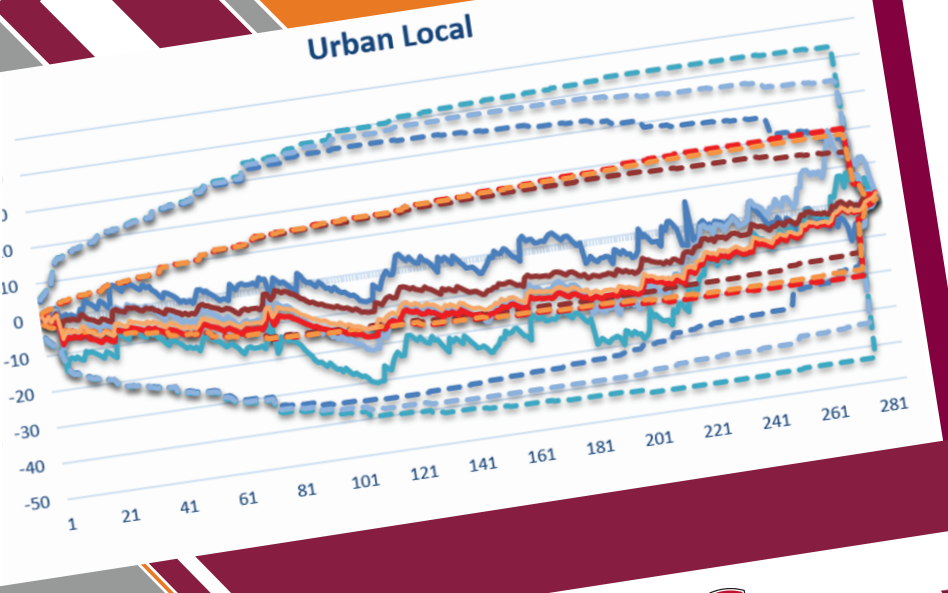
Use of Disruptive Technologies to Support Safety Analysis and Meet New Federal Requirements

March 2021

Final Report



SAFE-D
SAFETY THROUGH DISRUPTION

Texas A&M Transportation Institute

San Diego State University
Leadership Starts Here

VT
VIRGINIA TECH TRANSPORTATION INSTITUTE
VIRGINIA TECH

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report is funded, partially or entirely, by a grant from the U.S. Department of Transportation's University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. 04-113	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Use of Disruptive Technologies to Support Safety Analysis and Meet New Federal Requirements		5. Report Date March 2020 6. Performing Organization Code:
7. Author(s) Ioannis Tsapakis Subasish Das Ali Khodadadi Dominique Lord Jessica Morris Eric Li		8. Performing Organization Report No. Safe-D Project 04-113
9. Performing Organization Name and Address: Safe-D National UTC Texas A&M Transportation Institute The Texas A&M University System College Station, Texas 77843-3135 Virginia Polytechnic Institute and State University Virginia Tech Transportation Institute 3500 Transportation Research Plaza Blacksburg, Virginia 24061 USA		10. Work Unit No. 11. Contract or Grant No. 69A3551747115/04-113
12. Sponsoring Agency Name and Address Office of the Secretary of Transportation (OST) U.S. Department of Transportation (US DOT)		13. Type of Report and Period Final Research Report 14. Sponsoring Agency Code
15. Supplementary Notes This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program, and, in part, with general revenue funds from the State of Texas.		
16. Abstract States are required to have access to annual average daily traffic (AADT) for all public paved roads, including non-federal aid system (NFAS) roadways. The expectation is to use AADT estimates in data-driven safety analysis. Because collecting data on NFAS roads is financially difficult, agencies are interested in exploring affordable ways to estimate AADT. The goal of this project was to determine the accuracy of AADT estimates developed from alternative data sources and quantify the impact of AADT on safety analysis. The researchers compared 2017 AADT data provided by the Texas and Virginia Departments of Transportation against probe-based AADT estimates supplied by StreetLight Data Inc. Further, the research team developed safety performance functions (SPFs) for Texas and Virginia and performed a sensitivity analysis to determine the effects of AADT on the results obtained from the empirical Bayes method that uses SPFs. The results showed that the errors stemming from the probe AADT estimates were lower than those reported in a similar study that used 2015 AADT estimates. The sensitivity analysis revealed that the impact of AADT on safety analysis mainly depends on the size of the network, the AADT coefficients, and the overdispersion parameter of the SPFs.		
17. Key Words Annual average daily traffic, safety impact, safety analysis, low-volume roads, non-federal aid-system roads, safety performance functions	18. Distribution Statement No restrictions. This document is available to the public through the Safe-D National UTC website, as well as the following repositories: VTechWorks, The National Transportation Library, The Transportation Library, Volpe National Transportation Systems Center, Federal Highway Administration Research	

	Library, and the National Technical Reports Library.		
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages [insert # pages]	22. Price $0

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

Abstract

States are required to have access to annual average daily traffic (AADT) for all public paved roads, including non-federal aid system (NFAS) roadways. The expectation is to use AADT estimates in data-driven safety analysis. Because collecting data on NFAS roads is financially difficult, agencies are interested in exploring affordable ways to estimate AADT. The goal of this project was to determine the accuracy of AADT estimates developed from alternative data sources and quantify the impact of AADT on safety analysis. The researchers compared 2017 AADT data provided by the Texas and Virginia Departments of Transportation against probe-based AADT estimates supplied by StreetLight Data Inc. Further, the research team developed safety performance functions (SPFs) for Texas and Virginia and performed a sensitivity analysis to determine the effects of AADT on the results obtained from the empirical Bayes method that uses SPFs. The results showed that the errors stemming from the probe AADT estimates were lower than those reported in a similar study that used 2015 AADT estimates. The sensitivity analysis revealed that the impact of AADT on safety analysis mainly depends on the size of the network, the AADT coefficients, and the overdispersion parameter of the SPFs.

Acknowledgements

This project was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation – Office of the Assistant Secretary for Research and Technology, University Transportation Centers Program. The authors would like to thank StreetLight Data Inc. for providing probe data that were used in this study. We would also like to thank Dr. In-Kyu Lim and Wenling Chen from the Virginia Department of Transportation for their advice, comments, and help with this project.

Table of Contents

List of Figures

List of Tables

Introduction

The Federal Highway Administration (FHWA) requires states to report annual average daily traffic (AADT) through the Highway Performance Monitoring System for all federal-aid roads [1]. In March 2016, the United States Department of Transportation published the Highway Safety Improvement Program Final Rule [2]. According to the new Final Rule, states are required to have access to AADT along with other data elements for all public paved roads, including non-federal aid system (NFAS) roadways that include three roadway functional classes: rural minor collectors (6R), urban local roads (7U), and rural local roads (7R). States must have access to AADT data by September 30, 2026. The general expectation is to use AADT estimates in data-driven safety analysis and adopt advanced safety performance measures.

Most Departments of Transportation (DOTs) tend to focus their traffic data collection efforts on high-volume roads that typically pose significant safety challenges compared to NFAS roads. The latter account for 75% of the total roadway mileage in the US [3] and therefore, conducting an extensive number of short-term counts (STCs) on NFAS roads is financially difficult. Many agencies have raised concerns regarding the use of their limited budgets for data collection purposes on NFAS roads.

Because of these challenges, states are interested in exploring affordable ways to collect data and develop AADT estimates that are appropriate for use in safety analysis. Over the last few years, there has been an increasing interest in exploring whether passively collected data from mobile devices (e.g., smartphones, personal and commercial navigation devices, and fleet monitoring systems) that are already in the traffic stream can be used along with other types of data (e.g., census data) to estimate accurate traffic volumes. To address this need, the authors pursued two research objectives:

- Determine the accuracy of AADT estimates developed for NFAS roads from alternative data sources such as probe and census data. StreetLight Data Inc. (SLD), a third-party data vendor, provided the probe-based AADT estimates that were used in this project.
- Quantify the impact of AADT estimation errors on data-driven safety analysis, such as network screening, which is described in the Highway Safety Manual [4]. The purpose of network screening is to scan the transportation network and rank sites from most to least likely to realize a reduction in crash frequency by implementing one or more safety treatments.

To address these objectives, the researchers performed the following activities:

- Gathered crash, traffic, and roadway data for NFAS roads in Texas and Virginia, and integrated them with probe-based AADT estimates provided by SLD.
- Compared AADT values derived from permanent traffic stations and STCs against SLD AADT estimates.

- Developed safety performance functions (SPFs) for NFAS roads in Texas and Virginia.
- Conducted a sensitivity analysis to determine the impact of AADT estimation errors on safety analysis that involved applying the Empirical Bayes (EB) method that uses SPFs.

Methodology

This section describes the methodology that the researchers followed to determine the accuracy of SLD AADT estimates and quantify the impact of AADT estimation errors on safety analysis.

AADT Accuracy Measures

To quantify the accuracy of SLD AADT estimates, the authors calculated the following metrics:

$$MSD\ (vehicles) = \frac{1}{n}\sum_{i=1}^{n}\left(AADT_{Estimated,i} - AADT_{Observed,i}\right) \quad (1)$$

$$MAD\ (vehicles) = \frac{1}{n}\sum_{i=1}^{n}\left(\left|AADT_{Estimated,i} - AADT_{Observed,i}\right|\right) \quad (2)$$

$$MAPE\ (\%) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{\left|AADT_{Estimated,i} - AADT_{Observed,i}\right|}{AADT_{Observed,i}}\right) \times 100 \quad (3)$$

$$ACV\ (\%) = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Standard\ Deviation(AADT_{Estimated,i}, AADT_{Observed,i})}{(AADT_{Estimated,i} + AADT_{Observed,i})/2}\right) \times 100 \quad (4)$$

Where:

MSD = mean signed difference.

MAD = mean absolute difference.

MAPE = mean absolute percent error.

ACV = average coefficient of variation.

$AADT_{Estimated,\ i}$ = SLD AADT estimate for the i^{th} site.

$AADT_{Observed,\ i}$ = observed AADT or the i^{th} site. These AADT values were provided by the Texas DOT (TxDOT) and the Virginia DOT (VDOT).

n = total number of sites included in the evaluation.

In addition to these measures, the research team calculated the median absolute percent error (APE), as it is generally considered more appropriate than the MAPE in situations where outliers exist in the data or the data are not normally distributed. The accuracy measures were calculated at different levels of aggregation—such as by state, functional class, rural/urban code, AADT range—as well as a combination of these variables (e.g., functional class combined with rural/urban code).

Safety Impact Analysis

The safety impact analysis involved two steps. In the first step, the research team developed SPFs for NFAS roads in Texas and Virginia. In the second step, the authors conducted a sensitivity analysis by repeatedly applying the EB and the full Bayesian (FB) method using the SPFs developed in Step 1.

Step 1 – SPF Development

SPFs are typically negative binomial (NB) models (i.e., equations) that predict the mean crash frequency at a given facility as a function of AADT and roadway characteristics such as segment length, shoulder width, etc. A baseline SPF is often developed using AADT and segment length:

$$C_{Predicted} = \exp[\beta_0 + \beta_1 \times \ln(L) + \beta_2 \times \ln(AADT)] \quad (5)$$

Crashes (N) can be predicted by multiplying three components: predicted crash frequency ($C_{Predicted}$) from baseline SPF, a series of crash modification factors (CMFs), and a calibration factor, C:

$$N = C_{Predicted} \times \prod CMF \times C \quad (6)$$

The authors developed separate SPFs for Texas and Virginia. State- and local-specific SPFs are generally preferred [4] over "global" SPFs provided in the Highway Safety Manual or other sources. The two state datasets that the researchers compiled had significant differences in terms of size (i.e., number of roadway segments) and number of independent variables that could be included in the SPFs. NFAS roads usually have unique features such as limited mobility, shorter segments, and fewer crashes (compared to higher functional class roads), which make it challenging to accurately quantify their safety performance by applying "global" SPFs that may have been developed for different state transportation networks.

SPFs for Texas – Traditional NB Models With and Without Decision Trees

The authors used the Texas dataset to investigate whether decision trees can improve the SPF prediction accuracy. Conventional SPFs generally examine the mean effects of key contributing factors and ignore subgroups that may have different characteristics. Due to this generalized approach, they fail to capture specific subgroup effects and influential factors within a subset of roadway segments or intersections. New modeling approaches are needed to tackle the complexities of crash data and improve the accuracy of the predictions. Decision tree rule-based modeling is one of several emerging approaches that can address these limitations. These methods can identify subgroup effects without imposing any prior assumption or group of assumptions [5]. The rules provide a subset of SPFs that represent subsets of roadway segments or intersections by not only considering interactions between the contributing factors but also their ranges. Recursive partitioning is one of the simplest rules-based modeling techniques.

In this study, the authors used two open-source R (rpart, rattle) packages [6, 7] to develop decision trees using data from Texas. Then, traditional SPFs (NB models) were developed for each roadway

group (or cluster) that was produced from the decision trees. The decision tree-based SPFs were compared against functional class-specific SPFs that were developed without dividing the population into clusters (i.e., without using decision trees).

SPFs for Virginia – Traditional and Non-Traditional NB Models Using FB Method

Using data from Virginia and by applying the FB method, the authors developed traditional and non-traditional SPFs. The latter were developed to examine how different functional forms and dispersion structures can improve the performance of the traditional SPFs.

Functional Form of Traditional NB Models Using FB Method

In traditional NB models, there is a quadric association between the mean function and the variance through the over-dispersion parameter. This relationship is a natural result of the traditional formulation of the NB model. Therefore, modifying this relationship would lead to the transformations of the NB formulation. Three different NB parametrization of NB model, NB1, NB2, and NBP, corresponding to three different variance structures, have been proposed and examined in the literature [8]. These structures can be written as follows:

$$NB1: \quad var(y_i) = \mu_i + \frac{\mu_i}{\phi} \tag{7}$$

$$NB2: \quad var(y_i) = \mu_i + \frac{\mu_i^2}{\phi} \tag{8}$$

$$NBP: \quad var(y_i) = \mu_i + \frac{\mu_i^P}{\phi} \tag{9}$$

Where, ϕ is the inverse dispersion parameter, and μ_i is the mean crash frequency. Among these structures, NBP offers the most flexible association between the mean and the variance by including a learnable parameter (P) to the model. Using data from Virginia, the authors developed three traditional NB models (NB1, NB2, NBP) for each functional classification (6R, 7R, and 7U), as well as for all NFAS roads (as one group). These models were compared to non-traditional NB Lindley models, as described below.

Negative Binomial Lindley Models Using FB Method

Traditional NB models can account for over-dispersion effects, but in the case of crash distributions with long tails and a large number of zeros, NB is not flexible enough to capture all the variability in the data. NFAS roads typically have lower crash rates (compared to higher-volume roads), leading to a large number of zero responses and long tails in their crash distribution. The Negative Binomial Lindley (NB-L) model is a mixture of the NB and Lindley distribution, which offers a more flexible structure to the traditional NB model through the Lindley parameter. NB-L was proposed by [9] and then generalized and examined by [10, 11] in the field of crash analysis. The hierarchical representation of the NB-L model can be written as follows:

$$P(Y = y, \mu_i, \phi|\varepsilon) = NB(y; \phi, \varepsilon\mu_i) \tag{9}$$

$$\varepsilon \sim Lindley(\theta)$$

Where θ is the Lindley parameter. In addition to the three traditional NB models (NB1, NB2, NBP) described previously, the authors used data from Virginia to develop three NB-L models (NB1-L, NB2-L, and NBP-L) for each functional classification (6R, 7R, and 7U), as well as for all NFAS roads (as one group).

Dispersion Structure

Besides the variance structure, the over-dispersion parameter itself can affect the flexibility of the NB model. The over-dispersion parameter is the linkage between the mean and the variance of the NB model. Even though past studies assumed that the over-dispersion parameter is fixed and invariant of site characteristics, some research studies [12, 13] showed that the varying dispersion structure, as a function of factors such as AADT or segment length, provides a superior fit compared to the models with a fixed dispersion structure. In this regard, Geedipally and Lord analyzed 10 different functional forms of dispersion structure dependent upon segment length and AADT [14]. All of the aforementioned NB formulations were also developed and examined with different dispersion structures in line with the best structures found by [14].

FB Method

All of the Virginia SPFs were developed in an FB framework. We set a non-informative normal prior for the coefficients of both mean function and dispersion structure. Also, as recommended by [11, 15], a Beta (N/3, N/2) distribution was chosen for a function of the θ parameter. The authors performed a Markov chain Monte Carlo analysis using three different chains, each containing 50,000 draws from the joint posterior distribution. The first 10,000 draws, and two draws out of each three draws were ignored to ensure the convergence and independence of the samples.

Step 2 – Sensitivity Analysis

The authors performed a sensitivity analysis to examine the impact of AADT estimation errors on the expected crash frequencies estimated using the EB method and in some cases the FB method. The results from both methods were used to rank the sites. The sensitivity analysis included the following steps:

- Step 1: Apply the EB method (for all Texas SPFs and for the traditional NB models for Virginia) to calculate the expected number crashes for each site. The EB method considers both the number of crashes predicted using SPFs, based on the average conditions of the group, along with the observed number of crashes at a given facility. The EB approach is based on a weighted average concept. Many studies use this method to develop localized SPFs for different facility and crash types [16-25]. The EB method improves the estimation precision by using a weight factor, w, to combine observed ($C_{Observed}$) and predicted crash frequencies ($C_{Predicted}$):

$$C_{Expected} = w \times C_{Predicted} + (1 - w) \times C_{Observed} \quad (10)$$

$$w = \frac{1}{1 + C_{Predicted} \times OP} \quad (11)$$

Where:

w= a weight factor that depends on the over-dispersion parameter (OP) of the SPF

OP = over-dispersion parameter (OP)

$C_{Expected}$ = expected crash frequency

$C_{Observed}$ = observed crash frequency

Note that the EB method is an approximation to a more general framework, the FB method, which was used in the case of the NB-L models developed for Virginia. The FB method uses a posterior predictive distribution to calculate expected values in Bayesian inferences. The posterior predictive distribution is equivalent to the distribution of the future data given the existing data that have been used to develop a model. Expected crash frequencies from both the EB method (as described above) and the posterior predictive distribution of the NB-L distribution were calculated in Step 1.

- Step 2: Rank the segments based on expected number of crashes.
- Step 3: Determine rank percentile for each site.
- Step 4: Increase the AADT of each segment by 10, 50, 100, 250 and 500 percent by keeping the rest of the variables and segments fixed.
- Step 5: Repeat Steps 1–4 separately for each segment.
- Step 6: Calculate the percentile rank change of each site by differentiating the original ranking (no AADT change) against the rank obtained when AADT was increased by a certain percent.

The study data and the results of the analyses are described in the next section.

Results

Study Data

DOT Data

The authors assembled a comprehensive database of roadway, traffic volume, and crash data for NFAS roads in Texas and Virginia. Data were gathered for the five-year period of 2014–2018. Most of the datasets were downloaded online from public websites maintained by TxDOT and VDOT. Traffic volume data for Virginia were provided by VDOT staff. The traffic volume data were obtained from both permanent stations and STCs. To develop the database, the authors took the following steps.

- Removed intersection and intersection-related crashes.

- Excluded missing records and outliers. For example, the researchers filtered out counts that were missing at least one of the following attributes: station ID, latitude, longitude, rural/urban designation, roadway functional class, and count type (i.e., permanent or short term).
- Identified and excluded low-quality count data in the case of Virginia.
- Excluded very short segments (segments length > 0.099 mi.).
- Geolocated crashes on the state transportation networks and created a geodatabase in ArcGIS.
- Determined the number of crashes on each segment by injury type and year and developed the final dataset. Note that the development of SPFs requires a comprehensive crash database with geocode crash location information or relevant route information, roadway type, injury type, collision type, and other relevant information. The injury classification system, KABCO, used in this study divides crash severity into five major groups: 1) fatal injury (K), 2) incapacitating suspected serious injury (A), 3) non-incapacitating injury (B), 4) possible injury (C), and 5) no injury or property damage only (O).

It is worth stating that the dataset compiled for Virginia had significantly fewer attributes that could be used as independent variables to develop SPFs than the Texas dataset. Further, the sample size (i.e., number of segments) was much smaller, mainly due to the smaller size of Virginia's transportation network.

SLD Data

The Texas Transportation Institute (TTI) downloaded 2017 AADT estimates for 10,000 roadway locations from SLD's web-platform, Insight. SLD developed nearly all the analytics for estimating 2017 AADT values at no cost to this research project. SLD follows three main steps to develop AADT estimates [26]:

- Step 1: Process and combine GPS and location based services (LBS) data that SLD obtains from various data providers.
- Step 2: Normalize GPS and LBS trip counts (derived from Step 1) using non-traffic data, such as U.S. census socioeconomic and demographic data.
- Step 3: Calibrate the estimates developed in Step 2 using machine learning algorithms. SLD uses actual traffic volume data that public agencies collect primarily from continuous count stations that are permanently installed at select locations of the network.

All passively collected data and the details of the traffic volume estimation models are the intellectual property of SLD and are considered confidential.

AADT Accuracy

TTI compared SLD AADT estimates against permanent and short-term count data provided by the two DOTs. The AADT values extracted from permanent stations are typically considered to be representative of the actual traffic volumes, assuming the dataset is complete and free of erroneous values. Therefore, under certain circumstances, the AADT derived from permanent sites can be used for validation purposes. However, the AADTs derived from STCs have been estimated (not calculated) by applying one or more seasonal adjustment factors to the average daily traffic (ADT) of the counts. As a result, the short-duration AADT values have an inherent estimation error, which does not make them appropriate for validation purposes. In this study, they were merely used as a comparison device.

The research team calculated the metrics presented in the Methodology section to compare DOT-supplied AADT values against SLD AADT estimates. Table 1 shows these metrics aggregated by five traffic volume ranges that the authors developed based on DOT data. Table 5 of the appendix presents the results disaggregated by state and AADT range. Note that SLD did not produce AADT estimates that are less than 400 vehicles per day (vpd); therefore, the errors are high within the first volume range. As of the publication year of this report, SLD has made significant methodological improvements and is currently producing AADT estimates for all volume ranges, including low-volume roads (0–400 vpd).

Table 1. Accuracy of SLD AADT Estimates by AADT Range

AADT Range (vehicles/day)	Number of Records	MSD	MAD	MAPE	Median APE	ACV
0–399	5,545	NA	NA	NA	NA	NA
400–1,999	3,484	757	768	118%	84%	40%
2,000–4,999	365	498	888	32%	25%	19%
5,000-9,999	59	211	1,989	32%	21%	19%
≥10,000	32	(3,324)	4,934	30%	33%	26%
Grand Total	**9,485**	**691**	**831**	**108%**	**77%**	**38%**

NA = Not applicable

The main findings related to the accuracy of AADT estimates are summarized below:

- In general, the AADT accuracy gradually improves from lower to higher traffic volume roads.
- The grand average Median APE is 77%; which nonetheless decreases to 25% when the first two volume groups are excluded from the analysis (i.e., AADT>2000 vpd).
- SLD AADT estimates tend to be higher than DOT AADT values (i.e., positive mean signed difference) for the first four AADT ranges, but this trend is reversed for the last range (>10,000 vpd).

- The MAD gradually increases from low-volume roads to higher AADT roads. Not surprisingly, this finding is also observed when the MAD is aggregated by roadway functional class.

The wide range of AADT estimation errors was used as an input to construct the sensitivity analysis described in the Methodology section.

SPFs

Texas SPFs – Decision Trees

After performing a correlation analysis and determining variable importance measures, the segment length and AADT were found to be the best explanatory variables [27]. The decision trees developed in this study confirmed these findings. Figure 1 shows three decision trees developed for KABCO, KABC, and KAB crashes for rural minor collectors (6R) in Texas. In Figure 1a, the total number of KABCO crashes was used as the dependent variable of the decision tree. The independent variables included segment length (LEN_SEC), average daily traffic (ADT_CUR), shoulder width, and others. Figure 1(a) provides annotation of various statistics calculated during the development of the decision trees. A classification and regression tree (CRT) algorithm was applied to the dataset to determine the appropriate number of clusters. A maximum of three levels was used to limit the number of the final clusters. All the decision trees used in this study were validated by splitting the dataset into a training and a test dataset. This report only presents the results obtained for 6R roadways. For example, the decision rules generated for KABCO crashes on 6R roadways are:

- Class 1 rule: LEN_SEC (Segment Length) < 1.3 and ADT_CUR (AADT in the current year) < 612 (mean crash frequency = 0.28 crashes/year per segment).
- Class 2 rule: LEN_SEC < 1.3 and ADT_CUR ≥ 612 (mean crash frequency = 1 crash /year per segment)
- Class 3 rule: LEN_SEC ≥ 1.3 and ADT_CUR < 331 (mean crash frequency = 0.8 crashes/year per segment).
- Class 4 rule: LEN_SEC ≥ 1.3 and ADT_CUR ≥ 331 (mean crash frequency = 3.3 crashes/year per segment).

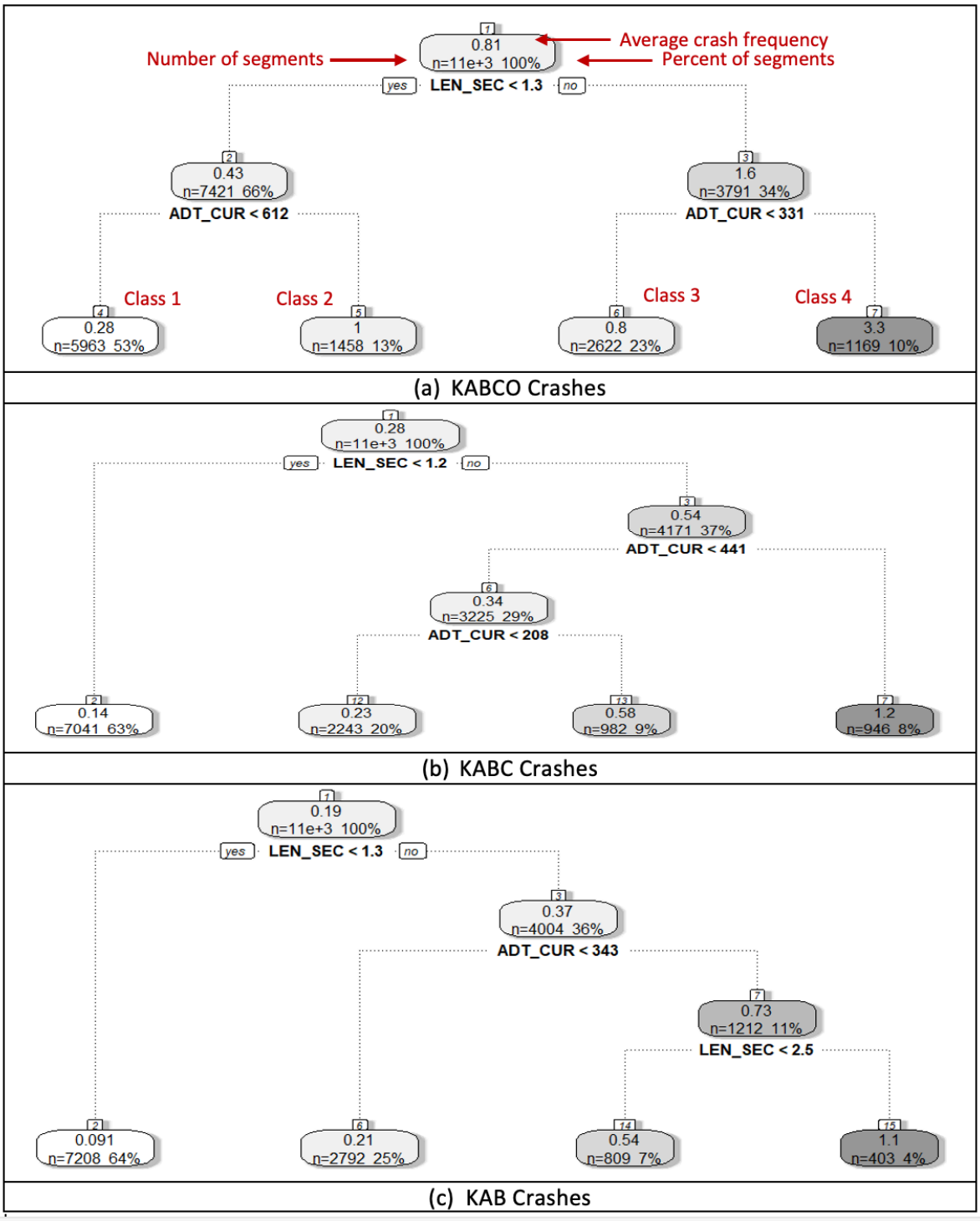


Figure 1. Decision trees for a) KABCO, b) KABC, and C) KAB crashes.

For rural minor collectors in Texas, there were 11,212 segments with available AADT data. lists the SPFs developed for the three crash severity groups (KABCO, KABC, and KAB) and the clusters created from the decision trees. The table shows the model equation along with the overdispersion parameter (b), and the loglikelihood of each model.

Table 2. SPFs (NB Models) for 6R in Texas

Crash Severity Group	Class	Safety Performance Functions	Over-Dispersion Parameter	Log-likelihood
KABCO	Class 1	**Rule: All Data** $N_{6R,tot,all} = exp(-4.759) \times Length^{0.900} \times AADT^{0.766}$	1.3075	-22215.077
KABCO	Class 2	**Rule: LEN_SEC < 1.3 & ADT_CUR < 612** $N_{6R,tot,class1} = exp(-4.170) \times Length^{0.898} \times AADT^{0.658}$	0.8020	-7317.370
KABCO	Class 3	**Rule: LEN_SEC < 1.3 & ADT_CUR > 611** $N_{6R,tot,class2} = exp(-4.298) \times Length^{0.958} \times AADT^{0.699}$	1.3710	-3676.017
KABCO	Class 4	**Rule: LEN_SEC > 1.2 & ADT_CUR < 331** $N_{6R,tot,class3} = exp(-4.627) \times Length^{0.764} \times AADT^{0.756}$	0.9225	-6036.538
KABCO	All	**Rule: LEN_SEC > 1.2 & ADT_CUR > 331** $N_{6R,tot,class4} = exp(-4.606) \times Length^{0.832} \times AADT^{0.763}$	2.129	-5071.186
KABC	Class 1	**Rule: All Data** $N_{6R,kabc,all} = exp(-5.636) \times Length^{0.940} \times AADT^{0.736}$	1.2247	-12386.275
KABC	Class 2	**Rule: LEN_SEC < 1.2** $N_{6R,kabc,class1} = exp(-5.335) \times Length^{0.953} \times \text{AADT}^{0.686}$	1.035	-5141.116
KABC	Class 3	**Rule: LEN_SEC > 1.1 & ADT_CUR < 208** $\text{N}_{6R,kabc,class2} = \exp(-5.323) \times \text{Length}^{0.649937} \times \text{AADT}^{0.649}$	0.4449	-2459.781
KABC	Class 4	**Rule: LEN_SEC > 1.1 & 207 < ADT_CUR < 441** $\text{N}_{6R,kabc,class3} = \exp(-4.043) \times \text{Length}^{1.023} \times \text{AADT}^{0.446}$	1.191	-1954.343
KABC	All	**Rule: LEN_SEC > 1.1 & ADT_CUR > 440** $\text{N}_{6R,kabc,class4} = \exp(-4.707) \times \text{Length}^{0.688} \times \text{AADT}^{0.639}$	2.518	-2738.366
KAB	Class 1	**Rule: All Data** $\text{N}_{6R,kab,all} = \exp(-5.949) \times \text{Length}^{0.960} \times \text{AADT}^{0.719}$	1.259	-9630.336
KAB	Class 2	**Rule: LEN_SEC < 1.3** $\text{N}_{6R,kab,class1} = \exp(-5.846) \times \text{Length}^{0.945} \times \text{AADT}^{0.699}$	0.961	-4002.588
KAB	Class 3	**Rule: LEN_SEC > 1.2 & ADT_CUR < 343** $\text{N}_{6R,kab,class2} = \exp(-6.169) \times \text{Length}^{0.963} \times \text{AADT}^{0.756}$	0.654	-2893.847
KAB	Class 4	**Rule: 1.2 < LEN_SEC < 2.5 & ADT_CUR > 342** $\text{N}_{6R,kab,class3} = \exp(-5.401) \times \text{Length}^{0.858} \times \text{AADT}^{0.649}$	3.420	-1552.351
KAB	All	**Rule: LEN_SEC > 2.5 & ADT_CUR > 342** $\text{N}_{6R,kab,class4} = \exp(-3.999) \times \text{Length}^{0.619} \times \text{AADT}^{0.501}$	1.935	-1137.224

As mentioned earlier, regression models examine the mean effects of the explanatory variables and ignore subclass effects in the entire population of all segments. This study applied decision trees to determine the subclass effect in the dataset. As the current model is completely based on the rural minor collector roadways in Texas, transferability of these models to other states should be carefully considered. The R^2 values range from 0.18 to 0.22 for all data (without splitting) for different injury level models. The prediction accuracies are improved in the decision tree-based models. For different class-specific models (based on injury levels), the R^2 values range from 0.25

to 0.41. To understand the goodness-of-fit, another quick diagnostic is the development of Cumulative Residual (CURE) plots. Residuals indicate the disparities between historical crash frequencies and predicted crash frequencies. Model fitting can be performed by examining the residuals. If the surrounding residuals of a model are close to zero, the model can be considered as a good-fit model. The CURE plot is a good visualization tool to examine the SPF predictions based on the individual explanatory variables used in the model. A horizontal stretch of the CURE plot infers to a region of the variable where the estimates are unbiased [28]. On the contrary, in locations where the CURE plot drifts up or down significantly, the estimates are not considered to be unbiased. The CURE plot for an unbiased SPF must be within the boundaries of two standard deviations [28].

Figure 2 shows the CURE plots for the SPFs developed using KABCO crashes. There are five CURE plots on each side of the figure. The CURE plots on the left side show the segment length on the horizontal axis, whereas those on the right side show the AADT on the x-axis.

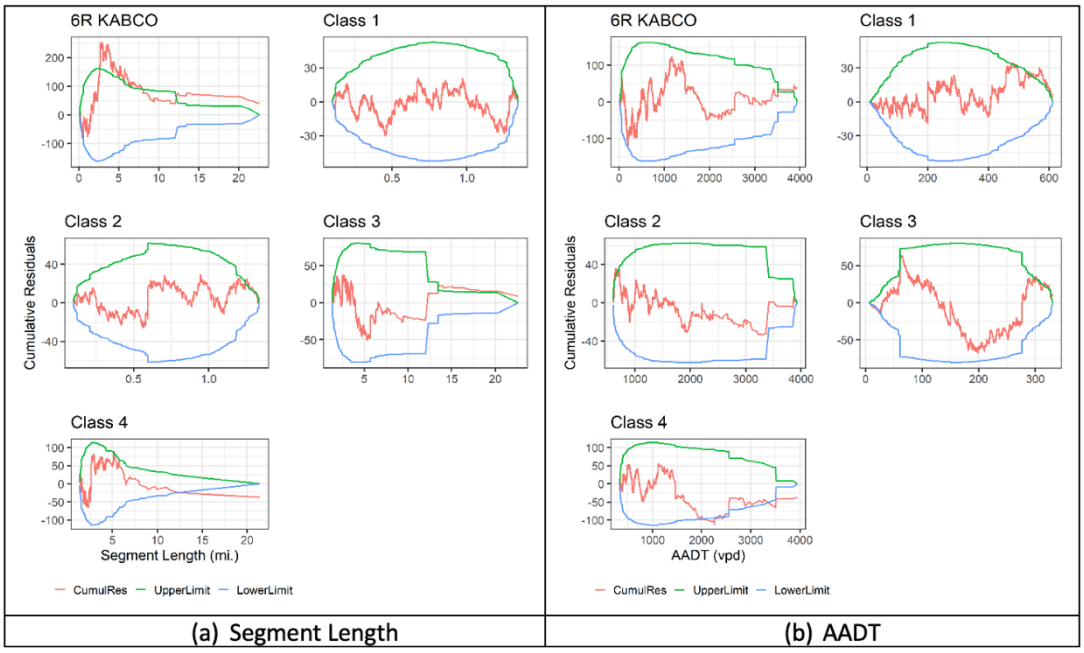


Figure 2. CURE plots for KABCO model.

Examining the projections of the residuals of each plot shows the improved performance measures of the class-based models—those developed based on the decision trees. A comparison between the CURE plots of the main models and those of the rules-based models indicates that the rules-based models are, for the most part, inside the confidence boundaries. For example, in Figure 2(a) the residual (red) line of the main KABCO model is outside of the confidence boundary in two zones, whereas the residual lines of the class-based models are, for the most part, within then confidence boundaries. This may also be due to the small sample size of the long segments in the database. The CURE plots for other models (not shown in this report) have similar trends. This example clearly shows the effectiveness of rules-based modeling in generating more robust SPFs.

Figure 3 shows the predicted KAB crash counts by the developed SPFs for different AADT values. It shows that predicted crashes in Class 4 differ from the rest. Similar plots for KABCO and KABC crashes are listed in the appendix (Figure 9 and Figure 10).

Figure 3. Predicted KAB crashes by SPFs.

Virginia SPFs – Different Functional Forms and Structures

As previously explained, in addition to decision trees, the authors explored how the functional form and (variance and dispersion) structures of SPFs can improve the latter's performance. Initially, a global SPF was developed for all three functional classes treated as one group. All six NB formulations were modeled with a fixed structure as well as four different dispersion structures to find the best model. The results of the generated SPFs that have a fixed dispersion parameter are provided in Table 3. The remaining results of the SPFs with non-fixed structures are provided in Table 6 and

Table 8 in the appendix.

Table 3. Model estimation results (fixed dispersion parameter) for all NFAS roads

Dispersion Structure	NB-1 – Fixed	NB-2 – Fixed	NB-P – Fixed	NB1-L – Fixed	NB2-L – Fixed	NBP-L – Fixed
Intercept ($\boldsymbol{\beta_0}$)	-5.08 (0.19)	-5.06 (0.20)	-4.99 (0.20)	-5.06 (0.24)	-5.11 (0.25)	-5.37 (0.27)
Ln(AADT) ($\boldsymbol{\beta_1}$)	0.73 (0.02)	0.73 (0.03)	0.73 (0.03)	0.72 (0.07)	0.73 (0.07)	0.76 (0.06)
Length ($\boldsymbol{\beta_2}$)	0.77 (0.02)	0.77 (0.02)	0.66 (0.02)	0.73 (0.02)	0.74 (0.02)	0.87 (0.03)
P	-	-	2.23 (0.11)	-	-	~~0.11 (0.10)~~
WAIC	8833	8725	8723	8315	8477	8471
LOO	8832	8725	8723	8678	8696	8559
MASE	0.67	0.62	0.61	0.22	0.23	0.33
MSPE	12.99	7.95	6.81	0.66	0.71	3.44
Log-Likelihood	-4413	-4360	-4361	-3655	-3711	-3679

The results showed that a) models with a varying dispersion structure outperformed models with a fixed dispersion parameter, and b) the appropriate dispersion structure depends on the NB formulation being used. The NB-L family models (NB1-L, NB2-L, and NBP-L) performed better with a length-only dependent dispersion structure. However, traditional NB models (NB1, NB2,

and NBP) showed a better performance when modeled with length and an AADT dependent dispersion structure. In the next step, six different models, each with the appropriate dispersion structure found in the previous step, were developed and examined for each classification, separately. The coefficient estimates and performance metric of the 6R model are provided in Table 4. The results of the 7R and 7U models are shown in Table 10 and Table 11 of the appendix.

Along with other commonly used metrics, two fully Bayesian metrics, widely applicable information criterion (WAIC) and leaving one out (LOO) cross-validation, were chosen for performance evaluation. WAIC was proven to perform better than DIC, especially in a hierarchical setting [29].

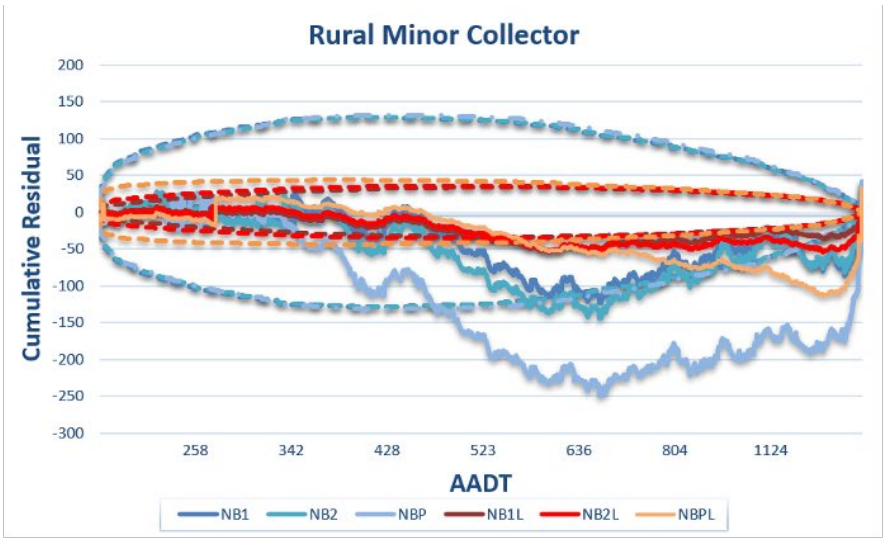
Table 4. Model Estimation Results for Rural Minor Collectors

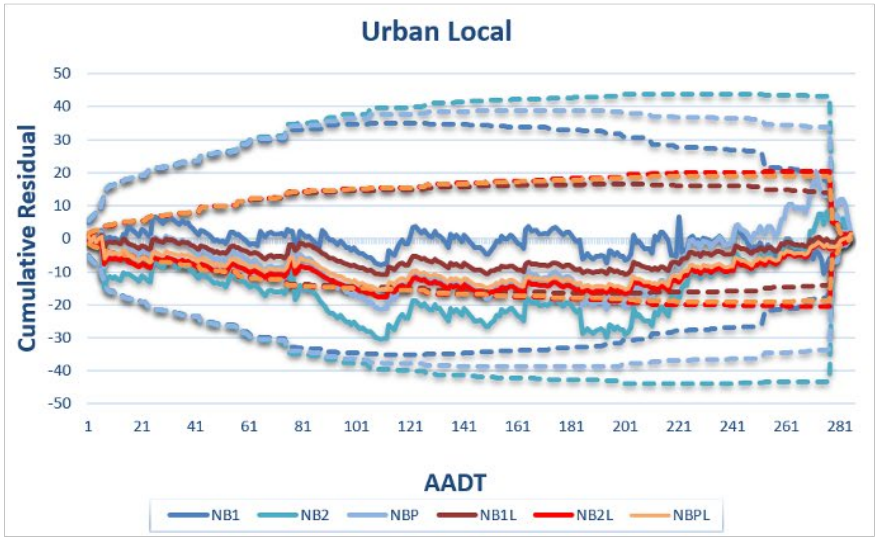
Dispersion Structure	$\emptyset_i = e^{\eta_0} AADT_i^{\eta_1} L_i^{\eta_2}$ **NB-1**	$\emptyset_i = e^{\eta_0} AADT_i^{\eta_1} L_i^{\eta_2}$ **NB-2**	$\emptyset_i = e^{\eta_0} AADT_i^{\eta_1} L_i^{\eta_2}$ **NB-P**	$\emptyset_i = e^{\eta_0} L_i^{\eta_2}$ **NB1-L**	$\emptyset_i = e^{\eta_0} L_i^{\eta_2}$ **NB2-L**	$\emptyset_i = e^{\eta_0} L_i^{\eta_2}$ **NBP-L**
Intercept ($\boldsymbol{\beta_0}$)	-4.88 (0.20)	-4.89 (0.21)	-4.88 (0.22)	-4.70 (0.31)	-4.81 (0.32)	-5.01 (0.35)
Ln(AADT) ($\boldsymbol{\beta_1}$)	0.81 (0.04)	0.79 (0.04)	0.78 (0.03)	0.71 (0.12)	0.73 (0.12)	0.78 (0.12)
Length ($\boldsymbol{\beta_2}$)	0.46 (0.01)	0.48 (0.01)	0.53 (0.02)	0.60 (0.02)	0.57 (0.02)	0.54 (0.03)
$\boldsymbol{\eta_0}$	4.22 (0.88)	~~-0.29 (0.83)~~	-7.41 (1.13)	5.91 (1.27)	4.71 (0.82)	5.81 (0.61)
$\boldsymbol{\eta_1}$	-0.69 (0.13)	~~0.15 (0.12)~~	1.53 (0.19)	-	-	-
$\boldsymbol{\eta_2}$	0.79 (0.14)	1.15 (0.11)	2.25 (0.15)	4.32 (0.84)	3.71 (0.53)	4.01 (0.49)
P	-	-	3.74 (0.18)	-	-	3.77 (0.2)
WAIC	6199	6161	6123	5753	5762	5784
LOO	6198	6161	6123	6123	6136	6124

According to the results, the NB-L family models provide a superior fit in comparison to the traditional NB models. In all functional classifications, almost all the performance evaluation metrics favored the NB-L models over the traditional NB models. The majority of roadways (56%) experienced less than two crashes during the five years of the analysis and the crash data are highly skewed (skewness > 2.8 for all classifications). These results are in line with previous research findings [11], according to which NB-L models provide a better fit in cases of excess zeros and long tails in crash distributions.

No considerable improvements were observed when using different variance structures. In rural minor collector SPF results, models with less flexible variance structures (i.e., NB1, and NB1-L) slightly outperformed their counterparts. In rural and urban local SPFs models, however, there were no significant differences between models with different variance structures. Adjusted cumulative residual plots are also provided to better represent the superiority of NB-L family models over traditional NB models. Figure 4 shows CURE plots for 6R, 7R, and 7U, respectively. Each figure includes six plots corresponding to the six SPF models. The confidence intervals are depicted by a dashed line of the same color as the corresponding residual line. The results show

that NB-L models have narrower confidence intervals and less periodicity compared to the traditional NB models.





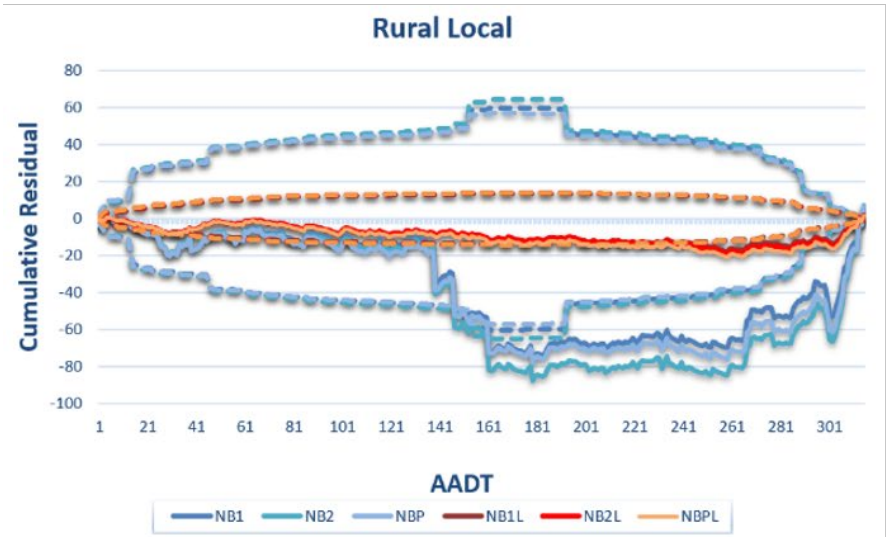


Figure 4. CURE plots for AADT.

Safety Impact Analysis

The purpose of this analysis was to determine whether and how AADT errors can affect the results of data-driven safety analysis, such as the final ranking of roadway segments sorted based on their safety risk. The sensitivity analysis was separately conducted for the decision tree-based SPFs developed for Texas, as well as for the NB-L-based SPFs developed for Virginia. Figure 5 shows the rank percentile changes in a box-violin format for five different AADT groups (illustrated in different colors), and five AADT percent increases starting with 10% (upper part) all the way to 500% percent (bottom part).

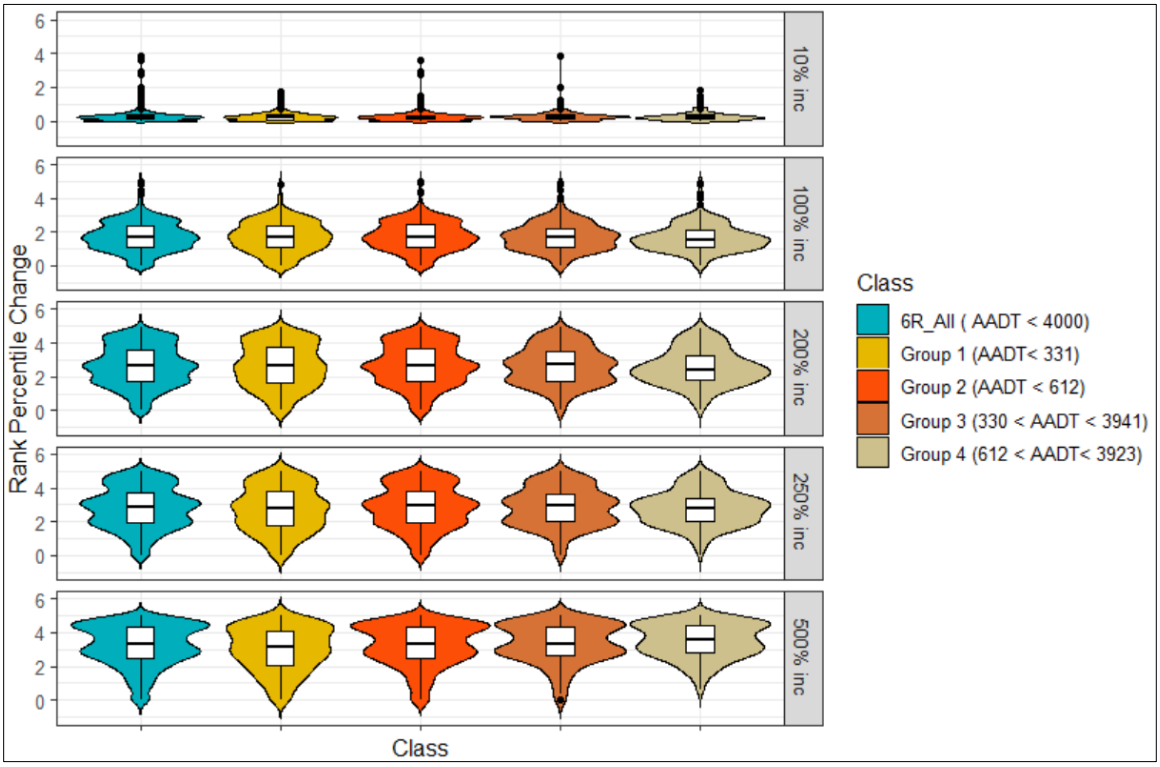


Figure 5. Sensitivity analysis results for 6R KABCO models.

These results were produced from the SPFs developed for KABCO crashes on 6R roadways in Texas. The figure shows that a 10% increase in AADT does not have a substantial effect on the expected crash frequencies and associated percentile rank changes. Higher percent increases in AADT result in slightly higher percentile rank changes; however, the latter are not proportional to the AADT percent increase. For example, the highest percentile rank changes were approximately 4% and were obtained when the AADT was increased by 500% (bottom part of Figure 5). The magnitude of this impact depends on several factors, such as the AADT coefficients of the SPFs (the smaller the coefficients, the smaller the impact), the sample size of the network (the bigger the network, the smaller the impact), and the overdispersion parameter (the higher the parameter, the smaller the impact), among others.

Figure 6 shows the rank percentile changes obtained from the EB method (three plots on the left) and the FB method (three plots on the right) that used the Virginia SPFs developed for 6R, 7R, and 7U.

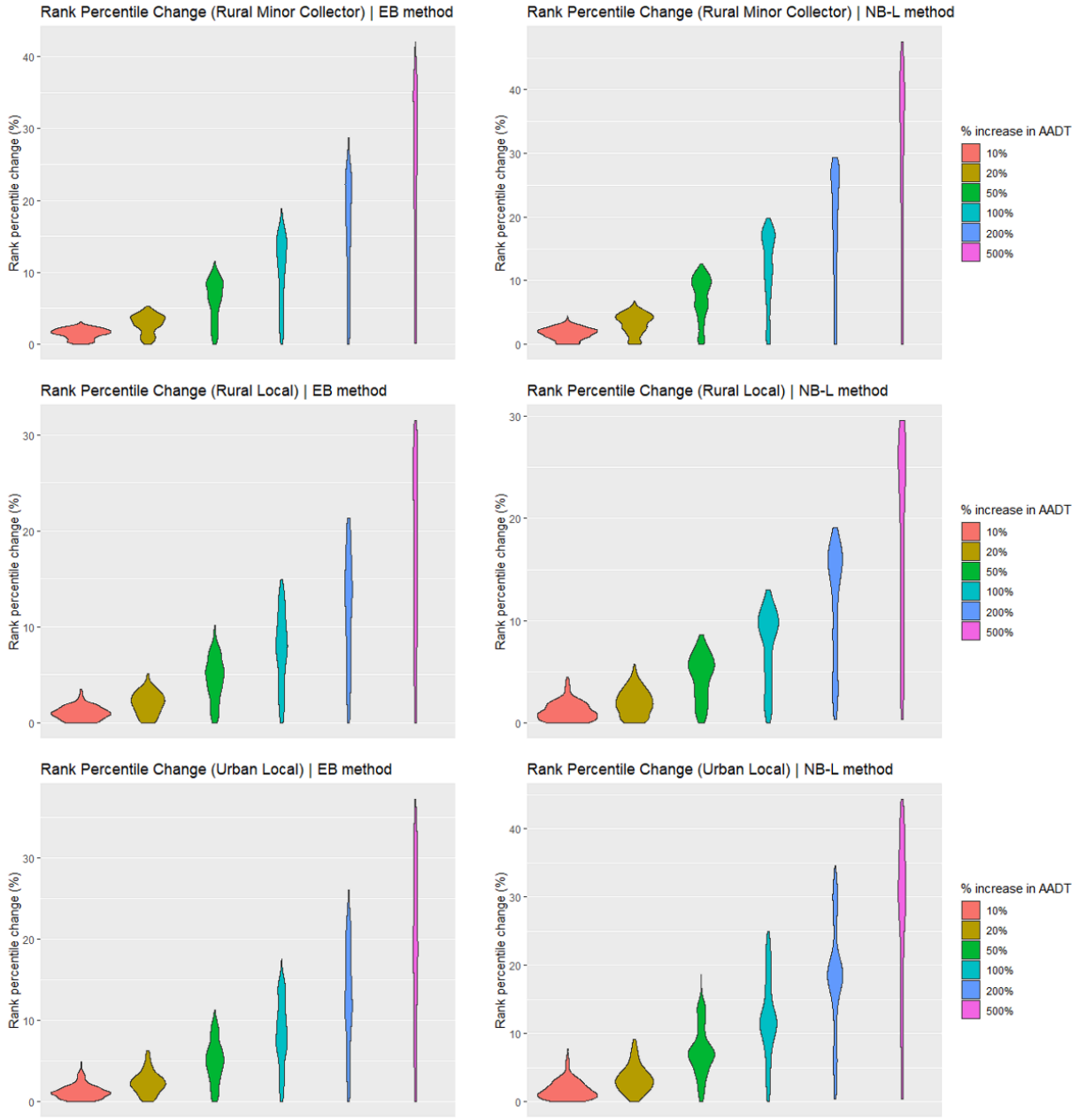


Figure 6. Rank percentile change using expected crash values from different methods (6R).

As illustrated in the legend of the figure, the five colored violin plots correspond to the five AADT percent increases: 10, 20, 50, 100, 200, and 500%. The results show that the NB-L method tends to be more sensitive (i.e., higher percentile rank changes) to AADT increases compared to the EB method. Note that the majority (96%) of the rank percentile changes that are greater than 20% are due to sites that had fewer than five crashes during the five-year study period. In general, the

percentile rank changes are significantly higher compared to those obtained for Texas (Figure 5). This can be partially attributed to the small size of the Virginia network and the NB-L models, which tend to be more sensitive to AADT changes than the traditional SPFs.

Conclusions and Recommendations

This study aimed to a) determine the accuracy of AADT estimates developed from alternative data sources, and b) quantify the impact of AADT estimation errors on data-driven safety analysis such as the EB method that uses SPFs. To address the first objective, the research team compared AADT values provided by TxDOT and VDOT against AADT estimates supplied by a third-party data vendor, StreetLight Data Inc., that develops various traffic products using a combination of traffic, probe, census, and other data. The comparison revealed that the median APE for roads with AADT greater than 2000 vpd is approximately 25%. The AADT accuracy gradually improves from lower to higher traffic volume roads. SLD tends to overestimate AADT within lower volume ranges (0-10,000 vpd) and underestimate it for roads that have an AADT higher than 10,000 vpd.

As the use of mobile devices continues to increase by the driving public and data providers continue to improve their analytical methods, the accuracy of AADT estimates is expected to increase. For example, the 2017 AADT estimates used in this project resulted in lower errors than those reported in a 2017 report that evaluated 2015 AADT estimates [26]. In 2020, MnDOT re-evaluated 2019 SLD estimates and found that the mean absolute error ranged from 8% to 10% for locations greater than 10,000 AADT and gradually increased to 42% for sites with less than 1,000 AADT [30]. Future evaluations of probe-based AADT estimates are needed using data from different states and regions that potentially have diverse traffic, geometric, demographic, socioeconomic, and weather characteristics. The ongoing FHWA pooled fund study "Independent Evaluation of Non-Traditional Methods to Obtain Annual Average Daily Traffic" is expected to shed light on this topic [31].

AADT estimates developed from alternative data sources can yield several benefits, such as time and cost savings by eliminating the need to conduct short-term counts and purchase and maintain expensive traffic equipment. These estimates could also reduce safety risks to employees and contractors who go out in the field to install sensor devices in and on roadways. Further, AADT estimates from alternative data sources can assist agencies in meeting new federal requirements mandating that states must have access to a series of data elements, including AADT, for all public paved roads by 2026.

To address the second study objective, the research team developed several SPFs for NFAS roads in Texas and Virginia, and then performed an extensive sensitivity analysis. A procedure was developed for using local roadway network data in estimating crash frequencies. The goodness-of-fit measures showed that the decision tree rule-based SPFs performed better than traditional SPFs in Texas.

Using data from Virginia, the study examined different functional forms and (variance and dispersion) structures of SPFs. The results revealed that the choice of the model formulation is of high importance in SPF development. The functional form of a model should match with the underlying data characteristics. Using mixture distributions, such as the NB-L and modifying variance and dispersion structures, allows different ways of introducing more flexibility into the model. We concluded that NB-L models provide a better fit when developing SPFs for NFAS roads. Also, the dispersion structure is highly dependent upon the underlying NB formulation. Different variance structures did not considerably change the model performance; however, in highly skewed datasets, flexible variance structures can provide more flexibility to the model. Inclusion of other variables considering different crash severity levels or crash types can further improve the proposed models.

The sensitivity analysis was performed to investigate the impact of AADT on the expected number of crash frequencies, and hence the impact on the final ranking of segments sorted by safety risk. The results suggest that higher-volume roads experience higher percentile rank changes compared to lower AADT roadway groups. Higher percent increases in AADT result in slightly higher percentile rank changes; however, the latter are not proportional to the AADT percent increase. The magnitude of this impact depends on several factors, such as the AADT coefficients of the SPFs (the smaller the coefficients, the smaller the impact), the sample size of the network (the bigger the network, the smaller the impact), and the overdispersion parameter (the higher the parameter, the smaller the impact), among others. We also observed that sites with low crash frequencies (e.g., one crash per year) are more sensitive to AADT increases than sites that exhibit more crashes. Overall, the NB-L models are much more flexible and tend to produce lower bias and therefore high variances, which in turn means that these models are more sensitive to any change in the data or the model parameters. On the other hand, as the expected values in the NB-L models comes from a full Bayesian procedure, both the expected values and their ranking are more reliable and more accurate than the ones derived from the EB method. As a conclusion, the trade-off between the sensitivity and reliability of the expected crash values needs to be accounted for before developing SPFs.

Additional Products

The Education and Workforce Development (EWD) and Technology Transfer (T2) products created as part of this project are described below and are listed on the Safe-D website here. The final project dataset is located on the Safe-D Dataverse.

Education and Workforce Development Products

Undergraduate and graduate courses:

- TTI/Texas A&M: CVEN 626 – Highway Safety (Fall 2021): Some of the material will be included in the slides and class notes for the graduate course CVEN 626. At the time this

report was written, the class notes had not been yet updated. They will be made available on Dr. Lord's website.

- TTI/Texas A&M: Some of the material has been included in Chapter 2 of the forthcoming textbook titled "Highway Safety Analytics and Modeling" co-written by Dr. Dominique Lord that will be published on March 1, 2021.

- UTC presentations: One presentation for the public will be hosted by the Safe-D UTC sometime in the spring of 2021.

Student Funding and Enrichment:

- TTI – one Ph.D. student, Ali Khodadadi, at the Texas A&M University. Title of dissertation to be determined. Status: anticipated December, 2021.

- TTI – one undergraduate student, Jessica Morris, at the University of Texas San Antonio.

For Ali Khodadadi, the project has been very beneficial. This project allowed Ali to enhance his knowledge in safety analysis and statistics, learn new programming languages, and publish papers. Jessica Morris learned how to assemble different types of traffic and roadway data, perform data quality control checks, process and analyze data in ArcGIS, link databases, and download data from SLD's Insight tool.

Technology Transfer Products

The main technology transfer products from this study include the following:

- New SPFs for NFAS roads that TxDOT and VDOT can use in data-driven safety analysis.

- Webinar – At the conclusion of this project, the researchers will conduct a webinar to present the methodology and project findings to students and stakeholders.

- Conference Paper – The research team prepared a conference paper that will be presented at the 100th Transportation Research Board annual meeting in 2021.

- Journal Article – The research team will prepare at least two more papers, which will be submitted to peer-reviewed transportation journals.

Data Products

The research team uploaded to the Safe-D Dataverse two databases (Texas_SPF_Data and Virginia_SPF_Data) along with their metadata for Texas and Virginia, respectively. The two datasets contain geometric (e.g., segment length, lane width, shoulder width), traffic volume, and crash counts for five years and different severity levels for NFAS roads in Texas and Virginia. The metadata describe the data, including the source, description and coding of categorical variables, and number of missing values.

References

1. Federal Highway Administration, *Highway Performance Monitoring System Field Manual*, Office of Highway Policy Information, Washington, D.C., December, 2016.

2. Federal Highway Administration, *Highway Safety Improvement Program Final Rule*, Docket No. FHWA-2013-0019. Federal Register, Vo. 81, No. 50, March 15, 2016.

3. Federal Highway Administration, *Highway Statistics 2016*, Office of Highway Policy Information, Washington, D.C., September, 2017.

4. American Association of State Highway and Transportation Officials (AASHTO). *Highway Safety Manual*. 1st Edition, 2010.

5. Haghighi, M., S. B. Johnson, X. Qian, K. F. Lynch, K. Vehik, and S. Huang. *A Comparison of Rule-Based Analysis with Regression Methods in Understanding the Risk Factors for Study Withdrawal in a Pediatric Study*. Scientific Reports, Vol. 6, No. 1, 2016, p. 30828.

6. Therneau, T. and B. Atkinson. *rpart: Recursive Partitioning and Regression Trees*. package version 4.1-15. 2019. https://CRAN.R-project.org/package=rpart Accessed August 1, 2020.

7. Williams, G. J. *Data Mining with Rattle and R: The Art of Excavating Data for Knowledge Discovery*. https://cran.r-project.org/web/packages/rattle/rattle.pdf Accessed August 1, 2020.

8. Wang, K. et al. Functional Forms of the Negative Binomial Models in Safety Performance Functions for Rural Two-Lane Intersections. *Accident Analysis & Prevention*, 124, 2019, pp. 193-201.

9. Zamani, H., and N. Ismail. Negative Binomial-Lindley Distribution and its Application. *Journal of Mathematics and Statistics*, 6(1), 2010, pp. 4-9.

10. Lord, D., and S. R. Geedipally. The Negative Binomial–Lindley Distribution as a Tool for Analyzing Crash Data Characterized by a Large Amount of Zeros. *Accident Analysis & Prevention*, 43(5), 2011, pp. 1738-1742.

11. Geedipally, S. R., et al. The Negative Binomial-Lindley Generalized Linear Model: Characteristics and Application Using Crash Data. *Accident Analysis & Prevention*, 45, 2012, pp. 258-265.

12. Lord, D., and P. Y.-J. Park. Investigating the Effects of the Fixed and Varying Dispersion Parameters of Poisson-Gamma Models on Empirical Bayes Estimates. *Accident Analysis & Prevention*, 40(4), 2008, pp. 1441-1457.

13. Cafiso, S. et al. Revisiting Variability of Dispersion Parameter of Safety Performance for Two-Lane Rural Roads. In *Transportation Research Record*: *Journal of the Transportation*

Research Board, No. 2148, TRB, National Research Council, Washington, D.C., 2010, pp. 38-46.

14. Geedipally, S. R., et al. Analyzing Different Parameterizations of the Varying Dispersion Parameter as a Function of Segment Length. In *Transportation Research Record*: *Journal of the Transportation Research Board, No. 2103,* TRB, National Research Council, Washington, D.C., 2009, pp. 108-118.

15. Lord, D., and S. R. Geedipally. Safety Prediction with Datasets Characterised with Excess Zero Responses and Long Tails. *Safe Mobility: Challenges, Methodology and Solutions*, Emerald Publishing Limited, 2018.

16. Avelar, R., S. Geedipally, S. Das, L. Wu, B. Kutela, D. Lord, and I. Tsapakis. Evaluation of Roadside Treatments to Mitigate Roadway Departure Crashes: Technical Report. Publication FHWA/TX-20/0-6991-R1. Texas Department of Transportation, 2020.

17. Das, S., Sun, X., Dixon, K., and M. Rahman. Safety Effectiveness of Roadway Conversion with a Two-Way Left Turn Lane. *Journal of Traffic and Transportation Engineering* (English Edition), Vol 5(4), 2018, pp. 309-317.

18. Das, S., M. Le, M. P. Pratt, and C. Morgan. Safety Effectiveness of Truck Lane Restrictions: A Case Study on Texas Urban Corridors. *International Journal of Urban Sciences*, 2020a. 24: pp. 35–49.

19. Das, S., S. R. Geedipally, and K. Fitzpatrick. Inclusion of Speed and Weather Measures in Safety Performance Functions for Rural Roadways. *IATSS Research*, 2020.

20. Dixon, K., K. Fitzpatrick, R. Avelar, and S. Das. Analysis of the Shoulder Widening Need on the State Highway System. *Publication FHWA/TX-15/0-6840-1*. Texas Department of Transportation, 2017.

21. Geedipally, S., Das, S., Pratt, M., and Lord D. Determining Skid Resistance Needs on Horizontal Curves for Different Levels of Precipitation. In *Transportation Research Record: Journal of the Transportation Research Board*, *No. 2674,* TRB, National Research Council, Washington, D.C., 2020.

22. Tsapakis, I., S. Sharma, B. Dadashova, S. Geedipally, A. Sanchez, M. Le, L. Cornejo, S. Das, and K. Dixon. Evaluation of Highway Safety Improvement Projects and Countermeasures: Technical Report. *Publication FHWA/TX-19/0-6961-R1*. Texas Department of Transportation, 2019.

23. Sun, X., and S. Das. A Comprehensive Study on Pavement Edge Line Implementation. *Report No. FHWA/LA.13/508*, 2013.

24. Pratt, M. P., S. R. Geedipally, B. Wilson, S. Das, M. Brewer, and D. Lord. Pavement Safety-Based Guidelines for Horizontal Curve Safety. *Publication FHWA/TX-18/0-6932-R1*. Texas Department of Transportation, 2018.

25. Zou, Y., Ash, J., Park, B., Lord, D., and L. Wu. Empirical Bayes Estimates of Finite Mixture of Negative Binomial Regression Models and its Application to Highway Safety. *Journal of Applied Statistics*, Vol. 45(9), 2017.

26. Turner, S., and P. Koeneman. Using Mobile Device Samples to Estimate Traffic Volumes. *Final Report 2017-49*, Prepared by the Texas A&M Transportation Institute for the Minnesota Department of Transportation, December 2017.

27. Greenwell, B., B. Boehmke, and B. Gray. *vip: Variable Importance Plots*. https://koalaverse.github.io/vip/index.html Accessed on August 1, 2020.

28. Hauer, E. *The Art of Regression Modeling in Road Safety*. Springer, 2015.

29. Watanabe, S., and M. Opper. Asymptotic Equivalence of Bayes Cross-Validation and Widely Applicable Information Criterion in Singular Learning Theory. *Journal of Machine Learning Research*, 11(12), 2010.

30. Turner, S., I. Tsapakis, and P. Koeneman. Evaluation of StreetLight Data's Traffic Count Estimates from Mobile Device Data. *Final Report 2020-30*, Prepared by the Texas A&M Transportation Institute for the Minnesota Department of Transportation, November, 2020.

31. Federal Highway Administration. *Independent Evaluation of Non-Traditional Methods to Obtain Annual Average Daily Traffic*. FHWA Pooled Fund Study, Performed by StreetLight Data Inc., the National Renewable Energy Lab, and Cambridge Systematics with the Texas A&M Transportation Institute, Washington, DC (ongoing project).

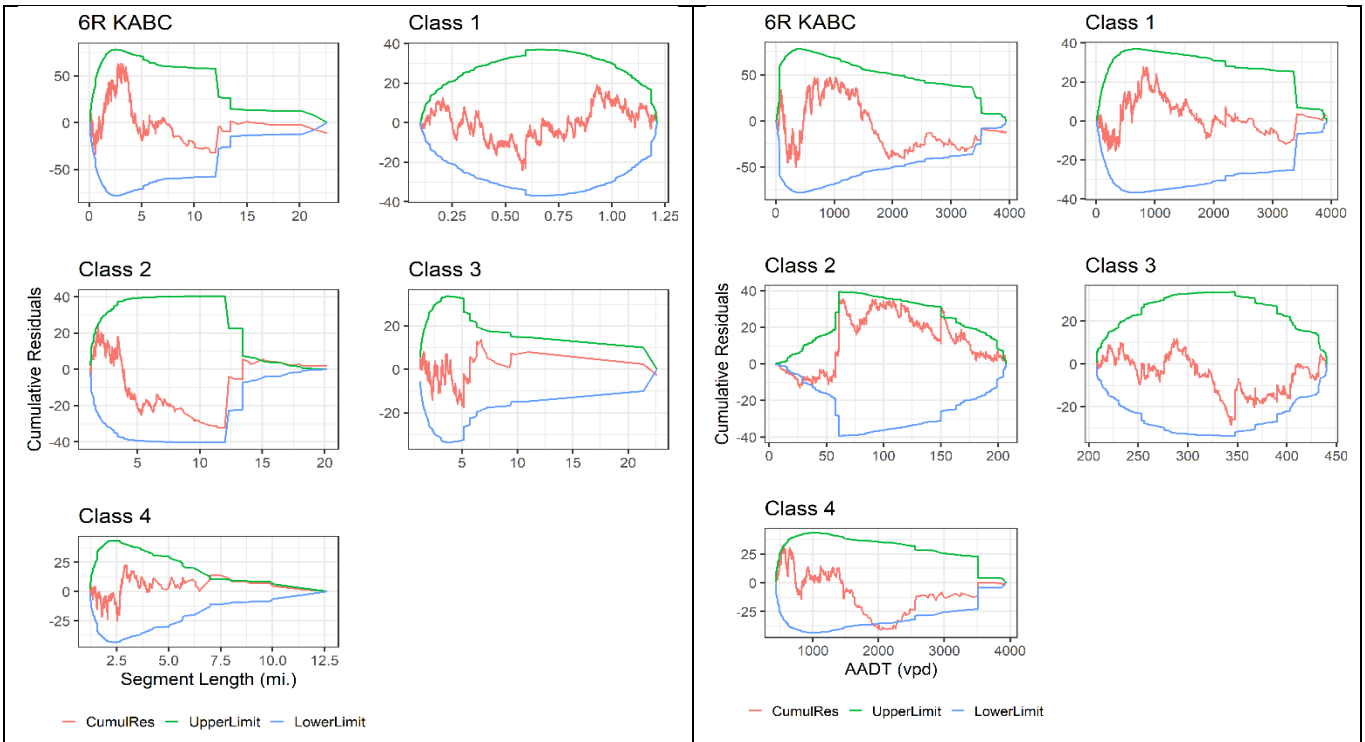
Appendix: Additional Analysis Results

This appendix includes various results from the analyses conducted in the study.

Table 5. Accuracy of StreetLight AADT Estimates by State and AADT Range

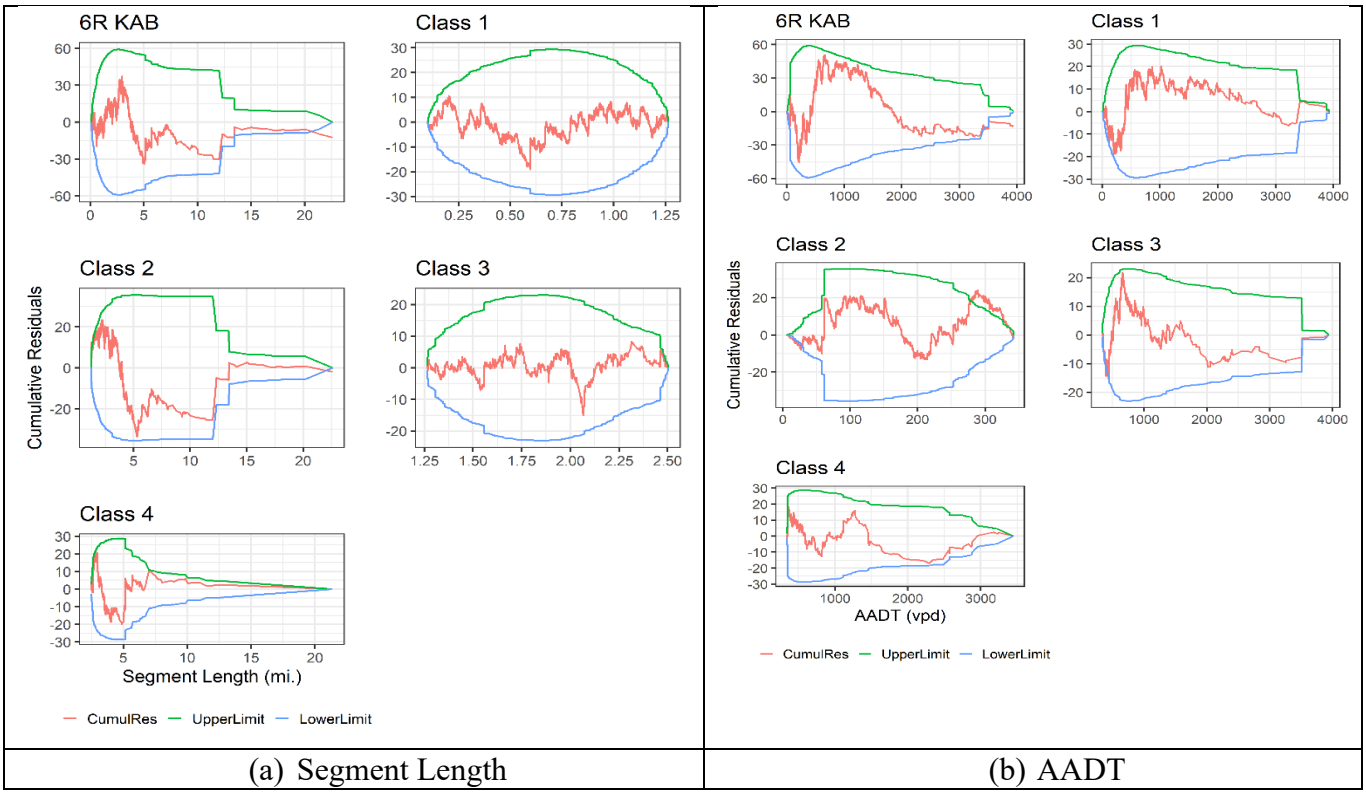
State	AADT Range (vehicles/day)	Number of Records	MSD	MAD	MAPE	Median APE	ACV
Texas	0–399	4,009	NA	NA	NA	NA	NA
	400–1,999	1,658	696	706	103%	79%	38%
	2,000–4,999	192	641	834	31%	25%	17%
	5,000-9,999	28	874	2,527	43%	18%	22%
	≥10,000	16	(1,838)	4,202	25%	28%	21%
Virginia	0–399	1,536	NA	NA	NA	NA	NA
	400–1,999	1,826	812	824	132%	89%	42%
	2,000–4,999	173	339	948	34%	26%	20%
	5,000-9,999	31	(387)	1,503	23%	25%	17%
	≥10,000	16	(4,810)	5,666	35%	37%	31%
Grand Total		**9,485**	**691**	**831**	**108%**	**77%**	**38%**

NA = Not applicable



(a) Segment Length (b) AADT

Figure 7. CURE plots for KABC model.



(a) Segment Length (b) AADT

Figure 8. CURE plots for KAB model.

Figure 9. Predicted KABCO crashes (Texas SPFs) against AADT.

Figure 10. Predicted KABC crashes (Texas SPFs) against AADT.

Table 6. Model Estimation Results (Length and AADT Dependent Dispersion Structure) for All NFAS Roads: Dispersion Structure $\emptyset_i = e^{\eta_0} AADT_i^{\eta_1} L_i^{\eta_2}$

	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Intercept ($\boldsymbol{\beta_0}$)	-4.89 (0.19)	-4.80 (0.18)	-4.91 (0.20)	-4.79 (0.26)	-4.84 (0.25)	-4.80 (0.27)
Ln(AADT) ($\boldsymbol{\beta_1}$)	0.77 (0.03)	0.77 (0.03)	0.77 (0.03)	0.74 (0.08)	0.73 (0.07)	0.75 (0.08)
Length ($\boldsymbol{\beta_2}$)	0.56 (0.02)	0.50 (0.01)	0.58 (0.02)	0.59 (0.03)	0.62 (0.02)	0.56 (0.03)
$\boldsymbol{\eta_0}$	-5.83 (0.64)	-1.98 (0.63)	-7.87 (0.85)	~~-2.70 (1.90)~~	~~0.90 (1.94)~~	-5.57 (1.95)
$\boldsymbol{\eta_1}$	1.15 (0.10)	0.40 (0.09)	1.56 (0.15)	1.17 (0.29)	~~0.58 (0.29)~~	1.69 (0.30)
$\boldsymbol{\eta_2}$	1.72 (0.06)	1.25 (0.07)	2.12 (0.12)	3.36 (0.30)	3.25 (0.36)	3.73 (0.33)
P	-	-	3.61 (0.16)	-	-	~~0.068 (0.07)~~
WAIC	8448	8509	8438	7914	7910	7938
LOO	8448	8509	8438	8401	8426	8293
MASE	0.56	0.56	0.57	0.21	0.20	0.24
MSPE	5.57	5.09	5.87	0.60	0.52	0.74
Log-Likelihood	-4221	-4251	-4215	-3519	-3503	-3538

Table 7. Model Estimation Results (Length and AADT Dependent Dispersion Structure) for All NFAS Roads: Dispersion Structure $\emptyset_i = e^{\eta_0} AADT_i^{\eta_1} L_i$

	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Intercept ($\boldsymbol{\beta_0}$)	-4.94 (0.20)	-4.86 (0.19)	-4.90 (0.19)	-5.3 (0.30)	-5.20 (0.27)	-5.03 (0.25)
Ln(AADT) ($\boldsymbol{\beta_1}$)	0.75 (0.03)	0.7 (0.03)	0.76 (0.03)	0.77 (0.08)	0.76 (0.07)	0.73 (0.07)
Length ($\boldsymbol{\beta_2}$)	0.63 (0.02)	0.53 (0.01)	0.56 (0.02)	0.74 (0.03)	0.69 (0.03)	0.69 (0.02)
$\boldsymbol{\eta_0}$	-2.55 (0.70)	-1.48 (0.61)	-2.12 (0.66)	~~2.33 (5.28)~~	~~6.15 (3.13)~~	~~4.98 (2.56)~~
$\boldsymbol{\eta_1}$	0.65 (0.11)	0.32 (0.09)	0.45 (0.11)	~~0.21 (0.76)~~	~~0.56 (0.45)~~	~~0.44 (0.38)~~
$\boldsymbol{\eta_2}$	-	-	-	-	-	-
P	-	-	1.77 (0.11)	-	-	3.12 (0.31)
WAIC	8584	8520	8519	8220	8110	8068
LOO	8584	8520	8518	8631	8565	8578
Mean Absolute Scaled Error	0.57	0.55	0.55	0.25	0.20	0.19
Mean Squared Prediction Error	6.51	5.19	5.40	1.43	0.55	0.50
Log-Likelihood	-4289	-4257	-4255	-3650	-3605	-3548

Table 8. Model Estimation Results (Length only Dependent Dispersion Structure) for All NFAS Roads: Dispersion Structure $\emptyset_i = e^{\eta_0} L_i^{\eta_2}$

	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Intercept ($\boldsymbol{\beta_0}$)	-4.80 (0.19)	-4.87 (0.18)	-4.88 (0.18)	-5.01 (0.25)	-4.95 (0.25)	-5.05 (0.27)
Ln(AADT) ($\boldsymbol{\beta_1}$)	0.76 (0.03)	0.78 (0.03)	0.78 (0.04)	0.77 (0.07)	0.75 (0.07)	0.78 (0.08)
Length ($\boldsymbol{\beta_2}$)	0.55 (0.02)	0.50 (0.01)	0.51 (0.02)	0.60 (0.03)	0.62 (0.02)	0.59 (0.03)
$\boldsymbol{\eta_0}$	1.67 (0.08)	0.61 (0.07)	0.80 (0.16)	5.14 (0.58)	4.170(0.70)	6.15 (0.56)
$\boldsymbol{\eta_1}$	-	-	-	-	-	-
$\boldsymbol{\eta_2}$	1.56 (0.07)	1.21 (0.07)	1.27 (0.08)	3.44 (0.31)	3.2 (0.34)	3.91 (0.33)
P	-	-	1.81 (0.16)	-	-	~~0.13 (0.12)~~
WAIC	8549	8522	8523	7920	7913	7920
LOO	8549	8522	8523	8396	8411	8384
MASE	0.55	0.56	0.56	0.21	0.20	0.22
MSPE	5.35	5.1	5.13	0.62	0.53	0.70
Log-Likelihood	-4271	-4257	-4257	-3518	-3505	-3514

Table 9. Model Estimation Results (Length only Dependent Dispersion Structure) for All NFAS Roads: Dispersion Structure $\emptyset_i = e^{\eta_0} L_i$

	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Intercept (β_0)	-4.87 (0.20)	-4.92 (0.19)	-4.91 (0.19)	-5.40 (0.26)	-5.09 (0.25)	-4.95 (0.25)
Ln(AADT) (β_1)	0.74 (0.03)	0.78 (0.03)	0.78 (0.03)	0.79 (0.07)	0.75 (0.08)	0.72 (0.07)
Length (β_2)	0.62 (0.02)	0.53 (0.01)	0.53 (0.02)	0.73 (0.02)	0.69 (0.02)	0.69 (0.02)
η_0	1.68 (0.08)	0.57 (0.06)	0.53 (0.13)	3.61 (0.30)	2.23 (0.24)	2.02 (0.25)
η_1	-	-	-	-	-	-
η_2	-	-	-	-	-	-
P	-	-	2.04 (0.12)	-	-	3.07 (0.30)
WAIC	8615	8529	8529	8214	8120	9070
LOO	8615	8528	8528	8623	8577	8572
MASE	0.56	0.55	0.56	0.25	0.20	0.19
MSPE	6.10	5.16	5.15	1.52	0.56	0.49
Log-Likelihood	-4305	-4262	-4262	-3469	-3616	-3551
Mean Absolute Scaled Error	0.54	0.53	0.54	0.15	0.16	0.18
Mean Squared Prediction Error	6.75	6.75	7.10	0.45	0.52	0.77
Log-Likelihood	-3095	-3077	-3058	-2536	-2555	-2574

Table 10. Model Estimation Results for Urban Local Roads

Dispersion Structure	$\emptyset_i = e^{\eta_0} AADT_i^{\eta_1} L_i^{\eta_2}$			$\emptyset_i = e^{\eta_0} L_i^{\eta_2}$		
	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Intercept (β_0)	-3.32 (0.69)	-3.34 (0.67)	-3.31 (0.67)	-3.34 (0.75)	-3.36 (0.76)	-3.36 (0.76)
Ln(AADT) (β_1)	0.45 (0.10)	0.44 (0.10)	0.44 (0.10)	0.45 (0.12)	0.44 (0.18)	0.48 (0.23)
Length (β_2)	0.85 (0.09)	0.98 (0.16)	0.93 (0.16)	1.02 (0.16)	1.06 (0.18)	1.05 (0.18)
η_0	~~-1.12 (2.38)~~	~~-0.48 (2.11)~~	~~0.42 (2.57)~~	4.16 (1.83)	4.70 (1.75)	4.58 (1.88)
η_1	~~-0.21 (0.33)~~	~~0.16 (0.30)~~	~~-0.02 (0.42)~~	-	-	-
η_2	-0.43 (0.44)	~~0.10 (0.37)~~	~~-0.06 (0.64)~~	~~-1.22 (1.86)~~	~~-0.89 (1.73)~~	~~-1.0 (1.81)~~
P	-	-	1.63 (0.80)	-	-	1.51 (1.04)
WAIC	880	880	881	813	813	815
LOO	880	880	882	859	860	860
Mean Absolute Scaled Error	0.68	0.71	0.70	0.32	0.33	0.33
Mean Squared Prediction Error	2.79	4.31	3.42	0.62	0.95	0.82
Log-Likelihood	-437	-437	-437	-350	-351	-351

Table 11. Model Estimation Results for Rural Local Roads

Dispersion Structure	$\emptyset_i = e^{\eta_0} AADT_i^{\eta_1} L_i^{\eta_2}$			$\emptyset_i = e^{\eta_0} L_i^{\eta_2}$		
	NB-1	NB-2	NB-P	NB1-L	NB2-L	NBP-L
Intercept ($\boldsymbol{\beta_0}$)	-3.89 (0.58)	-4.04 (0.68)	-3.81 (0.85)	-3.68 (0.78)	-3.80 (0.81)	-3.77 (0.84)
Ln(AADT) ($\boldsymbol{\beta_1}$)	0.51 (0.08)	0.53 (0.09)	0.51 (0.12)	0.51 (0.24)	0.52 (0.24)	0.51 (0.24)
Length ($\boldsymbol{\beta_2}$)	0.75 (0.06)	0.77 (0.08)	0.71 (0.09)	0.69 (0.08)	0.67 (0.08)	0.68 (0.08)
$\boldsymbol{\eta_0}$	~~14.54 (4.22)~~	~~9.88 (5.15)~~	~~0.76 (6.08)~~	4.66 (1.52)	4.79 (1.46)	4.98 (1.51)
$\boldsymbol{\eta_1}$	~~1.84 (0.59)~~	~~1.18 (0.69)~~	~~0.15 (0.88)~~	-	-	-
$\boldsymbol{\eta_2}$	~~3.73 (0.89)~~	~~1.63 (1.01)~~	~~1.09 (1.41)~~	~~3.23 (1.53)~~	3.66 (1.35)	3.55 (1.58)
P	-	-	~~3.16 (0.59)~~	-	-	~~1.81 (1.10)~~
WAIC	981	983	985	900	902	903
LOO	981	984	985	952	960	954
Mean Absolute Scaled Error	0.67	0.69	0.66	0.23	0.24	0.24
Mean Squared Prediction Error	7.45	8.68	6.77	0.38	0.40	0.40
Log-Likelihood	-487	-486	-484	-394	-397	-396