.4736	1.4198
Lane width (ft)	-0.2366	-0.50	-0.3169	0.7893	-0.2075	-0.41	-0.2804	0.8126
Outer (right side) shoulder width (ft)	0.0133	0.37	0.0178	1.0134	0.0084	0.21	0.0113	1.0084
Median width (ft)	-0.0016	-0.27	-0.0021	0.9984	-0.0022	-0.34	-0.0029	0.9978
Inner (left side) shoulder width (ft)	-0.1266	-2.24	-0.1696	0.8811	-0.1209	-1.97	-0.1633	0.8861
Speed limit (MPH)	-0.0142	-1.76	-0.0190	0.9859	-0.0162	-1.82	-0.0220	0.9839
Presence of Rumble Strip along Inner Shoulder (0/1)	-0.2462	-2.12	-0.3298	0.7818	-0.2700	-2.04	-0.3649	0.7634
Land use (0 if residential, 1 if commercial or mixed)	0.4127	2.22	0.5528	1.5109	0.4907	2.29	0.6630	1.6334
Constant	1.6214	0.28	---	---	1.6654	0.27	---	---
Segment length (mile): Exposure	1	---	---	1	1	---	---	1
Over-dispersion parameter (alpha)	---	---	---	---	0.1027	1.7539	---	0.1027
Test for alpha significantly different than 0								
Chi-square test statistics	---				5.14			
Prob. (Chi-square)	---				0.012			
Summary Statistics								
Log-likelihood (Convergence)	-352.9095				-350.3399			
AIC	723.8191				720.6797			
Degrees of freedom	9				10			
Sample Size (N)	271				271			

Note: The incidence rate ratios (IRR) for a specific explanatory variable in the enhanced TN-SPFs indicate the change (increase/decrease) in the average total crashes per year with a unit increase in a specific variable when all other variables are held at constant values in the model. If the value of IRR is less than or greater than 1, it shows the corresponding reduction or increase in total crashes per year by the time it is less than or greater than 1, respectively.

TABLE 3-16 MEAN ABSOLUTE ERROR: COMPARING MODELS FOR TOTAL CRASHES ON RURAL 4D SEGMENTS

Model	*Mean Absolute Error*					*Comparison: Mean MAE (%)*	*Model Summary*		
	Obs.	*Mean*	*S.D.*	*Min.*	*Max.*		*LL (Convergence)*	*AIC*	*Deg. of freedom*
HSM Enhanced-SPF	*271*	0.854	1.142	0.034	8.919	Base	---	---	---
Model 3A	*271*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Model 3B	*271*	0.88	0.99	0.005	6.51	3.49	-367.41	740.83	3
Model 3C	*271*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Model 3D	*271*	0.89	1.06	0.001	9.55	4.45	-350.33	720.67	10

Note: Model 3A and 3C indicate base-case and Enhanced TN-Specific Poisson models. Since their negative binomial counterparts (Model 3B and Model 3D respectively) were selected, the MAE for only Model 3B and 3D are shown. For comparing the MAE with the HSM Enhanced-SPF, the formula in equation (10) below was used.

$$\text{MAE } (\%) = \frac{\text{(Mean MAE of Model 3A/3B/3C/3D – Mean MAE of HSM Enhanced-SPF)}}{\text{(Mean MAE of HSM Enhanced-SPF)}} * 100\% \qquad (10)$$

3.2.2.2. TN-Specific SPFs for Fatal and Injury (FI) Crashes: Rural 4D Roadway Segments

First, researchers estimate base-case TN-SPFs for FI crashes on rural 4D roadway segments. Table 3.17 shows the NB model (Model 4B) as the final model for computing the values of regression parameters. Based on the TN-SPFs, the values of "a" and "b" are found to be -4.6137 and 0.4559, respectively, compared with the HSM (2010) default values -8.505 and 0.874 [1]. For FI crashes on rural 4D roadway segments, the value of "c" was found to be 15.641, which is substantially higher than the corresponding default HSM value of 1.740 [1].

To understand how key roadway variables, speed limit (MPH), and land use relate to average FI crashes per year on rural 4D roadway segments, the enhanced TN-SPFs for FI crashes were estimated. For consistency with the enhanced TN-SPFs for total crashes on rural 4D roadway segments, the same set of explanatory variables were used. The negative binomial was selected as the final model for consistency with HSM (2010) functional form (Table 3.18). In the final model, only AADT, inner (left) shoulder width, and indicator for commercial land use showed statistically significant relationships with average FI crashes per year on rural 4D segments (Table 3.18).

TABLE 3-17 BASE-CASE TN-SPECIFIC SPF FOR FI CRASHES ON RURAL 4D ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 4A)***			***NB (Model 4B)***		
	Coef.	***Std. Err.***	***t-stat***	***Coef.***	***Std. Err.***	***t-stat***
Average of 5-years AADT: ln form	0.4559	0.16	2.75	0.4559	0.16	2.75
Segment length (miles): Exposure	1	---	---	1	---	---
Constant	-4.6137	1.48	-3.11	-4.6137	1.48	-3.11
Over-dispersion parameter (alpha)	---	---	---	4.13e-07	0.00	0.00
Test for alpha significantly different than 0						
Chi-square test statistics	---			0.00		
Prob. (Chi-square)	---			0.500		
Summary Statistics						
Log-likelihood (Convergence)	-184.0819			-184.0819		
AIC	372.1639			374.1639		
Degrees of freedom	2			3		
Sample Size (N)	271			271		

TABLE 3-18 ENHANCED TN-SPECIFIC SPF FOR FI CRASHES ON RURAL 4D ROADWAY SEGMENTS

Explanatory Variables	Model 4C				Model 4D			
	Coef.	***t-stat***	***MEs***	***IRR***	***Coef.***	***t-stat***	***MEs***	***IRR***
Average 5-years AADT (in 1000s): In form	0.3492	2.12	0.1301	1.4180	0.3492	2.12	0.1301	1.4180
Lane width (ft)	0.4073	0.39	0.1518	1.5028	0.4073	0.39	0.1518	1.5028
Outer (right side) shoulder width (ft)	0.0645	1.01	0.0240	1.0667	0.0645	1.01	0.0240	1.0667
Median width (ft)	-0.0087	-0.81	-0.0032	0.9913	-0.0087	-0.81	-0.0032	0.9913
Inner (left side) shoulder width (ft)	-0.2024	-2.02	-0.0754	0.8167	-0.2024	-2.02	-0.0754	0.8167
Speed limit (MPH)	-0.0111	-0.72	-0.0041	0.9888	-0.0111	-0.72	-0.0041	0.9888
Presence of Rumble Strip along Inner Shoulder (0/1)	-0.2604	-1.18	-0.0970	0.7706	-0.2604	-1.18	-0.0970	0.7706
Land use (0= residential, 1= commercial or mixed)	0.8280	2.62	0.3086	2.2888	0.8280	2.62	0.3086	2.2889
Constant	-7.3042	-0.58	---	---	-7.3042	-0.58	---	---
Segment length (mile): Exposure	1	---	---	1	1	---	---	1
Over-dispersion parameter (alpha)	---	---	---	---	1.21e-07	.0009	---	---
Test for alpha significantly different than 0								
Chi-square test statistics	---				0.0e+00			
Prob. (Chi-square)	---				0.500			
Summary Statistics								
Log-likelihood (Convergence)	-173.8913				-173.8913			
AIC	365.7826				367.7826			
Degrees of freedom	9				10			
Sample Size (N)	271				271			

3.2.2.3. CMFs for Key Factors Using TN-SPFs for Total Crashes on Rural 4D Segments

Before discussing the CMFs, the distribution and parameter estimates (Model 4D) of key explanatory variables used in the final TN-SPF for total crashes on rural 4D segments are provided (Table 3.19). The findings indicate that with a unit decrease in lane width (ft) below the mean value (12 ft), the average total crashes per year increase by 18.73%. Note that for lane width equal to or greater than 12 ft, the CMF is 1.00. Similarly, with a unit decrease in median width (ft) and inner shoulder width (ft), the total crashes per year on rural 4D roadway segments are expected to increase by 0.2198% and 11.39%, respectively.

Referring to the speed limit (MPH), the results suggest that a unit decrease in speed limit below the average speed limit (~55 MPH) is associated with increases in total crashes per year of 1.61%. Note that if the speed limit (MPH) is equal to or greater than the mean value of the speed limit in the sample, the CMF is 1.00. Importantly, the parameter sign for right shoulder width in TN-SPF is consistent with the default value in the FHWA calibration spreadsheet; nevertheless, this result seems unintuitive suggesting that average total crashes per year may increase with an increase in width of the right shoulder.

For indicator variables included in the final TN-SPF for rural 4D roadway segments, the CMFs can be directly computed as shown in Table 3.19. For instance, based on the findings, if rumble strips are present along the inner shoulder on rural 4D roadway segments, then the CMF would be 0.7634 which indicates that total crashes may reduce by 23.64% (=1-0.7634).

For land use, the results indicate that compared to residential land use, if the rural 4D roadway segment passes through areas with commercial or mixed commercial and industrial use, the value of the CMF would be 1.6335 which indicates that the total crashes per year are higher by 63.35%. Using the abovementioned procedure, the CMFs for different values of specific factors used in the final TN-SPFs for rural 4D roadway segments are computed as shown in Table 3.20-3.26.

TABLE 3-19 DISTRIBUTION AND PARAMETER ESTIMATES OF KEY VARIABLES USED IN TN-SPFS FOR RURAL 4D SEGMENTS

Variable	*Mean*	*Min*	*Max*	*Coeff.*	*Exponent (Coeff.)*	*Percent change in average crashes*
Lane width (ft)	11.99	11	12	-0.2075	0.8126	18.7387
Outer (right side) shoulder width (ft)	10.16	2	12	0.0084	1.0084	-0.8435*
Median width (ft)	42.63	2	65	-0.0022	0.9978	0.2198
Inner (left side) shoulder width (ft)	3.78	0	8	-0.1209	0.8861	11.3877
Speed limit (MPH)	56.61	30	70	-0.0162	0.9839	1.6069
Presence of Rumble Strip along Inner Shoulder (0/1)	0.67	0	1	-0.2700	0.7634	0.7634
Land use (0 if residential, 1 if commercial or mixed)	0.10	0	1	0.4907	1.6335	1.6335

Note: * indicate that with a unit decrease in outer shoulder width, the average total crashes per year reduces by 0.8435%.

TABLE 3-20 CMF FOR LANE WIDTH ON RURAL 4D MULTILANE ROADWAY SEGMENTS

Lane width (ft)	CMF
11	*1.1874*
12 (Base)	*1*

TABLE 3-21 CMF FOR MEDIAN WIDTH (FT) ON RURAL 4D ROADWAY SEGMENTS

Median Width (ft)	CMF	Median Width (ft)	CMF
2	*1.0945*	*24*	*1.0461*
3	*1.0923*	*25*	*1.0440*
4	*1.0901*	*26*	*1.0418*
5	*1.0879*	*27*	*1.0396*
6	*1.0857*	*28*	*1.0374*
7	*1.0835*	*29*	*1.0352*
8	*1.0813*	*30*	*1.0330*
9	*1.0791*	*31*	*1.0308*
10	*1.0769*	*32*	*1.0286*
11	*1.0747*	*33*	*1.0264*
12	*1.0725*	*34*	*1.0242*
13	*1.0703*	*35*	*1.0220*
14	*1.0681*	*36*	*1.0198*
15	*1.0659*	*37*	*1.0176*
16	*1.0637*	*38*	*1.0154*
17	*1.0615*	*39*	*1.0132*
18	*1.0593*	*40*	*1.0110*
19	*1.0571*	*41*	*1.0088*
20	*1.0549*	*42*	*1.0066*
21	*1.0527*	*43*	*1.0044*
22	*1.0505*	*44*	*1.0022*
23	*1.0483*	*45 or above*	*1*

TABLE 3-22 CMF FOR OUTER (RIGHT SIDE) SHOULDER WIDTH (FT) ON RURAL 4D ROADWAY SEGMENTS

Width of right shoulder	***CMF***
2	*1*
3	*1*
4	*1*
5	*1*
6	*1*
7	*1*
8	*1*
9	*1*
10 (Base)	*1*
11	*1.0084*
12	*1.0169*

TABLE 3-23 CMF FOR INNER (LEFT) SHOULDER WIDTH ON RURAL 4D ROADWAY SEGMENTS

Inner Shoulder Width (ft)	***CMF***
0	*1.456*
1	*1.342*
2	*1.228*
3	*1.114*
4 (Mean value = 3.78 ft)	*1.000*
5	*1.000*
6	*1.000*
8	*1.000*

TABLE 3-24 CMF FOR SPEED LIMIT (MPH) ON RURAL 4D ROADWAY SEGMENTS

Speed limit (MPH)	***CMF***
30	*1.402*
35	*1.321*
40	*1.241*
45	*1.161*
50	*1.080*
55 (~56.6 MPH)	*1.000*
60	*1.000*
65	*1.000*
70	*1.000*

TABLE 3-25 CMF FOR RUMBLE STRIP ALONG INNER SHOULDER ON RURAL 4D ROADWAY SEGMENTS

Rumble Strip along inner shoulder	***CMF***
Yes	0.763
No	1

Table 3-26 CMF FOR LAND USE (0= RESIDENTIAL, 1= COMMERCIAL OR MIXED) ON RURAL 4D ROADWAY SEGMENTS

Land Use	***CMF***
Urban or Mixed	1.607
Residential	1

3.2.2.4. Comparison of Default FHWA with TN-Specific Calibration Spreadsheet (Rural 4D Roadways)

To compare the performance of the TN-Specific calibration spreadsheet with the default FHWA calibration spreadsheet for rural 4D roadway segments, the research team randomly selected four roadway segments of the rural 4D roadway type to see how well the two calibration spreadsheets could predict the number of total crashes per year compared to the observed crash counts per year. For segment #1 and segment # 2 which have almost base conditions, the FHWA (HSM-SPFs) calibration shows better prediction (Table 3.27). However, for segment #3 and segment #4 on which the conditions deviate from the base conditions, the new TN-specific calibration shows better prediction (Table 3.27). For details, please see the corresponding CMFs provided in the earlier section or refer to the corresponding calibration spreadsheet).

Table 3-27 COMPARING PERFORMANCE OF CALIBRATION WITH OUR NEW TN-SPECIFIC CALIBRATION (RURAL 4D)

Segment Details	*Randomly Selected Segments*			
	Segment 1	***Segment 2***	***Segment 3***	***Segment 4***
County (Route)	Weakley (SR043)	Rhea (SR029)	Obion (SR003)	Campbell (SR009)
BLM & ELM	0.43 & 0.98	11.01 & 13.2	10.87 & 11.48	1.31 & 1.748
Special Case (Co-Sequence)	0-None (1)	0-None (1)	0-None (1)	0-None (1)
Segment Attributes				
Average AADT Per Year	6462	14194	12728	3554
Segment Length	0.55	2.19	0.61	0.438
Lane Width	12	12	12	12
Right (outer) shoulder width	11	11	4	8
Median width	46	47	34	30
Streetlight (1/0)	No	No	No	Yes
Speed limit	60	55	45	45
Commercial land use (1/0)	Residential	Residential	Commercial	Commercial
Left (inner) shoulder width	4	4	3	2
Presence of rumble strip along inner shoulder	Yes	Yes	No	No
Observed Crashes Per Year	0	4	2	3
Predicted Crashes Per Year				
*HSM Base SPF**	0.658	5.977	1.485	0.280
*HSM Enhanced SPF (After Applying CMFs)**	0.638	5.798	1.619	0.255
*TN-Specific Base SPF***	1.097	6.124	1.628	0.676
*TN Enhanced SPFs (CMFs used from TN-Enhanced Models)***	1.106	6.176	2.156	0.995

Note: * indicates that FHWA default calibration procedure & CMFs are used. ** indicates that FHWA calibration spreadsheet was modified using TN-SPFs. See excel spreadsheet.

3.3. Rural Four-lanes Undivided (4U) Roadway Segments

3.3.1. Descriptive Statistics

Descriptive statistics of the key variables used in TN-Specific SPFs for total crashes on rural 4U roadway segments are provided in Table 3.28. According to statistics, the average of total crashes per year is found to be 2.56 where there was at least one roadway segment on which 0 or 39 crashes have occurred per year. For FI crashes, statistics reveal that 0.59 FI crashes have occurred per year per rural 4U roadway segment. The mean of average 5-years AADT (in 1000s) is found to be 6.94 whereas the maximum value was found to be 31.64 (31,640). To understand the distribution of rural 4U roadway segments based on the average of 5-years total crashes, average 5-years FI crashes, and the average of 5-years AADT, please see Figure 3.9, Figure 3.10, and Figure 3.11 respectively. Statistics reveal that the mean segment length of rural 4U roadway segments is found to be 0.42 miles; however, there was at least one roadway segment that has a length of 2.50 miles. For the distribution of other key explanatory variables on rural 4D roadway segments, please see Table 3.28.

TABLE 3-28 DESCRIPTIVE STATISTICS OF KEY VARIABLES (RURAL 4U ROADWAY SEGMENTS)

Variables	N	Mean	S.D.	Min	Max
Average 5-years total crashes	*81*	*2.56*	*5.60*	*0*	*39*
Average 5-years FI crashes	*81*	*0.59*	*1.25*	*0*	*6*
Average 5-years AADT (in 1000s)	*81*	*6.94*	*4.30*	*0.63*	*31.64*
Segment length (miles)	*81*	*0.42*	*0.54*	*0.10*	*2.50*
Presence of rumble strip along outer shoulder (1/0)	*81*	*0.27*	*0.45*	*0*	*1*
Speed limit (MPH)	*81*	*43.77*	*9.60*	*25*	*70*
Lane width (ft)	*81*	*11.59*	*0.63*	*10*	*12*
Land use					
Indicator for land use (1 if commercial or mixed, 0 if residential)	*81*	*0.44*	*0.50*	*0*	*1*
Lane width (ft)	*81*	*11.59*	*0.63*	*10*	*12*
Outer shoulder width (ft)	*81*	*4.49*	*3.41*	*0*	*12*

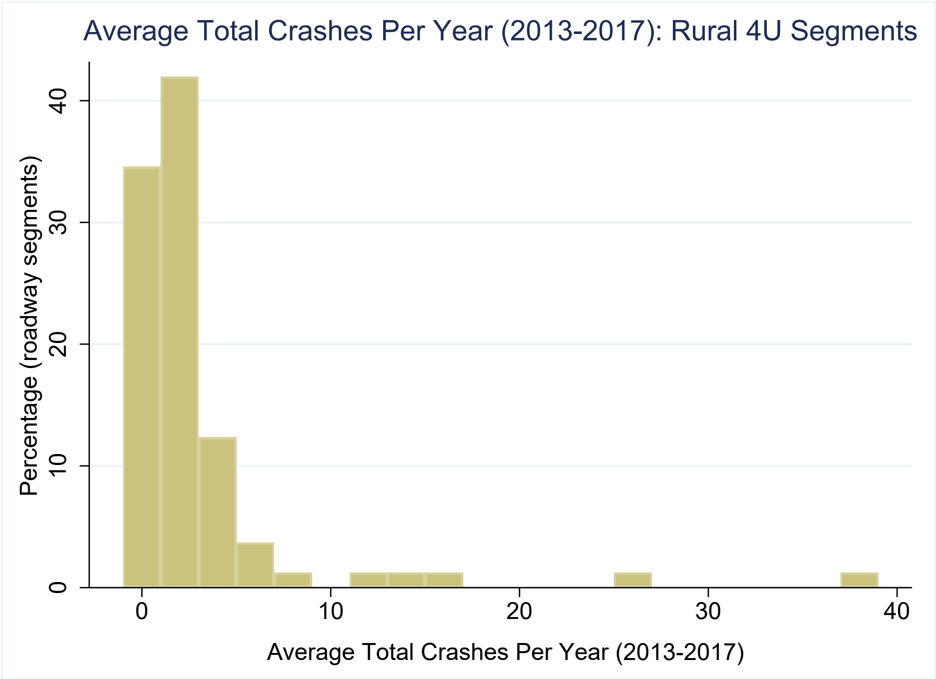


Figure 3.9 Distribution of Rural 4U Roadway Segments based on Average Total Crashes per Year

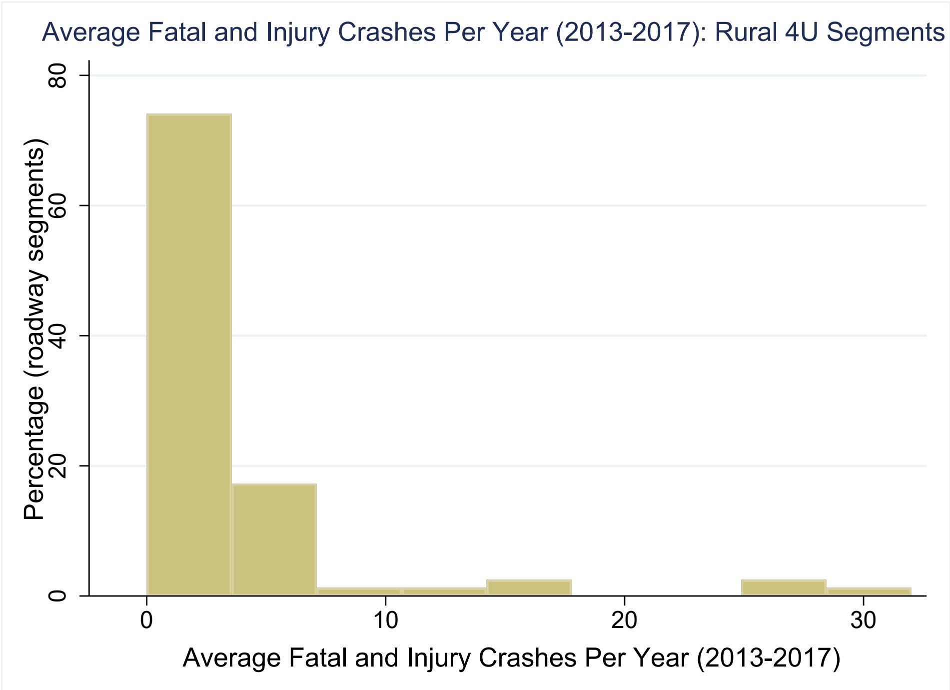


Figure 3.10 **Distribution of Rural 4U Roadway Segments based on Average FI Crashes per Year**

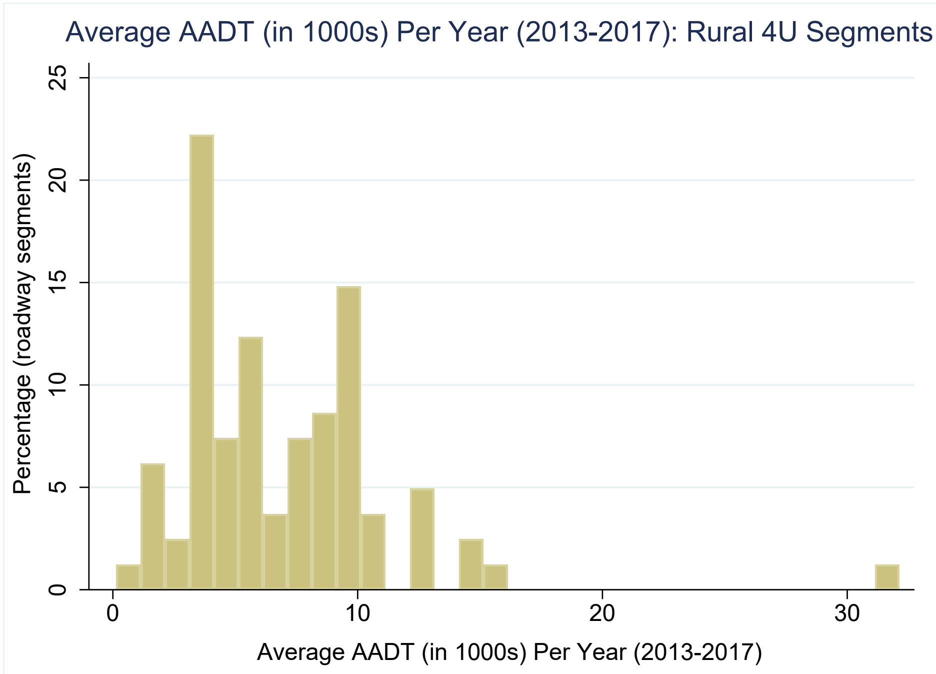


Figure 3.11 Distribution of Rural 4U Roadway Segments based on Average AADT per Year

3.3.2. Modeling Results

3.3.2.1. TN-Specific SPFs for Total Crashes

First, base-case TN-specific SPFs for total crashes on rural 4U roadway segments were developed by applying both Poisson and NB regressions (Table 3.29). Compared to Poisson (Model 5A), the negative binomial (Model 5B) showed significant improvement based on log-likelihood at convergence and AIC values. Also, the statistically significant overdispersion parameter suggests that NB (Model 5B) could be preferred instead of the Poisson (Model 5A). Based on the best model (Model 5B), the TN-Specific values of "a" and "b" parameters for total crashes on rural 4U roadway segments are found to be -7.9503 and 1.0919, respectively. Note that the TN-Specific values of "a" and "b" are closed to their default values in the HSM (2010) for the total crashes on rural 4U roadway segments which are -9.653 and 1.176, respectively [1]. Furthermore, the TN-Specific value of the "c" parameter for total crashes on the rural 4U roadway segment is found to be 2.0903 which is significantly different than the HSM default value of 1.675 [1].

TABLE 3-29 BASE-CASE TN-SPECIFIC SPF FOR TOTAL CRASHES ON RURAL 4U ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 5A)***			***NB (Model 5B)***		
	Coef.	***Std. Err.***	***t-stat***	***Coef.***	***Std. Err.***	***t-stat***
Average 5-years AADT: ln form	1.3797	0.1222	11.29	1.0919	0.2211	4.94
Segment length (mile): Exposure	1	---	---	1	---	---
Constant	-10.6938	1.1367	-9.41	-7.9503	1.9869	-4.00
Over-dispersion parameter (alpha)	---	---	---	0.4614	0.1567	2.94
Test for alpha significantly different than 0						
Chi-square test statistics	---			54.15		
Prob. (Chi-square)	---			0.000		
Summary Statistics						
Log-likelihood (Convergence)	-161.9846			-134.9073		
AIC	327.9692			275.8145		
Degrees of freedom	2			3		
Sample Size (N)	81			81		

Referring to the estimation results for the enhanced TN-SPFs, our analysis indicates that the NB model (Model 5D) did not show any significant improvement compared to Poisson (Model 5C). For consistency with HSM (2010), Model 4D (NB) was selected as the final model. Note that the indicator for the presence of streetlights shows intuitive parameter signs but is found to be statistically insignificant (Table 3.30). Given the intuitive parameter sign, an indicator for streetlights was included in the final model to compute the corresponding CMF as discussed in the subsequent sections. Referring to the estimation results of the final model (model 4D), total crashes per year were found to be lower on rural 4U roadway segments with wider lanes, higher speed limit, wider outer (right) shoulder, and those with streetlights (Table 3.30). Notably, the expected crash frequency per year is likely to be higher on rural 4U roadway segments which pass through areas with commercial or mixed land use compared to residential land use.

TABLE 3-30 ENHANCED TN-SPECIFIC SPFS FOR TOTAL CRASHES ON RURAL 4U ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 5C)***				***NB (Model 5D)***			
	Coef.	***t-stat***	***MEs***	***IRR***	***Coef.***	***t-stat***	***MEs***	***IRR***
Average 5-years AADT (in 1000s): ln form	0.5281	3.87	1.3496	1.6957	0.5281	3.87	1.3496	1.6957
Lane width (ft)	-0.4466	-3.28	-1.1413	0.6398	-0.4466	-3.28	-1.1413	0.6398
Speed limit (MPH)	-0.0250	-2.48	-0.0640	0.9753	-0.0250	-2.48	-0.0640	0.9753
Land use (0 if residential, 1 if commercial or mixed)	0.8045	2.98	2.0561	2.2357	0.8045	2.98	2.0561	2.2357
Outer (right side) shoulder width (ft)	-0.0626	-1.94	-0.1601	0.9393	-0.0626	-1.94	-0.1601	0.9393
Presence of streetlight (0/1)	-0.3670	-1.16	-0.9379	0.6928	-0.3670	-1.16	-0.9379	0.6928
Constant	3.2719	1.47	---	---	3.2719	1.47	---	---
Segment length (mile): Exposure	1	---	---	1	1	---	---	1
Over-dispersion parameter (alpha)	---	---	---	---	$2.94*10^{-8}$	0.001	---	---
Test for alpha significantly different than 0								
Chi-square test statistics	---				0.0e+00			
Prob. (Chi-square)	---				0.500			
Summary Statistics								
Log-likelihood (Convergence)	-106.7703				-106.7703			
AIC	227.5407				229.5407			
Degrees of freedom	7				8			
Sample Size (N)	81				81			

Referring to the comparison of TN-SPFs and enhanced HSM SPF for total crashes on rural 4U roadway segments, the research team compute and compare the MAE as shown in Table 3.31. The findings indicate that base-case TN-SPF (Model 5B) shows weaker prediction accuracy compared to the HSM enhanced SPF (Table 3.31). However, if enhanced TN-SPF (Model 5D) is used, the mean MAE reduces by 53.87% compared to the HSM enhanced SPF thus suggesting the superior performance of enhanced TN-SPFs compared to both the HSM Enhanced-SPF and base-case TN-SPF.

TABLE 3-31 MEAN ABSOLUTE ERROR: COMPARING MODELS FOR TOTAL CRASHES ON RURAL 4U ROADWAY SEGMENTS

Model (SPF)	*Obs.*	*Mean*	*S.D.*	*Min*	*Max*	*Comparison: Mean MAE (%)**
HSM Enhanced-SPF	81	1.7137	4.1701	0.0413	29.3170	Base
Model 5A	81	N/A	N/A	N/A	N/A	N/A
Model 5B	81	2.0382	3.7240	0.0677	19.5034	18.9337
Model 5C	81	N/A	N/A	N/A	N/A	N/A
Model 5D	81	0.7904	0.6522	0.0263	2.8045	-53.8755

Note: Model 5A and 5C indicate base-case and Enhanced TN-Specific Poisson models. Since their negative binomial counterparts (Model 5B and Model 5D respectively) were selected for consistency with HSM (2010); therefore, the MAE of only Model 5B and 5D are shown. For comparing the MAE with the HSM Enhanced-SPF, the formula in equation 11 below was used.

$$\mathrm{MAE}\ (\%) = \frac{(\text{"Mean MAE of Model 5A/5B/5C/5D"} - \text{"Mean MAE of HSM Enhanced-SPF"})}{(\text{"Mean MAE of HSM Enhanced-SPF"})} \times 100 \qquad (11)$$

3.3.2.2. TN-Specific SPFs for Fatal and Injury (FI) Crashes: Rural 4U Roadway Segments

According to the estimation results for FI crashes per year on rural 4U roadway segments, the overdispersion parameter in base-case TN-SPF is statistically insignificant due to which the negative binomial model simply reduces to the Poisson model (Table 3.32). This can also be noticed in the enhanced TN-SPF for average FI crashes on rural 4U roadway segments where the Poisson and negative binomial model show similar estimation results (Table 3.33). For consistency with HSM (2010), the negative binomial model was selected as the final model in both base-case and enhanced case TN-SPF for FI crashes (see Table 3.32 and Table 3.33, respectively). Based on the base-case TN-SPFs, the values of "a" and "b" are found to be -12.0980 and 1.3735, respectively (Table 3.32). The HSM default values of "a" and “b" for FI crashes on rural 4U roadway segments are found to be -8.577 and 0.938, respectively [1]. Using the TN-SPF, the value of "c" was found to be 18.4399 for FI crashes on rural 4U roadway segments which is significantly higher than the corresponding HSM default value of 2.003 [1]. Referring to the enhanced TN-SPF for FI crashes, only AADT and outer (right) shoulder width showed a statistically significant relationship with average FI crashes per year; for consistency with the enhanced TN-SPFs for total crashes on rural 4U roadway segments, all the relevant roadway variables were kept in the final models (Table 3.33).

TABLE 3-32 BASE-CASE TN-SPECIFIC SPFS FOR FI CRASHES ON RURAL 4U ROADWAY SEGMENTS

Explanatory Variables	***Poisson (Model 6A)***			***NB (Model 6B)***		
	Coef.	***Std. Err.***	***t-stat***	***Coef.***	***Std. Err.***	***t-stat***
Average of 5-years AADT: ln form	1.3735	0.2538	5.41	1.3735	0.2538	5.41
Segment length (miles): Exposure	1	---	---	1	---	---
Constant	-12.0976	2.3603	-5.13	-12.0980	2.3603	-5.13
Over-dispersion parameter (alpha)	---	---	---	3.66e-08	0.0001	0.0005
Test for alpha significantly different than 0						
Chi-square test statistics	---			0.00		
Prob. (Chi-square)	---			0.500		
Summary Statistics						
Log-likelihood (Convergence)	-52.6517			-52.65179		
AIC	109.3036			111.3036		
Degrees of freedom	2			3		
Sample Size (N)	81			81		

TABLE 3-33 ENHANCED TN-SPECIFIC SPF FOR FI CRASHES ON RURAL 4U ROADWAY SEGMENTS

Explanatory Variables	Poisson (Model 6C)				NB (Model 6D)			
	Coef.	***t-stat***	***MEs***	***IRR***	***Coef.***	***t-stat***	***MEs***	***IRR***
Average 5-years AADT (in 1000s): ln form	1.0069	2.97	0.5967	2.7372	1.0069	2.97	0.5967	2.7372
Lane width (ft)	0.0315	0.10	0.0187	1.0320	0.0315	0.10	0.0187	1.0320
Speed limit (MPH)	-0.0085	-0.46	-0.0050	0.9914	-0.0085	-0.46	-0.0050	0.9914
Land use (0 if residential, 1 if commercial or mixed)	0.6739	1.15	0.3993	1.9620	0.6739	1.15	0.3993	1.9620
Outer (right side) shoulder width (ft)	-0.1201	-1.78	-0.0711	0.8868	-0.1201	-1.78	-0.0711	0.8868
Presence of streetlight (0/1)	-0.6167	-0.96	-0.3654	0.5396	-0.6167	-0.96	-0.3654	0.5396
Constant	-8.191	-1.45	---	---	-8.191	-1.45	---	---
Segment length (mile): Exposure	1	---	---	1	1	---	---	1
Over-dispersion parameter (alpha)	---	---	---	---	4.77e-08	.0005	---	---
Test for alpha significantly different than 0								
Chi-square test statistics	---				7.5e-07			
Prob. (Chi-square)	---				0.500			
Summary Statistics								
Log-likelihood (Convergence)	-48.85425				-48.85425			
AIC	111.7085				113.7085			
Degrees of freedom	7				8			
Sample Size (N)	81				81			

3.3.2.3. CMFs for Key Factors Using TN-SPFs for Total Crashes on Rural 4U Segments

Before discussing the CMFs for the key explanatory variables, shown are the distribution and parameter estimates (based on Model 6D) of various variables which were used in the final TN-SPFs for the total crashes (average total crashes per year) on rural 4U roadway segments – for details, please see Table 3.34. The CMFs for the key factors on rural 4U roadway segments are computed using the parameter estimates from the TN-SPF for total crashes on rural 4U roadway segments. The results indicate that with a unit decrease in lane width (ft) below the mean value (~12 ft), the average total crashes per year increase by 36.02%. Having said this, the lane width on a particular rural 4U roadway segment is 11 ft and 10 ft, the CMFs would be 1.36 and 1.72 respectively (Table 3.35). A similar procedure is used for computing CMFs for other continuous variables including speed limit (MPH) and outer (right) shoulder width (ft). For the speed limit (MPH) on rural 4U roadway segments, the corresponding parameter estimate suggests that with a unit decrease in speed limit below the mean value (~45 MPH), the total crashes per year may increase by 2.47%. Similarly, with a unit decrease in outer shoulder width (ft) below the mean value (~5 ft), the total crashes per year are expected to increase by 6.07%. Given the nature of the indicator (dummy) variable, the CMF could be computed directly as shown in Table 3.34. For instance, our modeling results indicate that compared to no streetlights if streetlights are present on the rural 4U roadway segments, the CMF is found to be 0.6928 which indicates that the average total crashes per year may reduce by 32.72%. For commercial or mixed land use (compared to residential land use), the CMF is found to be 2.2356. To see the CMFs for different values of various important variables for average total crashes per year on rural 4U roadway segments, please refer to Table 3.35-3.39.

TABLE 3-34 DISTRIBUTION AND PARAMETER ESTIMATES OF KEY VARIABLES USED IN TN-SPFS FOR RURAL 4U ROADWAY SEGMENTS

Variable	*Mean*	*Min*	*Max*	*Coeff.*	*Exponent (Coeff.)*	*Percent change in average crashes*
Lane width (ft)	11.59	10	12	-0.4466	0.6398	36.02
Speed limit (MPH)	43.77	25	70	-0.0250	0.9753	2.47
Outer (right side) shoulder width (ft)	4.49	0	12	-0.0626	0.9393	6.07
Presence of streetlight (0/1)	0.69	0	1	-0.3670	0.6928	0.69
Land use (0 if residential, 1 if commercial or mixed)	0.44	0	1	0.8045	2.2356	2.23

TABLE 3-35 CMF FOR LANE WIDTH ON RURAL 4U ROADWAY SEGMENTS

Lane Width (ft)	*CMF*
10	1.7204
11	1.3602
12 (~11.59 ft)	1

TABLE 3-36 CMF FOR SPEED LIMIT (MPH) ON RURAL 4U ROADWAY SEGMENTS

Speed limit (MPH)	***CMF***
25	1.494
30	1.370
35	1.247
40	1.123
45 (~43.77 MPH)	1
50	1
55	1
60	1
65	1
70	1

TABLE 3-37 CMF FOR OUTER (RIGHT SIDE) SHOULDER WITH ON RURAL 4U ROADWAY SEGMENTS

Right Shoulder Width (ft)	***CMF***
0	1.3034
1	1.2427
2	1.1820
3	1.1214
4	1.0607
5 ft or more than 5 ft	1

TABLE 3-38 CMF FOR STREET LIGHT ON RURAL 4U ROADWAY SEGMENTS

Presence of Street Light	***CMF***
No (Base)	1
Yes	0.6928

TABLE 3-39 CMF FOR LAND USE ON RURAL 4U ROADWAY SEGMENTS

Land use	***CMF***
Commercial	2.236
Residential (Base)	1

3.3.2.4. Comparison of Default FHWA with TN-Specific Calibration Spreadsheet (Rural 4U Roadways)

To compare the performance of the TN-Specific calibration spreadsheet with the default FHWA calibration spreadsheet for rural 4U roadway segments, four roadway segments of the rural 4U roadway type were randomly selected to see how well the two calibration spreadsheets could predict the number of total crashes per year compared to the observed crash counts per year. For three (Segment #1, 3, and 4) of the randomly selected rural 4U roadway segments (Table 3.27), TN-specific calibration spreadsheet (either base-case TN-SPF or after applying CMFs) predict crash which is relatively closer to the observed crashes compared to the default FHWA calibration spreadsheets). For segment #3, the predicted crashes per year by enhanced TN-SPFs are very larger – because of CMFs for commercial land use, and speed limit (for details, please see corresponding CMFs provided in the earlier section or refer to the corresponding calibration spreadsheet).

TABLE 3.40 COMPARING PERFORMANCE OF CALIBRATION WITH NEW TN-SPECIFIC CALIBRATION (RURAL 4U)

Segment Details	***Randomly Selected 4U Roadway Segments***			
	Segment 1	***Segment 2***	***Segment 3***	***Segment 4***
County (Route)	Rhea (SR030)	Union (SR033)	Sevier (SR073)	Lake (SR078)
BLM & ELM	2.52 & 2.7	14.48 & 14.58	11.12 & 11.28	2.65 & 3.02
Special Case (Co-Sequence)	0-None (1)	0-None (1)	0-None (1)	0-None (1)
Segment Attributes				
Average AADT Per Year	3241	7065	15166	4052
Segment Length	0.18	0.10	0.16	0.37
Lane Width	12	12	11	12
Right (outer) shoulder width	5	6	1	4
Streetlight (1/0)	No	No	Yes	Yes
Speed limit	55	55	25	40
Commercial land use (1/0)	Residential	Residential	Commercial	Residential
Observed Crashes Per Year	1	0	4	1
Predicted Crashes Per Year				
*HSM Base SPF**	0.155	0.216	0.848	0.416
*HSM Enhanced SPF (After Applying CMFs)**	0.159	0.216	0.899	0.41
*TN-Specific Base SPF***	0.432	0.562	2.072	1.134
*TN Enhanced SPF (CMF Computed from TN-Enhanced Model)***	0.432	0.562	8.107	0.936

Note: * indicates that FHWA default calibration procedure & CMFs are used. ** indicates that the FHWA calibration spreadsheet was modified using TN-SPFs … See excel spreadsheet.

3.4. Comparing Performance of FHWA and TN-Specific Calibration Spreadsheets

Based on the TN-specific calibration procedures for the three rural roadway types which were developed, TDOT (especially Mr. David Duncan) provided valuable feedback. TDOT has tested the three predictive spreadsheets based on which trend of predicted crashes per year for each of the three rural roadways was developed. The trend (see Table 3.41 and Figure 3.12) for predicted crashes is shown on the three roadway types, which were created by Mr. David Duncan using the new TN-specific SPFs and calibration spreadsheets. Based on the empirical evidence (TN-SPF) for rural 5T roadway segments, it could be observed that speed limit (MPH) does not substantially affect crashes for 5T; however, this variable has been retained in the model specification for rural 5T roadway segments as it showed a statistically significant relationship with crashes. Overall, the trends were expected which show that rural 4U roadway is the most unsafe among the three rural roadways in TN (Figure 3.12). Furthermore, the trends reveal that rural 4D segments are safer (with a lower crash count per year) than rural 5T roadway segments for AADT above 7,000 which was again expected (Figure 3.12). However, in the lower range of AADT (below 7,000), the predicted crashes on rural 4D roadway segments are found to be higher than rural 5T roadway segments indicating that rural 5T roadway is safer than rural 4D in this range of AADT which seems counter-intuitive. To investigate this important and valid point and see the possible reason for this unexpected outcome, additional analyses was conducted and is documented in Appendix B.

TABLE 3.41 COMPARISON OF PREDICTED CRASHES VERSUS AADT VIA TN-SPFS

Speed limit (30 MPH)			
AADT	*5T (TN-SPF)*	*4U (TN-SPF)*	4D (TN-SPF)
4000	1.35	4.15	2.31
5000	1.84	5.29	2.55
6000	2.38	6.45	2.76
7000	2.96	7.64	2.94
8000	3.58	8.84	3.12
9000	4.22	10.05	3.28
10000	4.90	11.27	3.43
11000	5.60	12.51	3.58
12000	6.33	13.76	3.71
13000	7.09	15.01	3.84
14000	7.87	16.28	3.97
15000	8.68	17.55	4.09
16000	9.50	18.84	4.20
17000	10.35	20.12	4.31
18000	11.22	21.42	4.42
19000	12.11	22.72	4.52

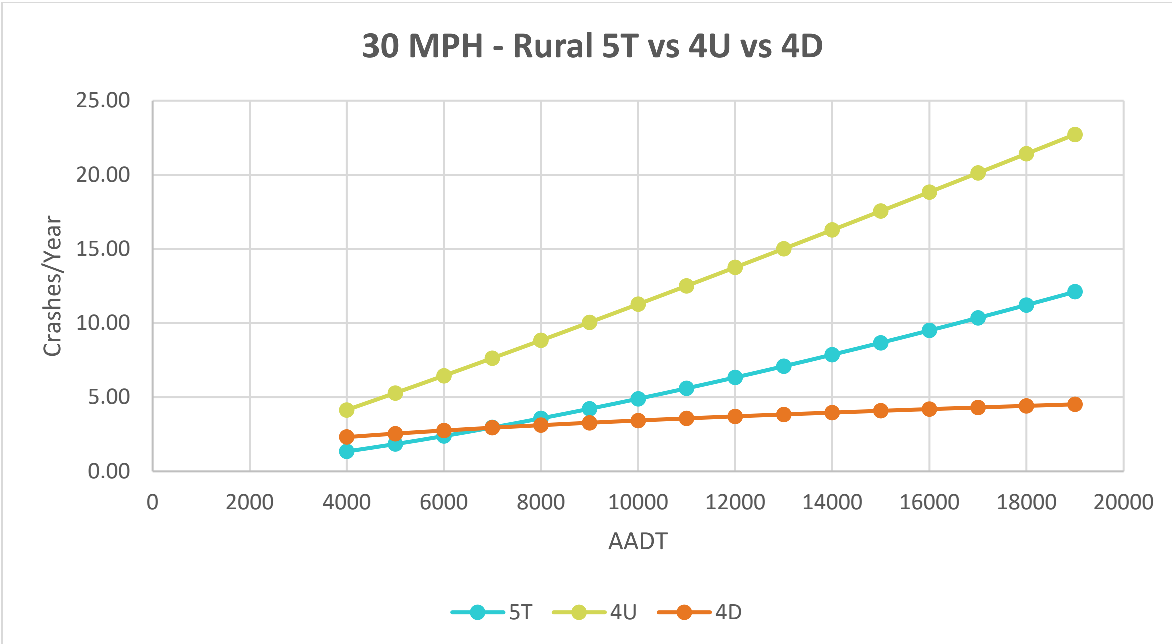


Figure 3.12 Comparison of Predicted Crashes versus AADT via TN-SPFs

Chapter 4 Conclusion

In modeling HSM Safety Performance Functions for rural roadways, this project has explored the role of conventional and new variables (e.g., speed limits and driveways) in road safety. The project required extensive data extraction, integration, and analysis, e.g., for rural 5T roadway segments, the crash, roadway, and traffic data were extracted using various sources and websites maintained by TDOT. The key objectives of this study included SPF estimation for Tennessee and the development of TN-specific calibration spreadsheets for future use by TDOT.

Statistics reveal that the observed average total crashes per year (from 2013 to 2017) on rural 4U, 5T, and 4U roadway segments are 2.56, 1.68, and 1.34, respectively. The average FI crashes per year on rural 4U, 5T, and 4D roadway segments are 0.59, 0.43, and 0.37, respectively. These statistics suggest that rural 4U segments have relatively more crashes, as expected. Furthermore, the research team estimated the values of a (coefficient of constant) and b (parameter of AADT) which can be used to predict the number of total and FI crashes for base conditions. Additionally, the CMF values for variables were estimated. The key findings are summarized below.

- Rural 5T roadway segments show that the average total crashes per year are lower with a wider center lane (2WLTL), higher speed limit (MPH), and wider regular lanes. Notably, the average of total crashes per year is higher on segments with more driveways. Furthermore, for the speed limit (MPH) of 45 MPH or higher, the CMF was 1.00. However, if the speed limit was lower, e.g., 40 and 35 MPH, crashes increase by 10.0% (CMF = 1.10) and 20.0% (CMF= 1.20).
- For the rural 4D roadway segment, the final TN-SPF suggests reductions in crashes with an increase in inner (left) shoulder width and speed limit (MPH) and if rumble strips are present on the inner shoulder. Furthermore, the average crashes per year are higher in locations with commercial or mixed land use compared with residential land use. The CMF for inner shoulder width is found to be 1.00 if it is equal to 4 ft or more. For every unit (1 foot) reduction below the mean value (4 ft), the average total crashes per year increase by 11.4%. Consequently, the CMFs for an inner shoulder with a width of 3 ft, 2 ft, and 1 foot are 1.114, 1.228, and 1.342, respectively.
- The average total crashes on rural 4U segments were lower with wider lanes, higher speed limits, and wider outer (right) shoulder width. The CMFs for the outer (right) shoulder width of 5 ft or higher was 1.00. Notably for every one-unit reduction in the width of the outer shoulder on rural 4U roadway segments below the mean values (5 ft), the average of total crashes per year increase by 6.07%.

Using the regression parameters for constant and AADT from base-case TN-SPFs and CMFs from the enhanced TN-SPFs, the research team developed the calibration spreadsheets for each of the three types of rural roadway segments. They can assist TDOT with estimating the safety performance of roadway segments. For improving the safety analysis in Tennessee, recommendations include:

- Use of HSM procedures and tools. The Tennessee Department of Transportation's adoption of the HSM procedures and corresponding investments in calibration procedures for safety improvements are ready for implementation. This study provides the data and tools for the analysis of segments on three rural roadway types. Specifically, TDOT can use these tools for rural 5T, 4D, and 4U roadway segments to predict crashes and any reductions associated with safety improvements in crash modification factors.

- Countermeasure selection. Based on the findings of this study, TDOT's Strategic Transportation Investments Division can use spreadsheet tools to explore countermeasures that can substantially improve safety on rural roads in Tennessee.
- Periodic updating of calibration factors and safety performance functions. The results of this study are based on crash data collected between 2013-2017. Given the substantial spatial and temporal variability in conditions, updating the calibration factors will help in applying the procedures embedded in the spreadsheets.
- For future research, TDOT can consider using emerging methods for crash prediction. These include artificial intelligence (AI) and machine learning (ML) methods to enhance safety outcome predictions. Notably, AI/ML methods can enhance prediction accuracy (though they are weaker in generating inferential knowledge). In this regard, the more robust heterogenous ensemble methods such as stacking can combine both traditional statistical models and AI/ML methods. Stacking is relevant for rural 5T roadway segments as the HSM (2010) does not provide crash prediction models for these roadway types. Note that stacking has shown promising performance for 5T urban and suburban arterials in Tennessee. For details, please see a relevant research paper in Appendix A.

References

1. AASHTO, Highway Safety Manual, in Washington, DC. 2010.
2. Wali, B., et al. Developing Safety Performance Functions for Rural Multilane Highways in Tennessee: Accounting for Unobserved Heterogeneity. in Transportation Research Board 99th Annual Meeting Transportation Research Board. 2020.
3. Wali, B., et al., Development of safety performance functions: incorporating unobserved heterogeneity and functional form analysis. Transportation research record, 2018. 2672(30): p. 9-20.
4. Khattak, A., et al., Highway safety manual safety performance functions & roadway calibration factors: Roadway segments phase 2, Part 1. 2020, Tennessee. Department of Transportation.
5. Akaike, H., B.N. Petrov, and F. Csaki, Second international symposium on information theory. 1973, Akadémiai Kiadó, Budapest.
6. Bozdogan, H., Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 1987. 52(3): p. 345-370.
7. Zhong, C., et al. Crash prediction on rural roads. in Proceeding of the 3rd International Conference on Road Safety and Simulation, Indianapolis, Indiana. 2011.
8. Thakali, L., L. Fu, and T. Chen, Model-based versus data-driven approach for road safety analysis: do more data help? Transportation research record, 2016. 2601(1): p. 33-41.
9. Shankar, V., F. Mannering, and W. Barfield, Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. Accident Analysis & Prevention, 1995. 27(3): p. 371-389.
10. Srinivasan, R. and D. Carter, Development of safety performance functions for North Carolina. 2011, North Carolina. Dept. of Transportation. Research and Analysis Group.
11. Caliendo, C., M. Guida, and A. Parisi, A crash-prediction model for multilane roads. Accident Analysis & Prevention, 2007. 39(4): p. 657-670.
12. Hauer, E., F.M. Council, and Y. Mohammedshah, Safety models for urban four-lane undivided road segments. Transportation research record, 2004. 1897(1): p. 96-105.
13. Abdel-Aty, M.A. and A.E. Radwan, Modeling traffic accident occurrence and involvement. Accident Analysis & Prevention, 2000. 32(5): p. 633-642.

Appendix- A Heterogeneously Ensemble Learning for Enhanced Crash Forecasts – A Statistical and Machine Learning based Stacking Framework

Abstract

This study aims to increase the prediction accuracy of crash frequency on roadway segments that can forecast future safety on roadway facilities. A variety of statistical and machine learning methods are used to model crash frequency on specific roadways – with machine learning methods generally having a higher prediction accuracy. Recently, heterogeneous ensemble methods (HEM), including "stacking," have emerged as more accurate and robust intelligent techniques and are often used to solve pattern recognition problems by providing more reliable and accurate predictions. In this study, we apply one of the key HEM methods, "Stacking," to model crash frequency on five-lane undivided segments (5T) of urban and suburban arterials. The prediction performance of "Stacking" is compared with parametric statistical models (Poisson and negative binomial) and three state-of-the-art machine learning techniques (Decision tree, random forest, and gradient boosting), each of which is termed as the base learner. By employing an optimal weight scheme to combine individual base learners through stacking, the problem of biased predictions in individual base-learners due to differences in specifications and prediction accuracies is avoided. Grid-search optimization and 10-fold cross validation procedures are used to obtain valid search ranges and optimal values of corresponding tuning parameters for individual machine learning and stacking methods. Data including crash, traffic, and roadway inventory were collected and integrated for the period of 2013 to 2017. The data are split into training (2013-2015), validation (2016), and testing datasets (2017). First, five individual base-learners are trained using training data. Next, prediction outcomes obtained from these five base-learners using the validation dataset are obtained, which are then used to train a meta-learner. Estimation results of statistical models reveal that besides other factors; crashes increase with density (number per mile) of different types of driveways. Individual machine learning methods show similar results – in terms of variable importance. Comparison of out-of-sample predictions of various models or methods confirms the superiority of "Stacking" over the alternative methods considered. From a practical standpoint, "stacking" can enhance prediction accuracy (compared to using only one base learner with a particular specification). When applied systemically, stacking can help identify more appropriate countermeasures.

Keywords: Count Data Models, Machine Learning, Base-learners, Meta-learner, Stacking.

A1. Introduction

Safety performance functions (SPFs) or crash prediction models are extensively used to predict the expected level of safety on specific roadway types. These models help evaluate the safety performance of specific countermeasures on a particular type of roadway or intersection. These practices are well discussed in the Highway Safety Manual (HSM, 2010) which presents SPFs for various roadway types [1]. HSM (2010) was developed by AASHTO to provide a coherent and rigorous methodology to evaluate the safety performance on national roads [1]. HSM SPFs were developed using data from specific states - and given the variations in geographical conditions, driving behaviors, and design practices nationally [2, 3], HSM highly recommends calibration of HSM SPFs to local conditions or developing jurisdiction-specific SPFs [1].

Traditionally, count data models (Poisson and negative binomial models) have been extensively used to model the relationships between crash frequency and key correlates, such as annual average daily traffic (AADT) and segment length, etc. [4-10]. Compared to Poisson models, negative binomial variants are well-suited to capture potential over-dispersion in crash data. These models provide rich inferential insights into the mechanisms through which associated factors correlate with safety outcomes. However, given the intrinsic parametric nature of the models and the subsequent assumptions, the prediction accuracy of count data models is often a concern. Growing evidence of the role of more accurate crash predictions in designing more appropriate safety countermeasures has led to an increased interest in machine learning methods. Unlike count data models, machine learning methods do not place strong restrictions on the specifications of the model [11]. Machine learning methods are more adequate for modeling complex non-linear relationships that frequently arise in crash data modeling. Tree-based regression (TBR) is one of the most popular and widely-used machine learning methods that does not require variable transformations and parametric assumptions [12, 13]. TBR determines significant non-linear relationships among various predictor variables as well as computes the relative influence of predictors on response outcome [12, 14]. However, the TBR technique is prone to instability leading to estimation results with higher variance [12, 15]. Ensemble methods like random forest regression (RFR) and gradient boosting regression (GBR) combine the estimates of numerous trees compared to a single tree, leading to improved stability and prediction accuracy [12, 16, 17]. GBR technique ensembles numerous trees in a sequential way with a slower learning rate that captures a higher variance in data compared to the RFR method [12]. While the prediction accuracy of machine learning methods usually is greater than the count data models, it lacks a holistic inferential framework providing little to no information about the safety mechanisms that link unsafe outcomes with key risk factors. Also, almost all the machine learning (ML) methods explicitly relate to the bias-variance trade off contour with different methods minimizing bias or variance. There is no escaping the relationship between bias and variance in machine learning models. Thus, the use of the single supervised or the unsupervised ML method could lead to relatively less accurate predictions.

While traditional count data models and ML methods have been extensively used in the safety literature, studies that combine the predictive (and inferential) strengths of both paradigms or the strengths of multiple ML methods are rare. The prediction performance of ML methods can be further improved by using more robust and heterogeneous ensemble methods (HEM), such as composite systems, stacking, or blending [18-22]. HEMs including "stacking" have emerged as more accurate and reliable intelligent techniques in pattern recognition issues. The idea of "Stacking" essentially helps in harnessing the gains simultaneously from less biased and low-variance predictions offered by different ML methods. For example, the gradient-boosting regression method builds on so-called "weak classifiers" - reducing prediction error mainly by reducing bias (and to some extent variance, by aggregating the predictions from many trees). Through heterogeneous ensemble methods such as "Stacking", predictive gains from different methodologies can be combined. For example, the predictive gains from low bias in the gradient boosting method can be combined with predictive gains from lowering variance through the random-forest method via stacking. Studies also suggest other sophisticated approaches like the information Entropy-Bayesian network method to enhance crash severity prediction [23]. In recent years, the stacked generalization approach – a more robust and accurate ML method, has been used in transportation safety [24-27]. However, very few studies have applied this robust ML method to solve problems related to road safety [24-27]. Note that the stacking approach can be used in the contexts of both regression and classification problems. However, most of the aforementioned studies applied stacking to solve classification (e.g., injury severity analysis) problems [24, 25]. For instance, [24] applied the stacked generalization approach to predict crash severity with the severity levels of no injury, invisible injury, no-capacitating injury, and highest injury severity. The predictions obtained from three individual ML classifiers like the random forest, adaptive boosting, and gradient boosting decision tree were combined via stacked model using logistic regression in the second layer [24]. Prediction accuracy of the stacked model was significantly higher compared to individual ML methods such as random forest classifier [24]. A similar stacked classification approach was used in one of the recent studies, which adopted a hybrid combination of homogeneous and heterogeneous ensemble methods to explore factors associated with fatal road crashes [25]. The aforementioned study reveals that prediction accuracy can significantly improve via the stacked generalization approach [25]. Studies also reveal that crash risk prediction can be significantly improved via stacking predictions from individual ML algorithms [27]. In another recent study, it was found that prediction of risky and aggressive driving behavior among taxi drivers can significantly improve via stacked generalization approach compared to individual ML classifier [26]. Stacking approach was mostly applied to solve classification problems related to transportation safety [24-27].

Note that ensembles including RFR, GBR, and stacking are used to improve out-of-sample prediction accuracy and can be classified into: (i) Homogeneous ensemble, and (ii) Heterogeneous ensemble [28-31]. The homogeneous ensemble (e.g., RFR and GBR) uses the same feature selection algorithm with different training or learning datasets distributed over various nodes [28, 29]. Instead, the heterogeneous ensemble (i.e., stacking) uses different feature selection algorithms (e.g., Poisson, Negative binomial, TBR, RFR, and GBR) where the stacking meta-learner (which can be any statistical or ML method) blends the optimal combinations of predictions by base-learners and acts as a single decision maker in the second-stage [29, 31, 32]. Both homogeneous and heterogeneous ensembles can be used in regression as well as classification contexts. Compared to homogeneous ensembles, heterogeneous ensembles typically show significant performance gains [29, 31]. Both types of ensembles are used in diverse fields (such as medicine) where their application provides more accurate and reliable predictions of a specific disease in patients [30, 33, 34]. Studies suggest that heterogeneous ensembles do not only outperform the conventional statistical models and other ML methods but also show superior prediction performance compared to homogeneous ensembles [30, 33, 34]. In transportation safety, homogeneous ensembles have been widely used for predicting crash frequency [12, 35, 36] and severity given a crash [37, 38]. Some studies used heterogeneous ensembles (e.g., stacking) in classification context to predict injury severity [24, 25]. However, the application of heterogeneous ensemble (stacking) to predict crash frequency on roadways has not been or very lightly explored to the best of the authors knowledge. Given the prevalent gaps in the literature discussed above, this study contributes by:

- Applying a rigorous and robust HEM scheme to model and predict crash frequency on five-lane (5T) undivided segments on urban and suburban arterials, including two-way left-turn lanes (2WLTL).
- Comparing the prediction performance of "Stacking" with traditional statistical models and three state-of-the-art machine learning techniques (decision trees, random forest, and gradient boosting regression).

The statistical (Poisson and negative binomial models) and ML models used in this study are considered as "base-learners." To obtain valid search ranges and optimal values of corresponding tuning parameters for individual base- and stacked learners, grid-search optimization and 10-fold cross-validation procedures are used. It is shown that using more accurate, reliable, and robust intelligent techniques can extract more useful information compared to individual count data or ML methods. To achieve the study objectives, detailed crash and roadway geometric data are extracted from the Enhanced Tennessee Roadway Information Management System (E-TRIMS).

A2. Methodology

A2.1. Conceptual Architecture: Heterogeneous Ensemble Methods (Stacking)

The idea of HEM, including "Stacking" was first introduced almost thirty years ago [39]. In stacked regression, predictions from various individual models (base-leaners) are combined and used as input in second-stage learning [40]. Stacking generally provides higher prediction accuracy compared to base-learners [40]. Suppose Y is the response outcome, X is the set of predictors used in individual models (briefly discussed in subsequent sections), and $g_1, g_2, \dots, g_L$ are the predictions obtained using base learners [40]. The prediction function for the linear ensemble (stacked) model can be given as [40]:

$$b(g) = (w_1 * g_1) + (w_2 * g_2) + \dots + (w_L * g_L) \tag{1}$$

Note that w_i indicates the weight assigned to an individual model in the stacking technique [40]. The model weights (w_i) are used to minimize MSE between actual response variable (y_i) and prediction outcome of meta-learner (stacked ensemble technique) as shown [40]:

$$\min \sum_{i=1}^{N} (y_i - (w_1 * g_{1i} + w_2 * g_{2i} + \dots + w_1 * g_{Li}))^2 \tag{2}$$

The conceptual design of this study is presented in Figure A.1. First, we manually extracted crash, traffic, and roadway geometry data using various software packages made available by the Tennessee Department of Transportation (TDOT) for a randomly selected subsample containing 304 roadway segments of 5T urban and suburban arterials for a period of five years (2013-2017). Next, we split data into training (2013-2015), validation (2016), and testing (2017) datasets (Figure A.1). Note that in all the three datasets, only crashes and average annual daily traffic may change while all other factors remain the same. We follow this splitting procedure to develop a crash prediction model which can be reused with updated data to forecast crashes in the future. First, five individual base-learners are trained using training data to model crash frequency per year (2013-2015). Next, prediction outcomes obtained from these five base-learners using the validation dataset are obtained and combined with actual crashes reported in 2016, which generates a new training dataset for the meta-learner (stacking). Note that grid-search optimization and 10-fold cross-validation procedures are used to obtain valid search ranges and optimal values of corresponding tuning parameters for individual machine learning techniques and stacking (Figure A.1). In the 10-fold cross-validation procedure, an algorithm splits the available data (used to train the model) into 10 subsamples of equal sizes, and nine of those subsamples are used for training, while one subsample is used for testing to determine the optimal model for prediction accuracy. The algorithm repeats the process 10 times, during which each of the subsamples is once used as a testing subsample. The results are finally averaged to get a single estimation. Note that studies commonly use 10-fold cross-validation procedure for the tuning of the machine learning models [41]. Finally, we apply individual base learners (trained using the training dataset) and meta-learners or the stacked model (trained using validation dataset) to the new data (2017) to accurately compare their prediction performance (Figure A.1).

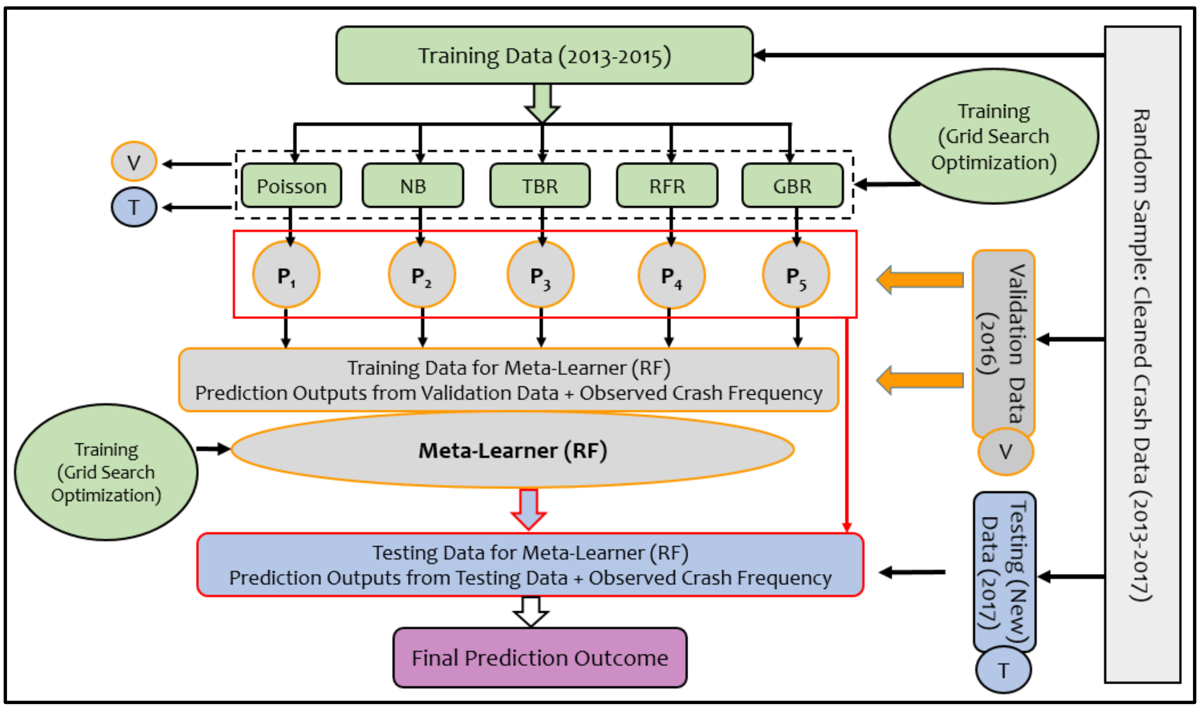


Figure A. 1 Conceptual Design of Stacking Ensemble Utilized for Crash Frequency Modeling

Notes: In Figure A.1, NB indicates a negative binomial model, TBR indicates a Tree-based regression, RFR indicates a random forest regression, and GBR indicates a gradient boosting regression. P_1, P_2, P_3, P_4, and P_5 are the prediction outcomes obtained while applying Poisson, NB, TBR, RFR, and GBR models to the validation dataset, respectively. V and T indicate validation and testing datasets, respectively.

This study applies stacking where a meta learner is used to combine multiple predictions obtained from various base learners, as explained below.

- **Base Learner:** Stacking is a two-stage process where individual statistical models and/or ML methods are applied in the first stage. Any statistical model or ML method when applied in the first stage of stacking is termed as a "base learner" in this study. For instance, this study applies five base learners which include two statistical models (Poisson and Negative binomial) and three ML methods (TBR, RFR, and GBR). The base learners applied in this study also include homogeneous ensembles (RFR and GBR), which use the same feature selection algorithm with different training datasets. In homogeneous ensembles, the results and/or predictions are averaged.

- **Meta Learner:** The stacking meta-learner algorithm is an ensemble technique that combines predictions from two or more than two base-learners specifically to further enhance prediction accuracy. This study uses three ML methods including TBR, RFR, and GBR as meta-learners to combine predictions for the five base-learners (Poisson, negative binomial, TBR, RFR, and GBR). Finally, after comparing the out-of-sample RMSE and MAE of all the base-learners and three meta-learners, we selected one model which has the lowest out-of-sample RMSE and MAE. Note: that stacking is termed as a "heterogeneous ensemble" that combines different feature selection procedures (Poisson, Negative binomial, TBR, RFR, and GBR). In stacking, a meta-learner can also be termed as a super-learner [42].

A2.2. Count Data Models: Poisson and Negative Binomial Regression

Studies suggest using count data (Poisson and negative binomial regression) models to explore the relationship of crash frequency with explanatory variables [2, 3, 43]. Poisson regression was first introduced by a French mathematician named Siméon-Denis Poisson in 1830. The mathematical formula of Poisson regression is given below [44].

$$\mathrm{P}(n_i) = \frac{exp\ (-\lambda_i)\lambda_i^n}{n_i!} \tag{3}$$

Where $P(n_i)$ is the probability of a crash occurring on a specific road segment (i), (n) is the frequency of a crash on a specific road segment at a particular time, and (λ_i) is the expected number of crashes occurring on a particular road segment (i) in a specific duration. The expected number of crashes (λ_i) is linked to its key contributing factors as below [2, 3, 44]:

$$ln(\lambda_i) = \beta(X_i) \tag{4}$$

Where X_i indicates a set of explanatory variables, and β are their associated parameter estimates.

The equations (3 and 4) can be maximized using the standard maximum likelihood procedure [3, 44]:

$$L(\beta) = \prod_i^n \frac{exp\ [-exp\ (\beta X_i)\ ]\ [exp\ (\beta X_i)\ ]^{n_i}}{n_i!} \tag{5}$$

In the case of over-dispersion, Poisson regression is not preferable due to violation of its basic assumption, therefore negative binomial regression is suggested as below [3, 44].

$$ln\lambda_i = \beta(X_i) + \epsilon_i \tag{6}$$

Where $exp\ (\epsilon_i)$ is an error term with gamma distribution "mean equals one and variance (α)" [3, 44]. The conditional probability for crashes can be given as [3, 45]:

$$P(\epsilon) = \frac{exp\ [-\lambda_i exp\ (\epsilon_i)\ ]\ [\lambda_i exp\ (\epsilon_i)\ ]^{ni}}{n_i!} \tag{7}$$

The error term (ϵ_i) can be integrated out to determine the unconditional distribution of n_i as given below [3, 45]:

$$P(n_i) = \frac{\Gamma(\theta+n_i)}{[\Gamma(\theta).n_i!]}.u_i^{\theta}(1-u_i)^{n_i} \tag{8}$$

Where u_i equals $\theta(\theta+\lambda_i)$ and $\theta = \frac{1}{\alpha}$, and Γ is a gamma function. In the case of (α) approaching zero, the negative binomial simply becomes a Poisson regression [44]. Negative binomial regression is preferred over Poisson regression when it is significantly different from zero [43, 44]. To evaluate the goodness of fit performance of the count data models, McFadden R^2 value [3], Akaike Information Criteria [3, 46, 47], and Bayesian Information Criteria [3, 48] can be used.

A2.3. Machine Learning Methods

A2.3.1. Decision-tree Regression

Decision tree uses a fast algorithm that recursively splits training data into smaller subsets [49]. However, instability and reliability issues are key weaknesses of this method [49, 50]. The algorithm searches to determine a splitting point with the lowest value of mean square error (MSE). At the optimal splitting point, the parent node is further split into two child nodes and the process continues until the optimal tree length is determined (reducing impurity associated with terminal node). The algorithm chooses the best splitter (S^*) considering deviance (D) or MSE at a particular node as:

$$D(t) = \sum_{x \in t}(Y_n - \hat{\mu})^2 \tag{9}$$

Where, $\hat{\mu}$ is a sample mean $(\bar{y})$ or mean estimate, t indicates a specific node, and X indicates a set of predictors. Referring to the generalized linear models, deviance (D) is also termed as log-likelihood ratio statistics and can be written as:

$$D = 2 * l(\mu_{max}; y) - l(\hat{\mu}: y) \tag{10}$$

Where, μ_{max} is the maximum likelihood estimate. Deviance of a tree (T) can be determined as below:

$$D(T) = \sum_{t \in \acute{T}} D(t) = \sum_{t \in \acute{T}} \sum_{x \in t}(Y_n - \bar{y}(t))^2 \tag{11}$$

Where T is the tree, $\acute{T}$ is a set of terminal nodes of T. For a binary partition via splitter (s), the difference is:

$$\Delta D(s,t) = D(t) - D(t_L) - D(t_R) \tag{12}$$

Where t_L and t_R indicate left and right child of parent node (t) respectively. The difference is maximized to determine the best splitter (s^*) as:

$$\Delta D(s^*,t) = max_{s \in S} \Delta D(s,t) \tag{13}$$

Note size selection and tree pruning is carried out using 10-fold cross-validation to select the optimal tree size with the lowest MSE.

A2.3.2. Random Forest Regression

Studies suggest ensemble methods like RFR and GBR to mitigate instability issues related to a single decision [51, 52]. The RFR algorithm works on a similar principle to the single decision tree;

however, the key difference is that RFR assembles an enormous number of trees. The RFR algorithm selects a predictor at each node to maximize homogeneity at successive nodes [53, 54]. Regularization parameters considered for RFR include [53, 55, 56]:

- Number of predictors selected at each node for split-up ($\boldsymbol{m_{try}}$)
- Number of trees in forest (n_{tree})
- Number of maximum nodes in the forest

Studies suggest using the following trials to select the optimal number of predictors ($\boldsymbol{m_{try}}$) at each node [50]:

- $m_{try} = \frac{p}{3}$
- $m_{try} = \frac{1}{2} * \frac{p}{3}$
- $m_{try} = 2 * \frac{p}{3}$

While p is the total number of predictor variables considered in RFR regression. Note that $\boldsymbol{m_{try}}$ indicates the number of variables/predictors available for splitting at each node. It is considered as an important regularized or tuning parameter [57]. To determine the optimal value of $\boldsymbol{m_{try}}$, we use an extended grid-search optimization and 10-fold cross validation procedure. To select an optimal pair of n_{tree} and $\boldsymbol{m_{try}}$, two performance criteria including MSE and R^2 values are usually used [53]:

$$MSE \approx MSE_{OOB} = \frac{1}{n}\sum_{i \in OOB}(y_i - \hat{y}_i)^2 \tag{14}$$

$$R^2 = 1 - \frac{MSE_{OOB}}{Var(y_i)} \tag{15}$$

Where MSE_{OOB} is the MSE for the out-of-bag (OOB) sample, y_i is the observed number of crashes occurring on i^{th} roadway segment in OOB sample, $\hat{y}_i$ is predicted crashes on i^{th} road segment in OOB sample, n is number of roadway segments in OOB sample, $Var(y_i)$ is the variance of response outcomes (y) determined as $\frac{1}{n}\sum_{i \in OOB}\ (y_i - \overline{y})^2$, while $\overline{y}$ is mean value of y_i in the OOB sample.

Similar to the TBR approach, variable importance relates to the reduction in node impurity at each split; however, the RFR technique uses the average reduction of all trees in the forest to determine the overall reduction in impurity. Importance $Imp\ (X_m)$ of any particular predictor variable X_m, is computed while summing the weighted reduction in node impurities, $\Delta_i(s,t)_{X_m}$, for all nodes t where X_m is used for splitting [51]:

$$Imp\ (X_m) = \frac{1}{N_T}\sum_T \sum_{t \in T} \frac{N_t}{N} \Delta_i(s,t)_{X_m} \tag{16}$$

Where N_T is the number of trees, N_t is the number of data points at a specific node (t), and N is the sample size.

A2.3.3. Gradient Boosting Regression

Similar to the RFR approach, GBR is a pool procedure to enhance prediction accuracy [17, 54, 58]. The algorithm calculates residuals after fitting the first tree to the l} due to which the GBR algorithm assigns more weight to such observations while fitting the next tree and so on [12]. In this straightforward and stage-wise process, the GBR algorithm keeps the existing tree unchanged while re-estimating residuals for every observation to reveal contributions to the new tree [12, 58]. Let $f(x)$ be an approximation function of response outcome (y) as predicted by a set of predictor variables (x). In GBR approach, an additive expansion of the basic functions $(x:\gamma_m)$ can be given as:

$$f(x) = \sum_m f_m(x) = \sum_m \beta_m b(x:\gamma_m) \tag{17}$$

Note that $\beta_m(m = 1,2,3, \dots, M)$ indicates the expansion coefficients, $b(x:\gamma_m)$ indicates single regression trees having parameter (γ_m) as a split variable, and β_m are the weights assigned to every tree [12]. The algorithm estimates parameters like β_m and γ_m to minimize loss function $L(y(f(s))$ indicating prediction performance in term of deviance [12]. Note that while GBR may nicely fit to the data, it can also lead to overfitting [12]. To cure this issue, studies suggest selecting appropriate regularization parameters including the number of trees, shrinkage (learning rate), and complexity which help in achieving a balance between variance and bias [12]. The learning rate is usually smaller ranging from 0.0001 to 0.1 [12]. Note that smaller values of shrinkage parameters are good but require more trees. The complexity parameter refers to tree depth which shows interactions among predictor variables [12].

A2.4. Model Performance

To evaluate the prediction performance of individual models (Poisson, negative binomial, TBR, RFR, and GBR) and stacked regression, we compare their Root Mean Square Error (RMSE) [3] and Mean Absolute Error (MAE) [3, 59] based on the testing dataset (2017):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2} \tag{18}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e| \tag{19}$$

The value of n is the total number of roadway segments, and f_i and y_i indicate predicted and observed crash frequency, respectively. Low values of RMSE and MAE indicate higher prediction accuracy.

A3. Results and Discussion

A3.1. Data Processing and Descriptive Statistics

Data used in this research was extracted from the Enhanced Tennessee Roadway Information Management System (E-TRIMS) which is a roadway inventory and crash database maintained by the Tennessee Department of Transportation (TDOT). We identified the five-lane (5T) roadway segments of urban and suburban arterials by selecting the attributes of interest including the number of through lanes (five lanes), presence of two-way-left-turn lanes (2WLTL), and functional class (urban arterials). The roadway segments are pre-defined in E-TRIMS where a segment refers to a portion of the roadway that either connects two nodes (i.e., intersections) or has uniform features, e.g., lane width, shoulder width, number of lanes, and median width as compared to neighboring (proceeding and succeeding) roadway segments. E-TRIMS showed a total of 3,208 (753.97 miles) segments of state-maintained 5T urban and suburban arterials in Tennessee. Figure A.2 shows the distribution of the total 5T roadway segments (N = 3,208) of urban and suburban arterials in TN identified in E-TRIMS which was first cleaned and then a random sample (N = 304) was selected for analysis.

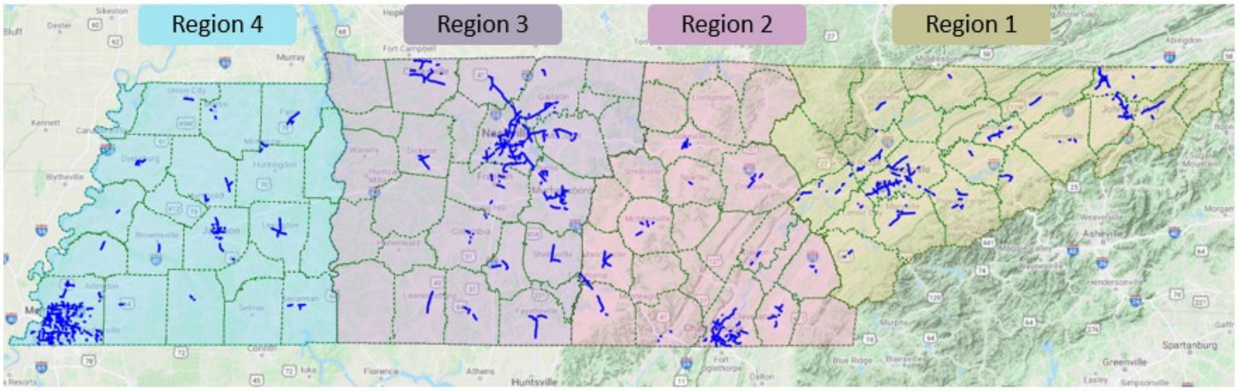


Figure A. 2 Distribution of the Overall 5T roadway segments of Urban and Suburban arterials in TN.
Note: Tennessee has 95 counties, which are divided into four TDOT regions, shown on the map.

Following the HSM guidelines [1], segments shorter than 0.1 miles were removed leading to a reduced dataset containing 1,519 segments (totaling 523.93 miles). First, we determined the sample size to be selected from the population (1,519 segments) using 95% confidence level criteria. A random sample of 317 segments (105.78 miles) was selected for which crash (2013-2017), roadway geometry, and traffic data (2013-2017) were extracted using E-TRIMS and TDOT Traffic History Application. Finally, 304 (103.27 miles) segments with complete data are considered in the analysis. The distribution of roadway segments of 5T urban and suburban arterials (random sample "N = 304") in Tennessee based on the total number of crashes that have occurred on these segments during the 5-years (2013-2017) period is shown in Figure A.3. The segments with a higher number of total crashes during the 5-years period are mostly located in TDOT Region 3, which contains the Nashville area (Figure A.3). Notably, segments with a low number of crashes over the 5-year period are mostly located in the suburbs of the major cities or other urban areas and small cities (Figure A.3). Note that each circle refers to a roadway segment of 5T urban and suburban arterials in TN, and the size of the circle depicts the total number of crashes that have occurred on a roadway segment during the 5-year (2013-2017) period.

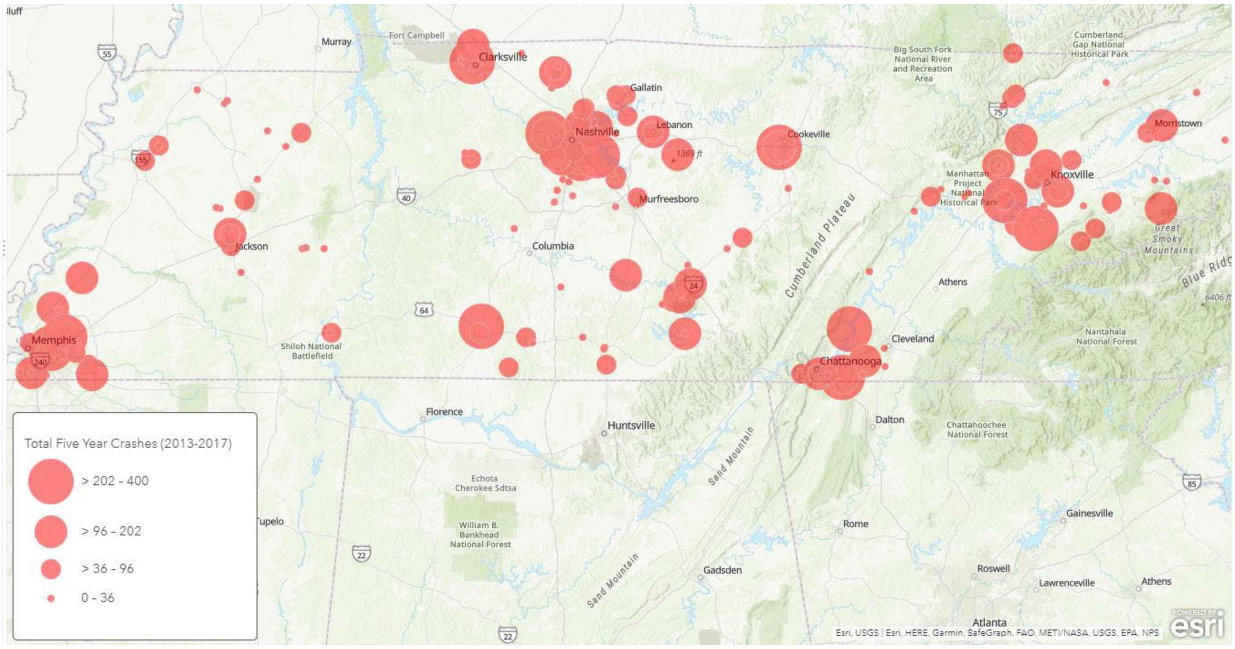


Figure A. 3 Distribution of 5T Urban and Suburban Arterial Segments based on 5-year (2013-2017) Crashes in TN.

A study by AASHTO has revealed that factors like density (number per mile) of major/minor driveways based on various land uses (e.g., commercial and industrial) and average offset distance to fixed objects significantly influence crash frequency on a roadway segment [1]. Fixed objects (utility poles, traffic signs, trees, and billboards) along roadway segments are considered as potential safety risks [60-63]. Such objects are more prevalent along urban roadway segments [61-63]. Distance to fixed objects along roadway segments is critical as the risk of fixed-object collisions increases as the offset distance to roadside fixed objects decreases [1]. In HSM (2010), SPFs for all types of urban roadways include offset to roadside fixed objects as an important factor to predict crashes on specific roadway segments [1]. We consider it important to include the average offset distance to fixed objects along the roadway segments of 5T urban and suburban arterials in the models. The offset distance (measured in ft) to every fixed object along the roadway segment may vary; therefore, we calculated and used the average of the offset distances to fixed objects along the roadway segments in the models.

To achieve our study objective, the data is split into three subsets: training, validation, and testing. Table A.1 presents descriptive statistics of key variables. Statistics reveal an average of 11.026 crashes (standard deviation of 14.020) across the three years on 5T segments of urban and suburban arterials. Crash distributions for validation (2016) and testing (2017) are shown in Table A.1 revealing similar distributions across the three (training, validation, testing) streams. Statistics for traffic measures and roadway geometric features are provided in Table A.1. In 2017, the mean AADT (in 1000s) was 19.903 which is slightly higher than the yearly AADT in 2016 and average AADT per year from 2013 to 2015 (Table A.1). This shows that on average, AADT per year has increased slightly compared to the previous years. The sample statistics show the mean segment length to be 0.340 miles including no segment with a length less than 0.1 miles (Table A.1). The mean offset distance to roadside fixed objects is found to be 14.26 ft (Table A.1). Referring to the density (i.e., frequency per mile) of driveways based on various land uses, the density of minor industrial driveways was found to be the highest density with a mean value of 1.286 driveways per mile (along both sides of the roadway segment) followed by minor commercial driveways (0.865 per mile) and major industrial driveways (0.461 per mile), as shown in Table A.1. The descriptive statistics seem reasonable because the dataset contains little to no outliers.

TABLE A. 1 DESCRIPTIVE STATISTICS OF KEY VARIABLES: 5T SEGMENTS OF URBAN AND SUBURBAN ARTERIALS

Variables	***Obs.***	***Mean***	***Std. Dev.***	***Min***	***Max***
Average Three-years Crashes (2013-15)	304	11.03	14.02	0.00	72.00
Total Crashes (2017)	304	11.07	14.62	0.00	100.00
Total Crashes (2016)	304	11.01	14.65	0.00	90.00
Average Annual Daily Traffic (AADT) per Year (2013-15) in 1000s	304	19.10	8.94	3.18	49.77
Average Annual Daily Traffic (AADT) (2017) in 1000s	304	19.90	9.15	3.81	54.56
Average Annual Daily Traffic (AADT) (2016) in 1000s	304	19.64	9.21	3.61	54.36
Segment length (mile)	304	0.34	0.28	0.10	1.81
Density (frequency per mile) of Major Commercial Driveways	304	0.35	0.78	0.00	6.00
Density (frequency per mile) of Minor Commercial Driveways	304	0.87	1.63	0.00	12.00
Density (frequency per mile) of Major Industrial Driveways	304	0.46	0.94	0.00	7.00
Density (frequency per mile) of Minor Industrial Driveways	304	1.29	1.84	0.00	11.00
Average Offset Distance (ft) to Roadside fixed objects	304	14.27	8.19	0.00	30.00

A3.2. Estimation Results

A3.2.1. Count Data Models: Poisson and Negative Binomial Regression

As a first step, we apply Poisson and negative binomial models to explore the average three-years (2013-2015) crash frequency. Both models come from a series of trials evaluated based on statistical significance, parsimony, and intuition. To select more appropriate models (with superior fit), several trials were made based on the specifications of explanatory variables. Initially, Poisson and negative binomial models were estimated including all the significant variables (including AADT and segment length) in their original forms (Model 1). Next, logarithmic forms of AADT and segment length were included while keeping all other covariates (e.g., the density of major/minor commercial and industrial driveways and average offset to roadside fixed objects) in their original forms (Model 2). Finally, logarithmic forms of all significant variables were tested. Including logarithmic forms of all variables in the model did not lead to improvements (results not shown for brevity). Poisson and Negative Binomial models with log-transformed AADT and segment length variables (Model 3 and Model 4) outperformed the counterparts with untransformed variables based on AIC, BIC, and log-likelihood values at convergence (Table A.2). Thus, Model 3 and Model 4 (including logarithmic forms of AADT and segment length) were selected as the best models compared to their counterparts. Similar specifications for the key variables (ln forms of AADT and segment length) were used while training machine learning methods. To quantify the effects of significant variables on crash frequency, we present marginal effects (MEs) in Table A.2. According to the estimation results, AADT (2013-2015) per year and segment length (mile) were positively correlated with the average three-years crash frequency (Table A.2). In terms of geometric factors, density (number per mile) of four key types of driveways including major commercial, minor commercial, major industrial/institutional, and minor industrial or institutional are also positively correlated with crash frequency on 5T segments of urban and suburban arterials (Table A.2). We also found that the average offset distance (ft) to fixed objects along these segments is negatively associated with the average three-years (2013-2015) crash frequency (Table A.2). The over-dispersion parameter in negative binomial models is found to be statistically significant, indicating that the negative binomial model is preferred over the Poisson regression (Table A.2).

To understand the relationship of key variables and crash frequency, we discuss the marginal effects of variables for the best statistical model (negative binomial model with AADT and segment length in ln forms), which has the best in-sample fit (Table A.2). Our findings indicate that yearly crash frequency increases by almost 13 units with a unit increase in yearly AADT in 1000s (ln form) while keeping all other variables at their means (Table A.2). Similarly, a unit increase in segment length (ln form) is associated with increases in crash frequency by 5.87, while keeping other variables at their mean values (Table A.2). The estimation results of the best-fit model suggest that major commercial driveways have a stronger association compared with other types of driveways with crash frequency i.e., a unit increase in density of major commercial driveways is associated with an increase in yearly crashes by 1.135 units (Table A.2). Moreover, yearly crash frequency is higher by 1.078, 0.744, and 0.505 with a unit increase in density of minor commercial driveways, major industrial driveways, and minor industrial driveways (Table A.2). Other studies suggest similar findings [1, 64, 65]. These findings were expected as an increase in commercial and industrial driveways increases potential conflict points and creates a potential for gap acceptance errors. These findings highlight the need for investigating proactive access management strategies which can potentially reduce crashes specifically on 5T roadway segments of urban and suburban arterials. Our findings indicate that higher offset distance to roadside fixed objects is associated with fewer crashes. Crash frequency is lower by 0.211 with a unit increase in average offset distance (ft) to roadside fixed objects (Table A.2). This was expected as fixed objects (i.e., utility pole, traffic sign, tree, and billboards) along roadway segments are potential safety risks specifically for errant vehicles [60-63]. Such objects are more prevalent along urban roadway segments [61-63]. The distance to fixed objects along the roadway segments is a critical factor because the risk of fixed-object collisions is lower with higher offset distances to roadside fixed objects [1].

TABLE A. 2 ESTIMATION RESULTS OF POISSON AND NEGATIVE BINOMIAL MODELS

Variables	*Poisson (Model 1)*			*Negative Binomial (Model 2)*			*Poisson (Model 3)*			*Negative Binomial (Model 4)*		
	Data (2013-2015)			*Data (2013-2015)*			*Data (2013-2015)*			*Data (2013-2015)*		
	Coeff.	*t-stat*	*MEs*	*Coeff.*	*t-stat*	*MEs*	*Coeff.*	*t-stat*	*MEs*	*Coeff.*	*t-stat*	*MEs*
Constant	0.8804	13.39	---	0.5274	2.96	---	-0.527	-3.44	---	-0.446	-1.28	---
Average Annual Daily Traffic (AADT) per Year (2013-15) in 1000s	0.0548	30.49	0.603	0.0603	10.41	0.707	---	---	---	---	---	---
Segment length (mile)	0.8223	13.85	9.066	0.9078	4.17	10.656	---	---	---	---	---	---
Density (frequency per mile) of Major Commercial Driveways	0.1156	6.87	1.274	0.1298	1.94	1.524	0.072	4.32	0.795	0.102	1.54	1.135
Density (frequency per mile) of Minor Commercial Driveways	0.0808	8.80	0.891	0.1183	3.53	1.388	0.063	7.03	0.693	0.097	2.91	1.078
Density (frequency per mile) of Major Industrial Driveways	0.0765	5.28	0.843	0.1163	2.07	1.364	0.045	3.09	0.496	0.067	1.22	0.744
Density (frequency per mile) of Minor Industrial Driveways	0.0495	5.65	0.545	0.0675	2.21	0.792	0.029	3.41	0.321	0.045	1.50	0.505
Average Offset Distance (ft) to Roadside fixed objects	-0.0215	-8.05	-0.237	-0.0148	-2.18	-0.173	-0.028	-10.12	-0.303	-0.019	-2.84	-0.211
Key Variables (ln form)												
AADT per Year (2013-15) in 1000s (ln form)	---	---	---	---	---	---	1.261	28.72	13.904	1.146	11.42	12.756
Segment length (mile) (ln form)	---	---	---	---	---	---	0.567	18.03	6.250	0.522	5.87	5.805
Over-dispersion Parameter	---	---	---	0.5581	9.63	---	---	---	---	0.528	9.49	---
Summary Statistics												
Sample Size	304			304			304			304		
Log likelihood at Convergence	-1470.852			-938.271			-1395.448			-931.1744		
AIC	2957.704			1894.541			2806.896			1880.349		
BIC	2987.44			1927.994			2836.632			1913.802		

Note: AIC is Akaike Information Criterion, BIC is Bayesian Information Criterion, while MEs indicate marginal effects.

A3.2.2. Machine Learning Techniques

A3.2.2.1. Single Decision Tree Regression

First, we apply single TBR to predict average crash frequency per year on 5T urban and suburban arterials using a training dataset (2013-2015). Using one standard-error rule, we do not observe a significant reduction in error after a tree size of 7 (with cost complexity ~ 0.01858831). Using the mentioned optimal values of tuning parameters, an optimal tree is grown as shown on the right side in Figure A.4. The key predictor variables used in developing the optimal tree include AADT (2013-2015) per year (ln form), segment length "mile" (ln form), the density of minor commercial driveways, and density of major commercial driveways (Figure A.4). Note that the single decision tree is easily interpretable. For instance, if AADT (2013-2015) is greater than 29,964 ($e^{3.4}$ = 29.964 AADT in 1000s) and the segment length is greater than 0.69 miles, the estimated number of crashes on average is 58 (Figure A.4). The optimal TBR model may assign only one of the nine values (5.4, 5.1, 8.1, 22, 29, 16, 42, 20, and 58) of crashes to roadway segments based on the attributes (mean AADT, segment length, the density of major commercial driveways, and density of minor commercial driveways) selected by the optimal tree-based regression model. Note that logarithmic forms of AADT (1000s) and segment length (miles) along with other key covariates (e.g., density of major/minor commercial driveways) in their original forms were used to train the tree-based model. Once the results from tree-based regression were obtained, we took the anti-log of the values of segment length and AADT (1000s) to interpret the results – as shown in Figure A.4).

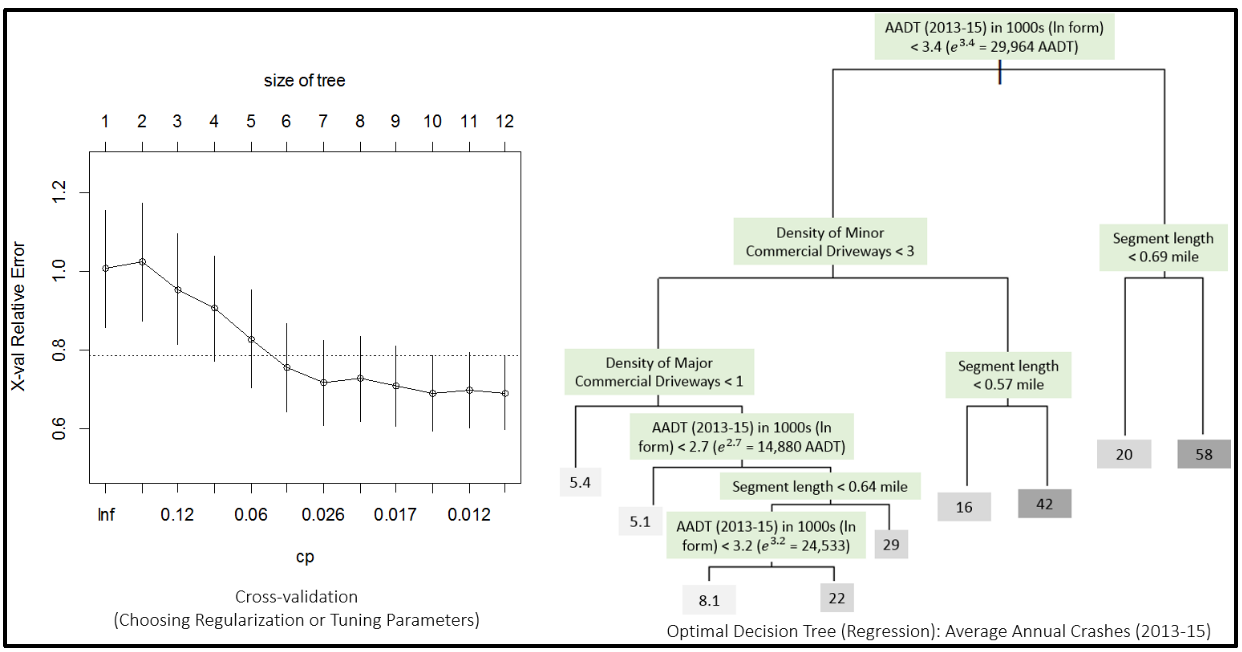


Figure A. 4 Illustration of Cross-validation (Regularization) and Optimal Decision-Tree Regression

A3.2.2.2. Random Forest Regression

To select optimal values of tuning parameters including the number of predictors considered at each split, the number of trees and the maximum number of nodes in the random forest, an extended grid-search optimization, and 10-fold cross-validation procedure were used (Figure A.5). Based on RMSE, our comprehensive grid search indicates that optimal values for the number of predictors considered in each split, number of trees, and number of maximum nodes are found to be 5, 250, and 14 respectively (Figure A.5). Using these tuning parameters, we apply RFR model to predict crash frequency per year using training data (Figure A.5).

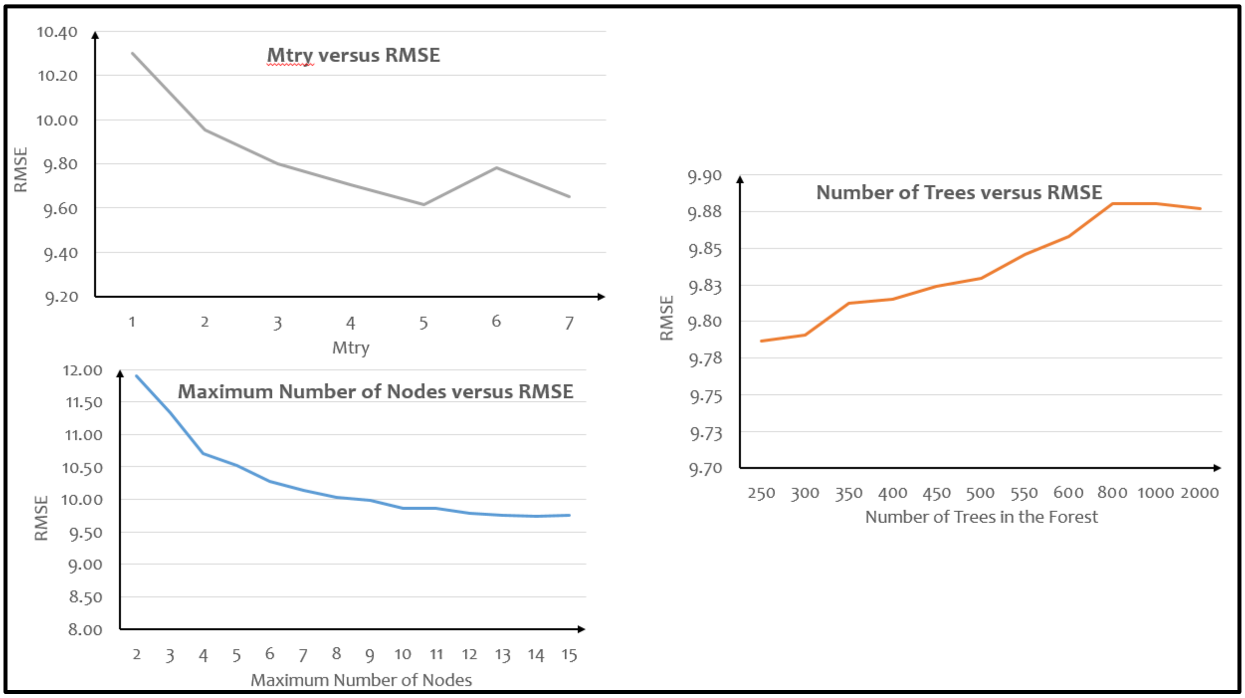


Figure A. 5 Selecting Optimal Values of Regularization Parameters for Random Forest

The relative importance of predictor variables used in the final random forest model is illustrated in Figure A.6. On basis of relative importance, AADT per year (2013-2015) and segment length (mile) are found to be the most important predictor variables (Figure A.6). Similarly, density (number per mile) of minor commercial driveways and density of major industrial/institutional driveways are ranked at 3rd and 4th as per the final RFR model using their relative importance (Figure A.6).

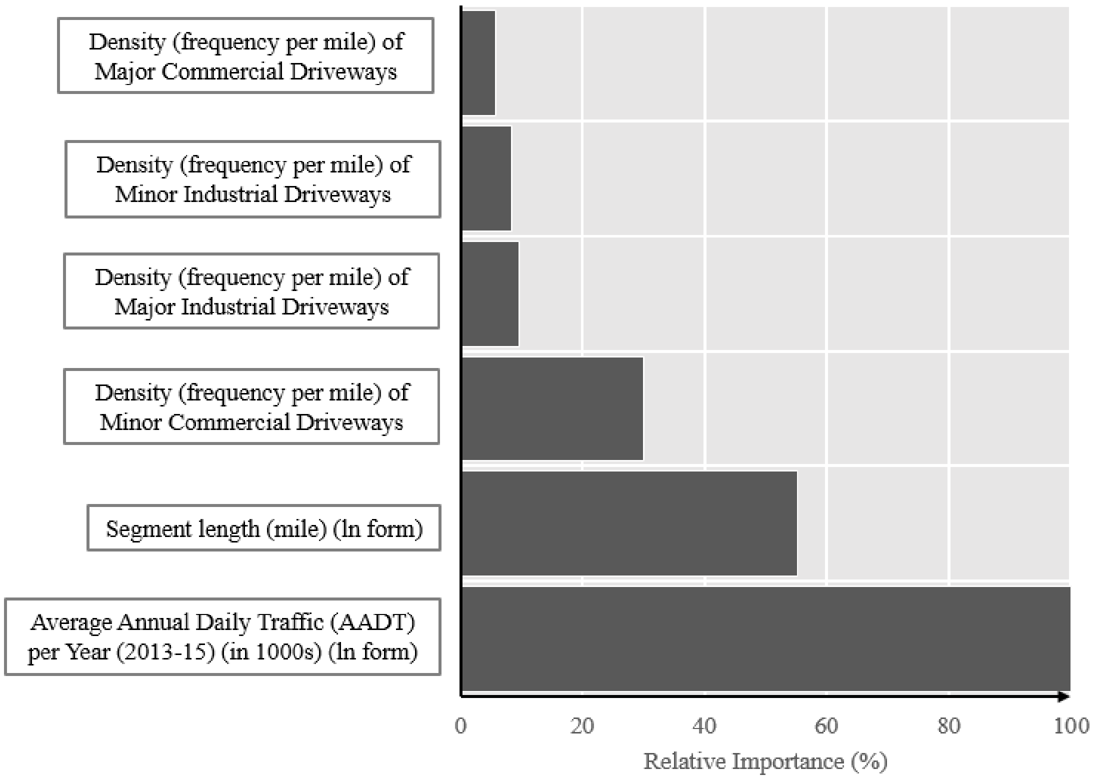


Figure A. 6 Variables Relative Importance Plot: Optimal RFR (Base Learner)

A3.2.2.3. Gradient Boosting Regression

As discussed earlier, GBR is prone to overfitting, which can be minimized while achieving a balance between variance and bias through the selection of optimal regularization parameters, such as the number of trees, learning rate (shrinkage), and complexity parameter (interaction depth). Again, extended grid search and 10-fold cross-validation procedures are used to select optimal values of the regularized parameters. After conducting a grid search with all possible combinations of the number of trees, shrinkage, and interaction depth, a minimum RMSE is achieved when the number of trees, shrinkage, and complexity parameters are equal to 100, 0.1, and 3 respectively (Table A.3). The performance of some key combinations of regularization parameters is shown in Table A.3.

TABLE A. 3 SELECTING OPTIMAL COMBINATION OF REGULARIZATION PARAMETERS FOR GRADIENT BOOSTING

Shrinkage	***Interaction Depth***	***Number of trees***	***RMSE***	***R^2***
0.1	3	100	9.8098	0.5079
0.1	3	100	9.9705	0.5014
0.1	10	100	10.0821	0.4860
0.1	7	100	10.0887	0.4890
0.1	3	100	10.1875	0.4820
0.1	1	100	10.2423	0.4733

Note: The above six combinations are the combinations with smaller RMSE compared to all other combinations. Note that in our grid search, we assigned a range of values to shrinkage (0.1 to 1), interaction depth (1, 3, 7, and 10), and the number of trees (100, 300, 500, 1000).

Once the optimal values of the regularization are determined, a final GBR model is trained. The relative importance of key variables in predicting crash frequency per year on 5T segments of urban and suburban arterials is shown in Figure A.7. Similar to the RFR model, average three-years AADT (2013-2015) in 1000s and segment length (mile) are the most important predictor variables (Figure A.7). Moreover, the density of minor commercial driveways and average offset distance to roadside fixed objects are ranked as 3rd and 4th in terms of their relative importance in predicting crash frequency (Figure A.7).

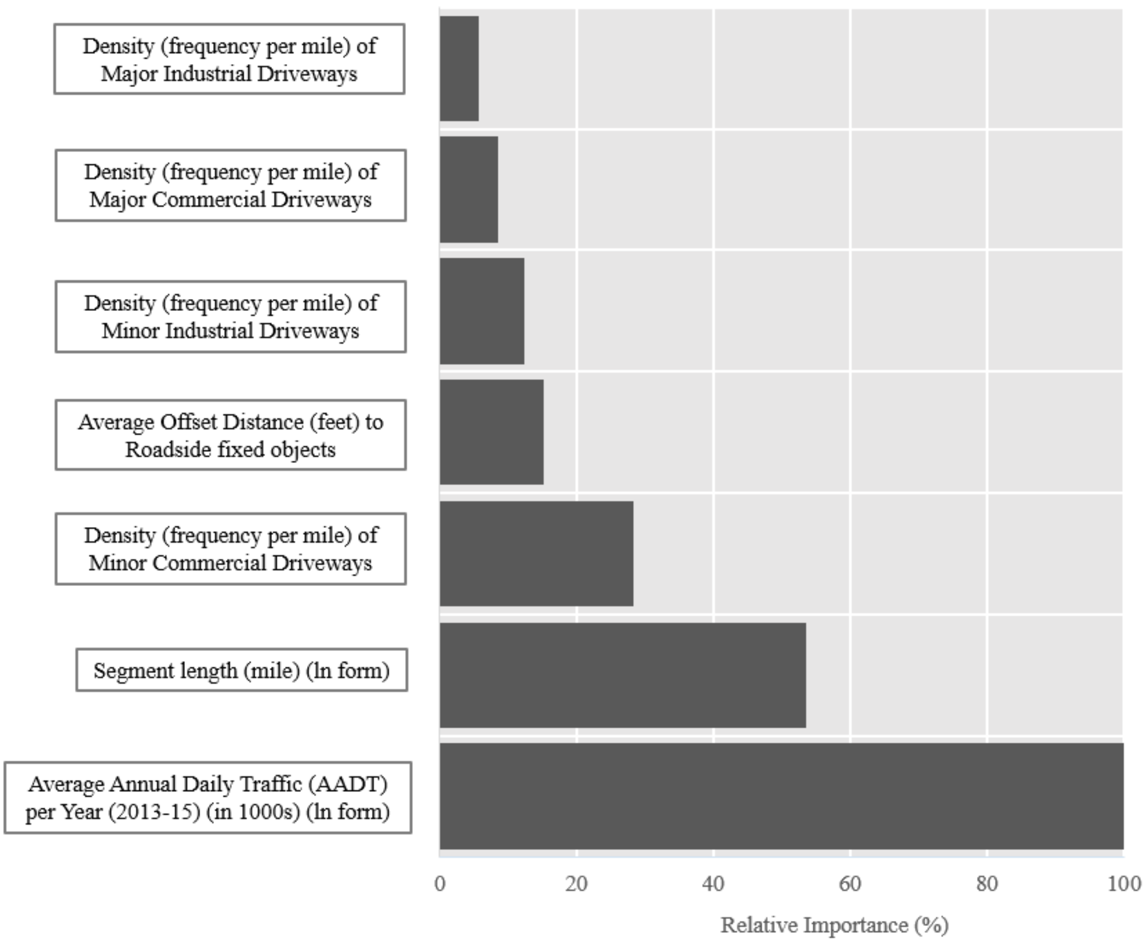
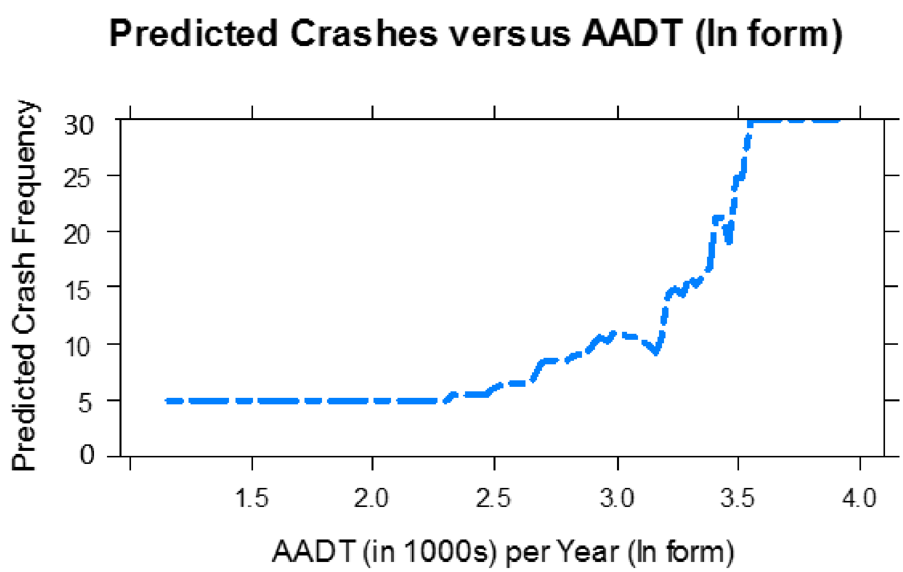


Figure A. 7 Variables Relative Importance Plot: Optimal GBR (Base Learner)

Note that GBR (the best performing base-learner) provides variable importance but does not show the magnitude or nature of the relationship between the response outcome and specific explanatory variables [66]. We present the partial dependence plots, which are similar to marginal effects in statistical models, for the two key variables, AADT and segment length (Figure A.8). For consistency with the best statistical model (details can be found in Section 3.2.1 and Table A.2), we used the natural log forms of segment length and AADT per year (1000s), respectively in all statistical and ML base-learners. The partial dependence plots reveal a non-linear association of average yearly AADT and segment length with average yearly crash frequency (Figure A.8). For instance, there is a sharp increase in crash frequency beyond an AADT of 3000 (Figure A.8). With higher average yearly AADT, the frequency of average crashes (including both injury and non-injury crashes) per year increases. While previous studies reveal that total crash frequency increases with AADT [1], the interesting aspect of the current study is that it captures non-linearities in such a relationship through ML methods. Referring to the partial dependence plots of segment length, the values of -2.3026 and 0.5928 along the x-axis indicate a segment length of 0.1 (= $e^{-2.3026}$) and 1.809 (= $e^{0.5928}$) miles respectively (Table A.1). From the plots, predicted crashes per year (GBR) increase with an increase in yearly AADT between 9,974 (= $e^{2.3}$) and 33,115 (= $e^{3.5}$). Interestingly, if yearly AADT decreases or increases beyond the values of 9,974 and 33,115 respectively, the number of predicted crashes by optimal GBR base-learners remain constant (5 and 30 crashes per year respectively) (Figure A.8). Similarly, predicted crashes by GBR increase with segment length till 0.8187 (= $e^{-0.2}$) beyond which it exhibits a constant pattern no matter if segment length increases (Figure A.8).



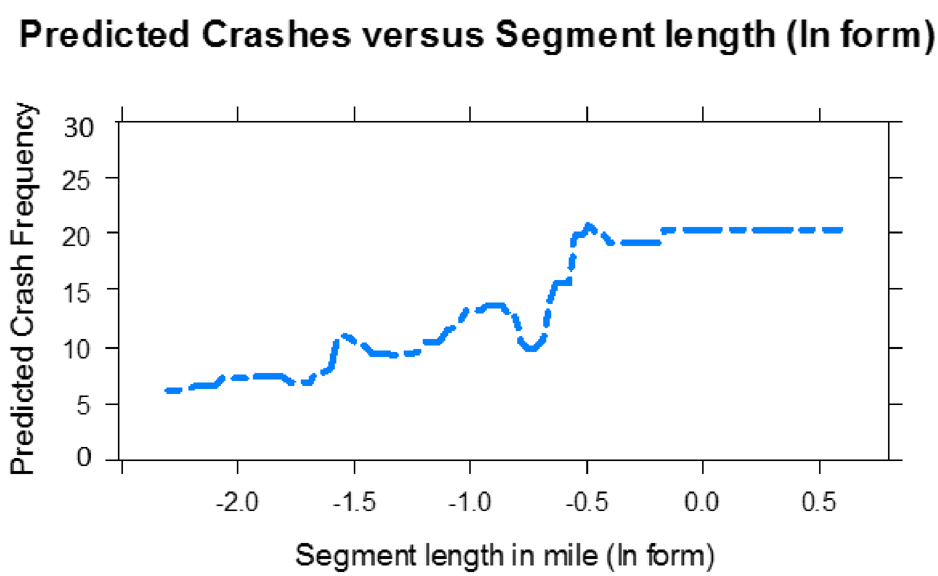


Figure A. 8 Yearly Predicted Crashes by GBR (best performing base-learner) for AADT and Segment length

A3.2.3. Stacking

After the five individual count data and machine learning-based models are developed using training data (2013-2015), the performance is evaluated using a validation dataset (2016). In the next step, the stacked model is trained on the validation dataset (2016) for which observed crash frequency (2016) is used as a response variable. Eventually, the predictions obtained from the five base-learners applied to the validation dataset are used as inputs (predictors) to train the stacked model. Descriptive statistics of predicted and observed crashes for the validation dataset are shown in Table A.4. Note that the mean number of crashes (2016) predicted by individual models such as count data models (Poisson and Negative Binomial model) and machine learning models such as TBR, RFR, and GBR (P_3, P_4, and P_5 respectively) are very similar to the mean number of observed crashes occurred during 2016 (Table A.4). While using the validation dataset including five new predicted values (P_1, P_2, ..., P_5) and observed crashes, we train an RFR model as a meta-learner (stacked ensemble model) in second-stage regression. Several techniques ranging from a simple linear regression to more robust ensemble methods like RFR and GBR can be used to train the stacked model.

The three ML methods (TBR, RFR, and GBR) were used as stacking meta-learners in the second-stage regression to predict crashes using the optimal combination of the base-learners. Our findings suggest that stacking meta-learners including RFR and GBR significantly reduced the out-of-sample RMSE, and MAE compared to homogeneous ensembles (RFR and GBR) used as base-learners (Table A.5). However, TBR when used as a stacking meta-learner, could not outperform the two homogeneous ensembles used as base-learners (Table A.5). Briefly, the reason is that the optimal TBR (used as a stacking meta-learner) assigns one of the nine values (4, 10, 18, 27, 28, 42, 53, 82, and 90) of crashes to roadway segments. While the observed crash frequency on roadway segments (in the sample used for analysis) has a wide range, the optimal TBR, irrespective of the range of predictor variables, was restricted to these nine values. This affects the accuracy of out-of-sample prediction. As mentioned, we used the three ML methods (TBR, RFR, and GBR) as meta-learners to predict crashes; however, we present and discuss the results of RFR as a meta-learner (stacked ensemble method) because it led to maximum improvement in out-of-sample prediction accuracy.

Similar to individual machine learning models (TBR, RFR, and GBR), grid search optimization and 10-fold cross-validation procedures were used to select optimal values for regularization parameters of the Stacked model (RFR-meta learner). The tuning parameters in the random forest model include the number of predictors considered at each split, number of trees, and maximum number of nodes in the random forest, which were found to be 2, 9, and 1000 respectively (Figure A.9). To select the best number of nodes, an initial grid search was specified with a range of 2 to 15; the RMSE value was at a minimum when the maximum number of nodes equaled 9 (Figure A.9). Similarly, to select the optimal number of trees, a grid-search with the range (250, 300, 350, 400, 450, 500, 550, 600, 800, 1000, 2000) was applied and showed that 1000 trees led to the lowest RMSE (Figure A.9).

TABLE A. 4 DESCRIPTIVE STATISTICS OF PREDICTED AND OBSERVED CRASHES (VALIDATION DATASET: 2016)

Variables	*Obs.*	*Mean*	*Std. Dev.*	*Min*	*Max*
Total Crashes (2016)	304	11.010	14.651	0.000	90.000
Predicted Crashes via Poisson Model (P_1)	304	11.357	10.869	0.754	70.664
Predicted Crashes per Negative Binomial Model (P_2)	304	11.430	12.119	0.839	97.733
Predicted Crashes per Decision Tree Model (P_3)	304	10.898	10.521	5.067	58.143
Predicted Crashes per Random Forest Model (P_4)	304	11.202	9.310	2.002	49.383
Predicted Crashes per Gradient Boosting Model (P_5)	304	11.345	11.359	0.000	53.928

The relative importance plot of the predictors (obtained from the five base-learners) for meta-learner (stacked ensemble model) is shown in Figure A.10. The predicted crashes obtained from the individual RFR model (P_4) is found to be the most important predictor variable followed by those predicted via gradient boosting (P_5) (importance = 75.45%), negative binomial model (P_2), and Poisson model (P_1) (Figure A.10). Note that the prediction outcome of the single-decision tree (P_3) was not used to determine crash frequency via the stacked RFR model (Figure A.10). Specifically, no weight is assigned to the predicted values by the TBR model (P_3) in the stacked model because there is no significant variation in P_3. The optimal TBR model assigns one of the 9 different values (5.4, 5.1, 8.1, 22, 29, 16, 42, 20, and 58) of crashes to segment(s) based on the attributes (shown in Figure A.4) which were considered by the algorithm developing an optimal tree regression model.

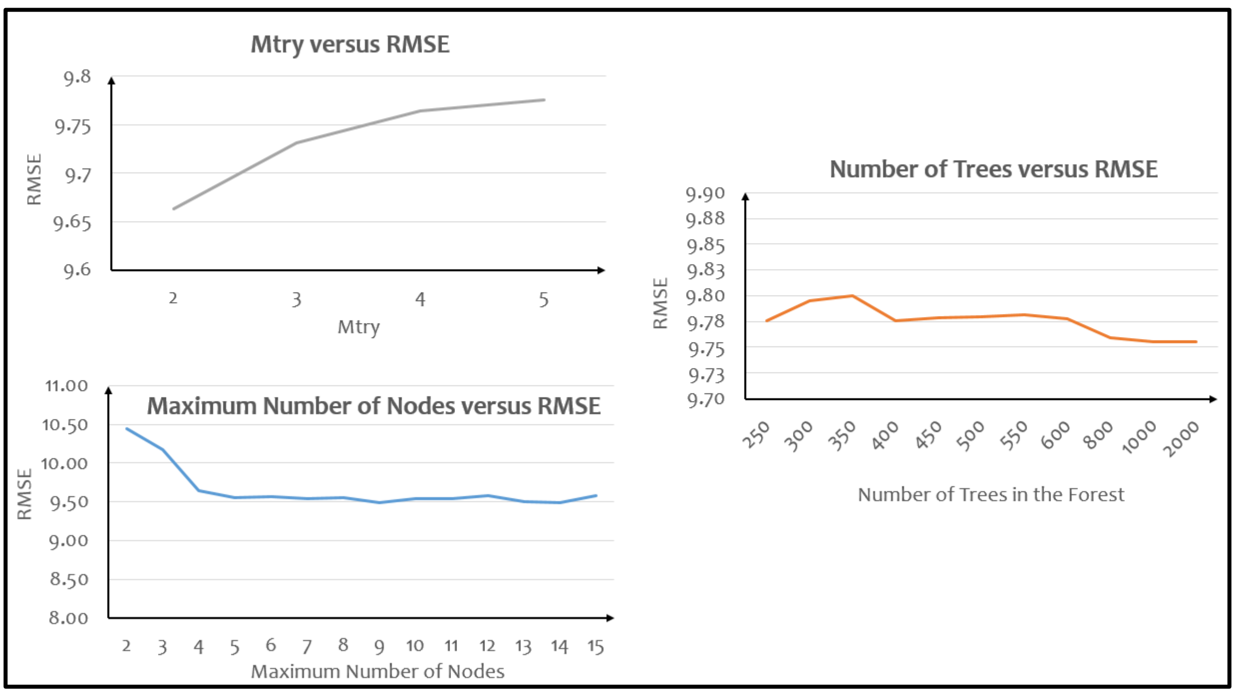


Figure A. 9 Selecting Optimal Tuning Parameters for Stacked RFR Model (Second-Stage Regression)

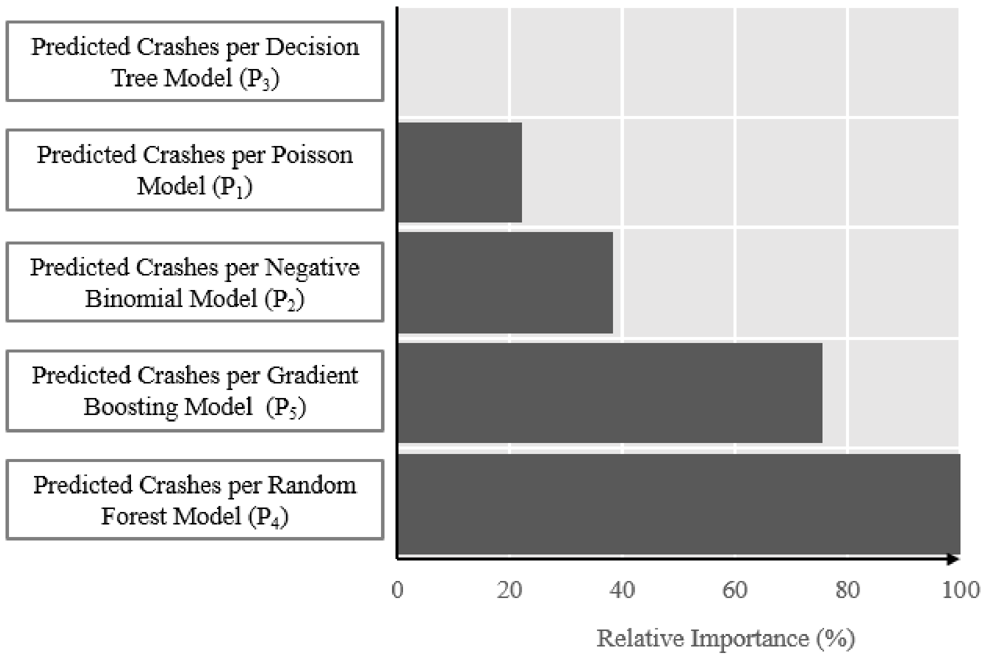


Figure A. 10 Variables Relative Importance Plot: Optimal Stacked RFR Model (Meta Learner)

A3.2.4. Comparing Out-of-Sample Prediction Performance

For evaluating the out-of-sample prediction performance of the stacked vs. un-stacked models, crash data in 2017 are used – this data were neither used to train base-learners nor the meta-learner (stacked model). Prior to comparing the out-of-sample prediction performance of various stacking meta-learners, notice that each of the three ML methods (TBR, RFR, and GBR) were used as stacking meta-learners. To select optimal values of regularization parameters for a particular ML meta-learner, we used the same procedure when the method was used as an ML base-learner. Using the optimal values of specific regularization parameters obtained through 10-fold cross-validation and extended grid-search, the three ML methods including TBR, RFR, and GBR were trained as stacking meta-learners. To compare the predictive performance of the five base-learners and meta-learners based on the new dataset, we computed out-of-sample RMSE and MAE (Table A.5). Our findings indicate that GBR has the lowest out-of-sample RMSE and MAE among all base-learners (Table A.5). Referring to the predictive performance of meta-learners, both RFR and GBR as stacking meta-learners further reduced out-of-sample RMSE and MAE compared to the best performing base-learner (GBR) (Table A.5). However, TBR as a stacking meta-learner showed poor out-of-sample prediction performance compared to the best performing base-learners (GBR) as well as RFR base-learner (Table A.5). This was expected as the optimal TBR model (when used a stacking meta-learner) could assign one of the nine values (i.e., 4, 10, 18, 27, 28, 42, 53, 82, and 90) of crashes to roadway segments. To conclude, RFR, as a stacking meta-learner, is found to have the lowest out-of-sample RMSE and out-of-sample MAE among all the base-learners and meta-learners and is selected as the best performing model for out-of-sample crash prediction (Table A.5). For brevity, we only discuss the results of RFR as a stacking meta-learner in the paper.

TABLE A. 5 COMPARISON PREDICTION PERFORMANCE (OUT-OF-SAMPLE): RMSE AND MAE

Model	***Type***	***RMSE***	***% Difference in RMSE compared to GBR Model***	***MAE***	***% Difference in MAE compared to GBR Model***
TBR	Meta Learner	9.515	6.98	5.889	0.43
Gradient Boosting	Meta Learner	8.404	-5.51	5.609	-4.33
Random Forest	Meta Learner	8.312	-6.54	5.383	-8.19
Poisson	Base Learner	10.123	13.82	6.315	7.71
Negative Binomial	Base Learner	11.118	25.01	6.589	12.38
TBR	Base Learner	10.251	15.26	6.757	15.25
Random Forest	Base Learner	9.023	1.45	5.951	1.50
Gradient Boosting	Base Learner	8.894	Base	5.863	Base

Note: The percent difference in RMSE compared to GBR model and the percent difference in MAE compared to GBR model are calculated using the following equations:

$$\text{\% Difference in RMSE compared to GBR Model} = \frac{(\text{RMSE}_{Model\,X} - \text{RMSE}_{GBR\,Model})}{\text{RMSE}_{GBR\,Model}} * 100\%$$

$$\text{\% Difference in MAE compared to GBR Model} = \frac{(\text{MAE}_{Model\,X} - \text{MAE}_{GBR\,Model})}{\text{MAE}_{GBR\,Model}} * 100\%$$

Note: Meta-learner refers to the model when it is applied in the second stage to use predictions from the optimal combinations of different base learners. In stacking, "meta-leaner" is also termed as "super-learner" [42].

To have a deeper understanding of the out-of-sample prediction errors, we also provide distributional statistics of out-of-sample absolute prediction error (Table A.6). RFR as a stacking meta-learner leads to the lowest out-of-sample absolute prediction error (Table A.6). The standard deviations of absolute prediction errors for GBR and RFR as stacking meta-learners are found to be the lowest indicating less spreading out around the mean value of the error (Table A.6).

TABLE A. 6 SUMMARY OF ABSOLUTE PREDICTION ERRORS (OUT-OF-SAMPLE) FOR BASE AND META LEARNERS

Model	***Type (used as)***	***N***	***Absolute (observed crashes - predicted crashes)***			
			Mean	***Std. Dev.***	***Minimum***	***Maximum***
TBR	Meta Learner	304	5.889	7.487	0.119	81.000
Gradient Boosting	Meta Learner	304	5.609	6.268	0.015	59.010
Random Forest	Meta Learner	304	5.383	6.345	0.043	62.137
Poisson	Base Learner	304	6.315	7.926	0.033	77.058
Negative Binomial	Base Learner	304	6.589	8.970	0.008	82.822
TBR	Base Learner	304	6.757	7.722	0.038	78.333
Random Forest	Base Learner	304	5.951	6.794	0.046	71.128
Gradient Boosting	Base Learner	304	5.863	6.698	0.000	68.084

Note: Meta-learner refers to the model when it is applied in the second stage to use predictions from the optimal combinations of different base learners. In stacking, "meta-leaner" is also termed as "super-learner" [42].

To visualize and compare the out-of-sample prediction performance of individual models (base-learners) and the stacked ensemble technique (meta-learner), we present plots of predicted versus observed crashes based on the testing data (2017) (Figure A.11). Notably, the RFR ensemble model, when used as a meta-learner, shows the best fit, followed by individual GBR (base learner) and RFR (base learner) as shown in Figure A.11. Similar findings are obtained in other fields where prediction accuracy for the stacked ensemble model (used for classification) improved by 2%-4% [40]. To conclude, we found that the application of the stacked ensemble technique can help in obtaining more accurate crash predictions in the future. Note that none of the studies have evaluated the applicability of more accurate, reliable, and intelligent heterogeneous ensemble procedures to determine the crash frequency. Similar rigorous stacked ensemble techniques can be used in determining or predicting crash frequency on other types of roadways using local data.

The plot of predicted versus actual crash frequency for the tree-based regression model seems unusual compared to the other five regression models. Note that tree-based regression models only assign a specific number of values to response outcomes for individual observations (roadway segments in this case) based on their attribute values and conditions (for details, please refer to Figure A.4) assigned by the optimal tree-based regression model. In our case, the optimal tree-based regression model indicates that one of the nine values (5.4, 5.1, 8.1, 22, 29, 16, 42, 20, and 58) of crashes may be assigned to any segment based on its attributes (mean AADT, segment length, density of major commercial driveways, and density of minor commercial driveways) based on which the optimal TBR model was developed. Hence, it can be seen in Figure A.11 that crashes are spotted at a few specific points by the TBR model, which seems unusual compared to the remaining models.

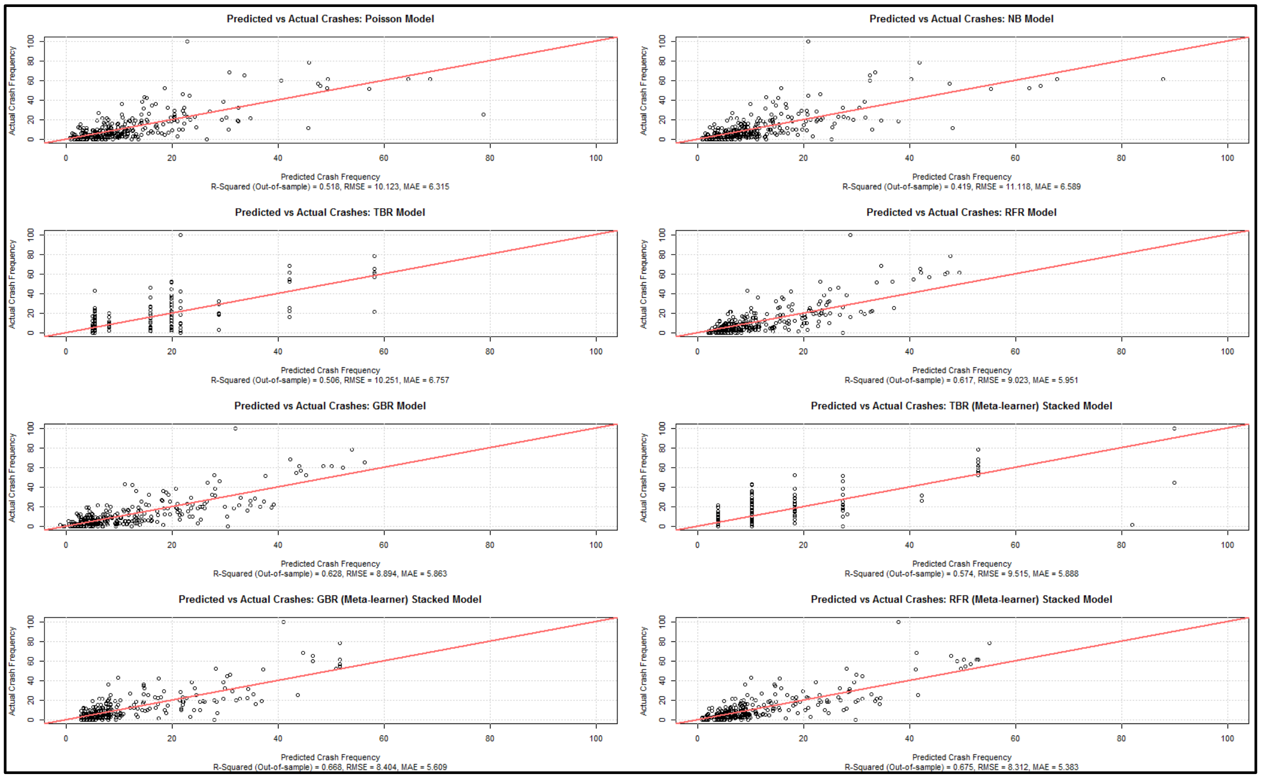


Figure A. 11 Out-of-sample prediction: Observed versus Predicted Crashes

A4. Limitations and Future Directions

This study uses 5T segments of urban and suburban arterials in Tennessee and may not be extended to other states due to variations in driving behavior, socio-demographic, and roadway conditions. Compared to individual machine learning techniques, we recommend using heterogeneous ensemble methods like stacking which are more accurate, reliable, and intelligent techniques and can help in accessing crash forecasts in the future. While stacking may significantly improve the out-of-sample prediction accuracy, it does not provide the variable importance for the actual predictor variables (e.g., segment length and AADT). Note that in this study, stacking is applied to combine multiple predictions (as opposed to combining distributions of coefficients, such as in the Bayesian setup). Thus, the inference is not relevant in Stage-2. However, the inferences are provided in Stage-1 by individual base-learners that include statistical models (like Poisson and negative binomial) and ML methods including TBR, RFR, and GBR. Note that based on our study objectives, we split five years of crash data into training (2013-2015), validation (2016), and testing (2017) datasets - indicting that only crash frequency and AADT may vary across the datasets while roadway geometry remains similar. In the future, data splitting can also be done using standard splitting procedures rather than the year-wise split, depending on study design and objectives. The application presented herein is based on year-wise splits, assuming temporal transferability of the models over years. As part of future work, variants of the methods presented herein that relax this assumption can be examined.

A5. Conclusions

Safety performance functions are core tools necessary for the accurate prediction of crashes and subsequent development of place-based countermeasures. Traditional count data models and machine learning methods have been extensively used in the safety literature for the development of statistical relationships between crash frequency and associated factors. This study contributes by presenting a rigorous and novel heterogeneous ensemble methods (HEM) scheme to "stack" predictions from competing frequentist and ML models – eventually leading to a more accurate prediction of crashes. By using a more accurate and reliable intelligent pattern recognition scheme, the "Stacking" methodology harnesses the inferential framework provided by traditional count data models and the predictive power offered by ML methods. The objectives are achieved using 5-years crash, traffic, and roadway geometric data for urban and suburban arterials extracted from the Enhanced Tennessee Roadway Information Management System (E-TRIMS). To the best of the authors' knowledge, no study to date has applied heterogeneous ensemble methods to pool multiple predictions from frequentist and ML methods.

The results suggest the significant potential of "Stacking" in providing more accurate predictions by heterogeneously assembling crash forecasts from individual statistical (Poisson and negative binomial) and machine-learning based base-learners (tree-based regression, random forests, and gradient boosting regression). Using out-of-sample prediction performance, the gradient boosting model led to the lowest RMSE and MAE values among all the individual base-learners. While individual ML-based base learners can provide greater predictive accuracy, there is no escaping the relationship between bias and variance underpinning most machine learning models. In other words, using a single supervised or unsupervised ML method could lead to relatively less accurate predictions due to the compromised bias or variance. By superimposing a machine-learning based meta learner on predictions obtained from the five statistical and ML based base-learners, the RMSE and MAE values of crash forecasts were further reduced by 6.54% and 8.19% respectively compared to the prediction accuracy of the best-fit gradient boosting based individual base-learner. From an inferential standpoint, the individual base-learners offer insights into the links between crash frequency and associated factors. Count data models show that besides exposure variables (AADT and segment length), higher accessibility correlates with higher crash frequency. Contrarily, a larger offset distance to fixed object correlates with lower crash frequency. In terms of variable importance, the three ML-based base-learners rank AADT, segment length, and density of minor commercial driveways as the three top predictors of crash frequency.

The results of this study have important implications. By using heterogeneous ensemble methods such as Stacking, even more accurate crash forecasts can be obtained compared to those obtained from individual frequentist or ML methods. With more accurate crash forecasts, roadway segments can be better prioritized in terms of the need for place-based safety countermeasures. From a practical standpoint, the straight-forward heterogeneous ensemble method technique can be easily automated for more accurate crash prediction. From a research perspective, the methodology can be expanded by other researchers to include an even broader set of ML methods or consider more rigorous simulation-assisted statistical methods accounting for methodological issues like observed and unobserved heterogeneity.

A6. Acknowledgments

This study was funded by the Tennessee Department of Transportation (Grant No. RES2020-04) and the U.S. DOT through the Collaborative Sciences Center for Road Safety; a consortium led by The University of North Carolina at Chapel Hill in partnership with The University of Tennessee. The views in this paper are solely those of the authors who are responsible for the content of this publication. The authors also wish to sincerely thank Amin Mohammadnazar (Graduate Student in Civil Engineering at University of Tennessee, Knoxville) for his help in collecting a portion of the data used in this study. The authors also extend their sincere gratitude to Ms. Meredith King, undergraduate student in Civil and Environmental Engineering at University of Tennessee (Knoxville), for proofreading.

A7. Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Numan Ahmad, Behram Wali, Asad Khattak; data collection: Numan Ahmad; analysis and interpretation of results: Numan Ahmad, Behram Wali; draft manuscript preparation Numan Ahmad, Behram Wali, Asad Khattak. All authors reviewed the results in the final version of the manuscript.

A8. References

1. AASHTO, Highway Safety Manual, in Washington, DC. 2010.
2. Wali, B., et al. Developing Safety Performance Functions for Rural Multilane Highways in Tennessee: Accounting for Unobserved Heterogeneity. in Transportation Research Board 99th Annual Meeting Transportation Research Board. 2020.
3. Wali, B., et al., Development of safety performance functions: incorporating unobserved heterogeneity and functional form analysis. Transportation research record, 2018. 2672(30): p. 9-20.
4. Zhong, C., et al. Crash prediction on rural roads. in Proceeding of the 3rd International Conference on Road Safety and Simulation, Indianapolis, Indiana. 2011.
5. Thakali, L., L. Fu, and T. Chen, Model-based versus data-driven approach for road safety analysis: do more data help? Transportation research record, 2016. 2601(1): p. 33-41.
6. Shankar, V., F. Mannering, and W. Barfield, Effect of roadway geometrics and environmental factors on rural freeway accident frequencies. Accident Analysis & Prevention, 1995. 27(3): p. 371-389.
7. Srinivasan, R. and D. Carter, Development of safety performance functions for North Carolina. 2011, North Carolina. Dept. of Transportation. Research and Analysis Group.
8. Caliendo, C., M. Guida, and A. Parisi, A crash-prediction model for multilane roads. Accident Analysis & Prevention, 2007. 39(4): p. 657-670.
9. Hauer, E., F.M. Council, and Y. Mohammedshah, Safety models for urban four-lane undivided road segments. Transportation research record, 2004. 1897(1): p. 96-105.
10. Abdel-Aty, M.A. and A.E. Radwan, Modeling traffic accident occurrence and involvement. Accident Analysis & Prevention, 2000. 32(5): p. 633-642.
11. Pan, G., L. Fu, and L. Thakali, Development of a global road safety performance function using deep neural networks. International journal of transportation science and technology, 2017. 6(3): p. 159-173.
12. Saha, D., P. Alluri, and A. Gan, Prioritizing Highway Safety Manual's crash prediction variables using boosted regression trees. Accident Analysis & Prevention, 2015. 79: p. 133-144.
13. Breiman, L., et al., Classification and regression trees. 1984: CRC press.
14. Karlaftis, M. and I. Golias, Effects of road geometry and surface on speed and safety. Accid. Anal. Prev, 2002. 34(34): p. 357-365.
15. Zhang, M., et al., Application of boosting to classification problems in chemometrics. Analytica Chimica Acta, 2005. 544(1-2): p. 167-176.
16. Das, A., M. Abdel-Aty, and A. Pande, Using conditional inference forests to identify the factors affecting crash severity on arterial corridors. Journal of safety research, 2009. 40(4): p. 317-327.
17. De'Ath, G., Boosted trees for ecological modeling and prediction. Ecology, 2007. 88(1): p. 243-251.
18. Dietterich, T.G., An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. Machine learning, 2000. 40(2): p. 139-157.

19. Thapa, R., et al., Use of geospatial technology for delineating groundwater potential zones with an emphasis on water-table analysis in Dwarka River basin, Birbhum, India. Hydrogeology Journal, 2018. 26(3): p. 899-922.
20. Tewari, S. and U. Dwivedi, A comparative study of heterogeneous ensemble methods for the identification of geological lithofacies. Journal of Petroleum Exploration and Production Technology, 2020: p. 1-20.
21. Bhatt, S., et al., Improved prediction accuracy for disease risk mapping using Gaussian process stacked generalization. Journal of The Royal Society Interface, 2017. 14(134): p. 20170520.
22. Sigletos, G., et al., Combining information extraction systems using voting and stacked generalization. Journal of Machine Learning Research, 2005. 6(Nov): p. 1751-1782.
23. Zong, F., et al., Analyzing traffic crash severity with combination of information entropy and Bayesian network. IEEE Access, 2019. 7: p. 63288-63302.
24. Tang, J., et al., Crash injury severity analysis using a two-layer Stacking framework. Accident Analysis & Prevention, 2019. 122: p. 226-238.
25. Ghandour, A.J., H. Hammoud, and S. Al-Hajj, Analyzing Factors Associated with Fatal Road Crashes: A Machine Learning Approach. International Journal of Environmental Research and Public Health, 2020. 17(11): p. 4111.
26. Zahid, M., et al., Predicting Risky and Aggressive Driving Behavior among Taxi Drivers: Do Spatio-Temporal Attributes Matter? International Journal of Environmental Research and Public Health, 2020. 17(11): p. 3937.
27. Bugusa, Y. and S. Patil, An Improved Accident Crash Risk Prediction Model Based On Driving Outcomes Using Ensemble Of Prediction Algorithms. INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH, 2019. 8(11): p. 603-611.
28. Rooney, N., et al. Random Subspacing for Regression Ensembles. in FLAIRS Conference. 2004.
29. Chali, Y., S.A. Hasan, and M. Mojahid. Complex question answering: homogeneous or heterogeneous, which ensemble is better? in International Conference on Applications of Natural Language to Data Bases/Information Systems. 2014. Springer.
30. Fernández-Alemán, J.L., et al. Homogeneous and heterogeneous ensemble classification methods in diabetes disease: a review. in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2019. IEEE.
31. Sabzevari, M., G. Martínez-Muñoz, and A. Suárez, Pooling homogeneous ensembles to build heterogeneous ones. arXiv preprint arXiv:1802.07877, 2018.
32. Elish, M.O., T. Helmy, and M.I. Hussain, Empirical study of homogeneous and heterogeneous ensemble models for software development effort estimation. Mathematical Problems in Engineering, 2013. 2013.
33. Petrakova, A., M. Affenzeller, and G. Merkurjeva, Heterogeneous versus homogeneous machine learning ensembles. Information Technology and Management Science, 2015. 18(1): p. 135-140.
34. Osareh, A. and B. Shadgar, An efficient ensemble learning method for gene microarray classification. BioMed Research International, 2013. 2013.
35. Wang, K., et al., How the choice of safety performance function affects the identification of important crash prediction variables. Accident Analysis & Prevention, 2016. 88: p. 1-8.

36. Farid, A., M. Abdel-Aty, and J. Lee, Comparative analysis of multiple techniques for developing and transferring safety performance functions. Accident Analysis & Prevention, 2019. 122: p. 85-98.
37. Iranitalab, A. and A. Khattak, Comparison of four statistical and machine learning methods for crash severity prediction. Accident Analysis & Prevention, 2017. 108: p. 27-36.
38. Yu, R. and M. Abdel-Aty, Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. Safety science, 2014. 63: p. 50-56.
39. Wolpert, D.H., Stacked generalization. Neural networks, 1992. 5(2): p. 241-259.
40. Güneş, F., R. Wolfinger, and P.-Y. Tan. Stacked ensemble models for improved prediction accuracy. in Proc. Static Anal. Symp. 2017.
41. McLachlan, G.J., K.-A. Do, and C. Ambroise, Analyzing microarray gene expression data. Vol. 422. 2005: John Wiley & Sons.
42. Van der Laan, M.J., E.C. Polley, and A.E. Hubbard, Super learner. Statistical applications in genetics and molecular biology, 2007. 6(1).
43. Anastasopoulos, P.C. and F.L. Mannering, A note on modeling vehicle accident frequencies with random-parameters count models. Accident Analysis & Prevention, 2009. 41(1): p. 153-159.
44. Washington, S., Karlaftis,. MG, and Mannering, FL, Statistical and Econometric Methods for Transportation Data Analysis. 2010, Chapman & Hall/CRC New York.
45. Poch, M. and F. Mannering, Negative binomial analysis of intersection-accident frequencies. Journal of transportation engineering, 1996. 122(2): p. 105-113.
46. Bozdogan, H., Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 1987. 52(3): p. 345-370.
47. Akaike, H., B.N. Petrov, and F. Csaki, Second international symposium on information theory. 1973, Akadémiai Kiadó, Budapest.
48. Schwarz, G., Estimating the dimension of a model. The annals of statistics, 1978. 6(2): p. 461-464.
49. Torgo, L.F.R.A., Inductive learning of tree-based regression models. 1999.
50. Breiman, L., Bagging predictors. Machine learning, 1996. 24(2): p. 123-140.
51. Louppe, G., et al. Understanding variable importances in forests of randomized trees. in Advances in neural information processing systems. 2013.
52. Malekipirbazari, M. and V. Aksakalli, Risk assessment in social lending via random forests. Expert Systems with Applications, 2015. 42(10): p. 4621-4631.
53. Liaw, A. and M. Wiener, Classification and regression by randomForest. R news, 2002. 2(3): p. 18-22.
54. Hastie, T., R. Tibshirani, and J. Friedman, The elements of statistical learning: data mining, inference, and prediction. 2009: Springer Science & Business Media.
55. Chung, Y.-S., Factor complexity of crash occurrence: An empirical demonstration using boosted regression trees. Accident Analysis & Prevention, 2013. 61: p. 107-118.
56. Heung, B., C.E. Bulmer, and M.G. Schmidt, Predictive soil parent material mapping at a regional-scale: a random forest approach. Geoderma, 2014. 214: p. 141-154.
57. Strobl, C., et al., Conditional variable importance for random forests. BMC bioinformatics, 2008. 9(1): p. 307.

58. Elith, J. and J. Leathwick, Boosted Regression Trees for ecological modeling. R Documentation. Available online: https://cran. r-project. org/web/packages/dismo/vignettes/brt. pdf (accessed on 12 June 2011), 2017.
59. Washington, S.P., M.G. Karlaftis, and F.L. Mannering, Statistical and econometric methods for transportation data analysis. 2010: CRC press.
60. Safety, T.O.T.F.f.R., Roadside design guide. 2011: AASHTO.
61. Albuquerque, F.D. and D.M. Awadalla, Roadside design assessment in an urban, low-density environment in the Gulf Cooperation Council region. Traffic Injury Prevention, 2019. 20(4): p. 436-441.
62. Albuquerque, F.D.B.d. and D.M. Awadalla, Roadside Fixed-Object Collisions, Barrier Performance, and Fatal Injuries in Single-Vehicle, Run-Off-Road Crashes. Safety, 2020. 6(2): p. 27.
63. Wolf, K.L., Urban trees and traffic safety: Considering the US roadside policy and crash data. Arboriculture and Urban Forestry. 32 (4): 170-179., 2006. 32(4): p. 170-179.
64. Dixon, K., et al., Quantifying Safety Performance of Driveways on State Highways. 2012, Oregon. Dept. of Transportation. Research Section.
65. Khattak, A., et al., HIGHWAY SAFETY MANUAL SAFETY PERFORMANCE FUNCTIONS & ROADWAY CALIBRATION FACTORS: ROADWAY SEGMENTS PHASE 2, PART. 2020.
66. Friedman, J.H., Greedy function approximation: a gradient boosting machine. Annals of statistics, 2001: p. 1189-1232.

Appendix B. Comparison or Safety Performance

As discussed in Section 3.4, using TN-SPFs, the key point is that based on the TN-data and estimated SPFs, the predicted crashes on rural 5T roadways were found to be a bit lower than rural 4D roadways when AADT drops below 7,000. Intuitively, we expect that crashes on 5T will remain above 4D for all AADT values, but this is not what the TN data shows us. We checked the actual TN data and if we stay with the standard HSM procedures (negative binomial model) and use the TN data that we collected, then it seems that the expectation of fewer crashes on 4D compared with 5T does not hold in the 4,000 to 7,000 range of AADTs. Though there are some nuances that emerged when we did deeper analysis, as explained below.

- First, we investigated if we have sufficient data in the below 7000 AADT range? The answer is Yes. There are sufficient cases in that range for rural 5T and rural 4D roadways where AADT is below 7,000 in TN data. For rural 4D, we have 125 segments (out of a total of 271) with AADT lower than 7,000 and for rural 5T, we have 76 (out of a total of 205) roadway segments below 7,000.
- Second, are we confident that the sample contains the true population means of crashes for the 2 roadway types? For this, we examined whether the confidence intervals overlap in the lower AADT range. They do for some part in the lower AADT range, as shown in Figures B.1 to B.3 below (which quantify variations in the data). Figures B.1 and B.2 show the predicted crashes and the confidence intervals. Note that the prediction line for rural 4D is almost flat, whereas the line for 5T has a steeper slope, indicating that the results in the lower AADT range may be an artifact of fitting regression lines to the entire dataset. Figure B.3 shows the confidence intervals for 5T and 4D overlayed. The confidence intervals overlap for the 2 lines in a significant portion of the lower AADT range, which suggests that the prediction ranges overlap. And they may not be statistically significantly different for part of the lower range for these 2 types of roadways. Notably, the magnitude of difference is relatively small (about ½ crash per year, i.e., not substantial) in the lower AADT range (up to 4000 AADT).
- Third, we explored whether the relationship between AADT and crashes is non-linear? If so, this could change the predicted crashes for 5T and 4D in the lower AADT range. Note that in some cases, HSM does consider non-linear relationships, so we decided to explore non-linearity. One way to do this is to use AADT and AADT-squared in the model (polynomial form). After using the squared term of the original AADT variable in a negative binomial model, we found evidence of non-linearity (for results, please refer to Table B.1). However, the predicted crashes on rural 5T roadways based on the new model were still lower than the ones on rural 4D roadway segments (see Figure B.4). Furthermore, non-linearity can also be explored using a technique called regression splines. We fitted splines for AADT to capture non-linearity for 5T. However, the predictions for crashes were still lower than rural 4D segments for AADT below 7,000.

After doing a thorough empirical investigation and based on the SPF slopes found (between crashes and AADT), the rural 5T roadways show lower predicted crashes when compared with 4D in the below 7,000 AADT range. Therefore, we recommend keeping the TN-specific models embedded in the analysis spreadsheets (the updated spreadsheets are provided with the report - we have added rural 4D and 4U results for Fatal and Injury crashes (as TDOT had requested).

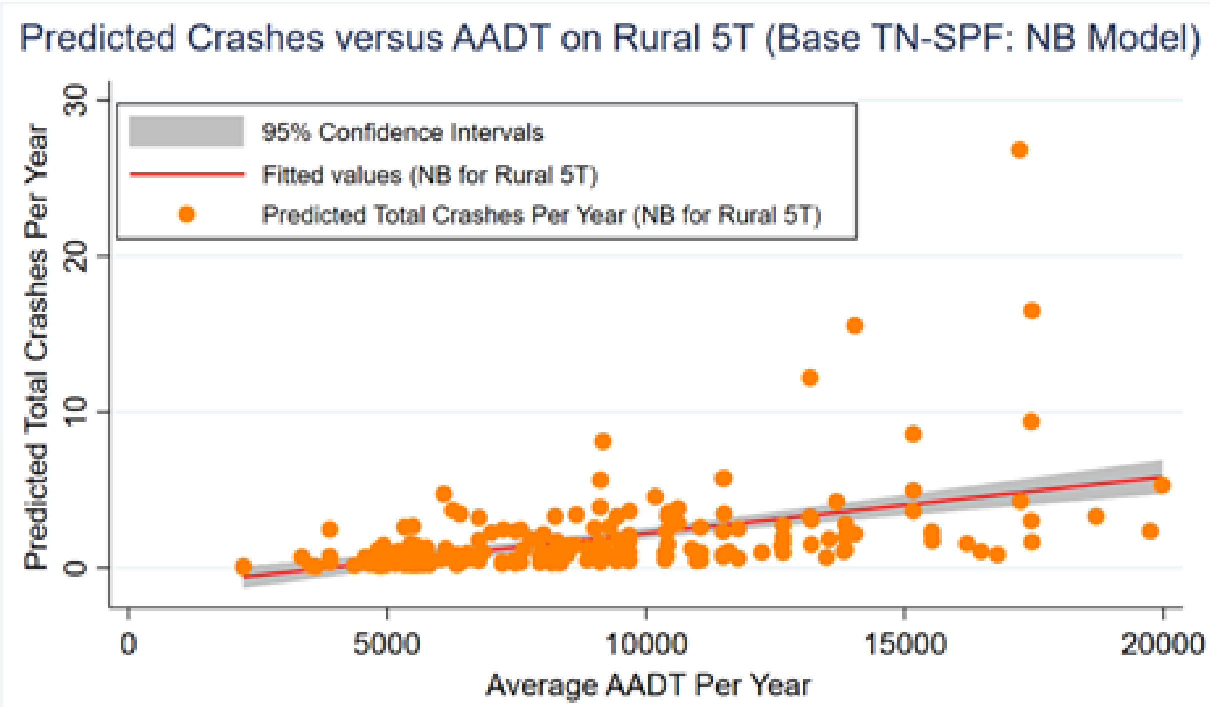


Figure B.1. Predicted Crashes versus AADT on Rural 5T: Predictions and Confidence Intervals (TN-SPF)

Note: The minimum and maximum AADT/year on rural 5T are 2219 and 19,978 respectively.

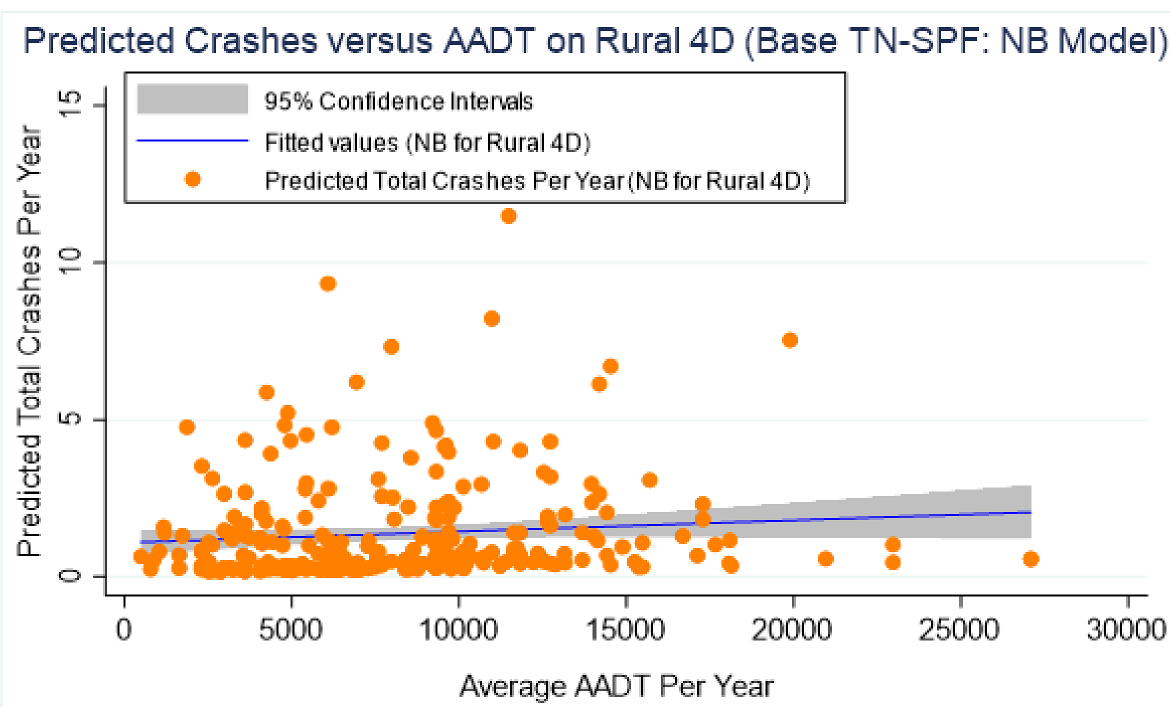


Figure B.2. Predicted Crashes versus AADT on Rural 4D: Predictions and Confidence Intervals (TN-SPF)

Note: The minimum and maximum AADT/year on rural 4D are 490 and 27,085 respectively.

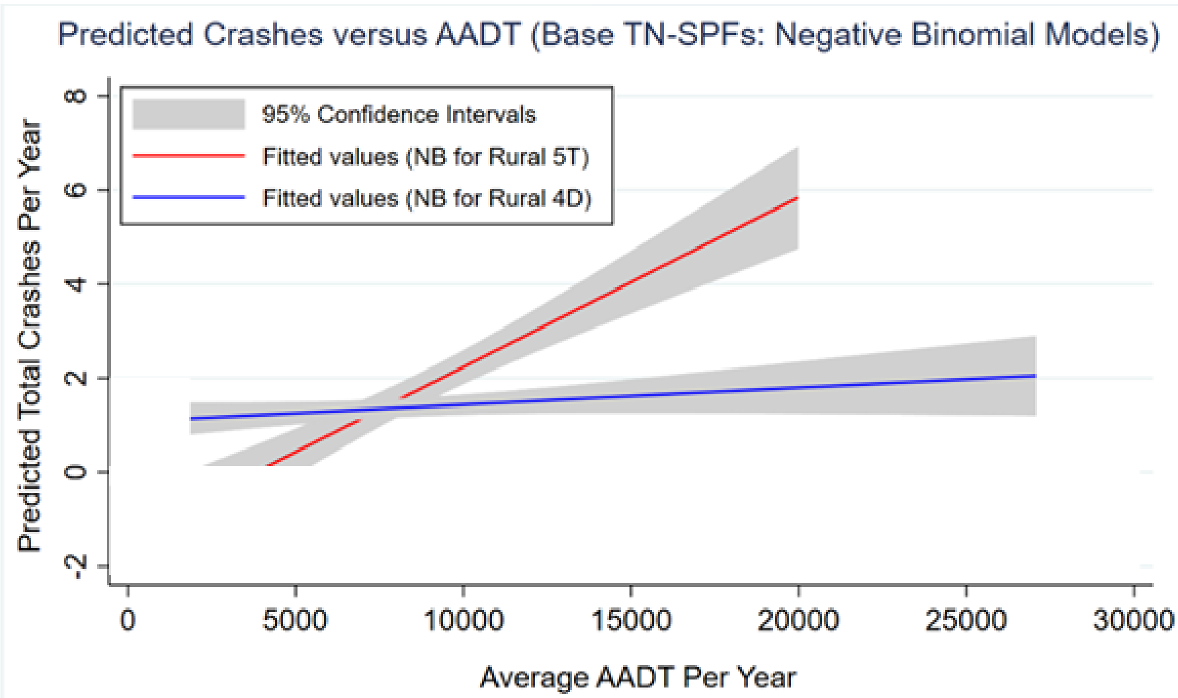


Figure B.3. Predicted Crashes versus AADT (TN-SPFs): Check for Confidence Intervals Overlap

TABLE B. 1 NB TN-SPF FOR TOTAL CRASHES ON RURAL 5T: ACCOUNTING FOR NON-LINEAR RELATIONSHIPS

Explanatory Variables	***NB (Rural 5T)***	
	Coef.	***t-stat***
Average 5-years AADT	0.0003	3.41
*Average 5-years AADT*Average 5-years AADT*	-8.71E-09	-2.01
Segment length (mile): Exposure	1	---
Constant	-1.1364	-2.21
Overdispersion parameter (alpha)	0.2332	3.29
Test for alpha significantly different than 0		
Chi-square test statistics	---	
Prob. (Chi-square)	---	
Summary Statistics		
Log-likelihood (Convergence)	-296.0749	
AIC	600.1498	
Degrees of freedom	4	
Sample Size (N)	205	

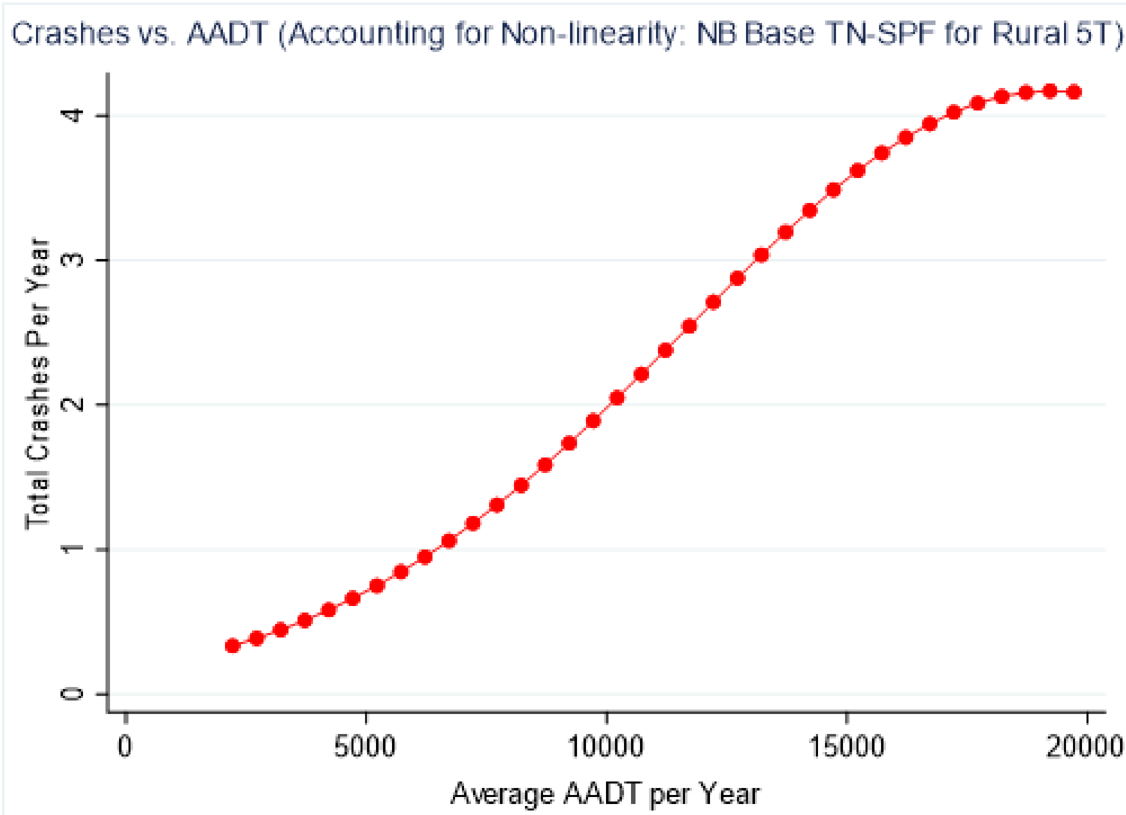


Figure B.4. Predicted Crashes vs. AADT on Rural 5T (Non-HSM Functional form: NB with AADT & AADT²)

Note: The minimum and maximum AADT/year on rural 5T are 2219 and 19978 respectively.