Investigating the Impact of Lack of Motorcycle Annual Average Daily Traffic Data in Crash Modeling and the Estimation of Crash Modification Factors

PUBLICATION NO. FHWA-HRT-16-054

OCTOBER 2016

U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

FOREWORD

Advanced crash prediction methods, including the use of safety performance functions and the Empirical Bayes method, have become the standard for traffic safety decisionmakers. Safety practitioners employ these methods as a part of their safety decisionmaking process for many roadway and crash types. However, there has been very little research conducted to extend these methods specifically to the prediction of motorcycle crashes. This deficiency may be in part related to a lack of traffic count data that specifically identifies motorcycles. Motorcycle-focused average annual daily traffic (AADT) information is critical to these types of assessments in order to properly account for the exposure of drivers and motorcycle riders. However, motorcycle crashes continue to be a significant safety concern on U.S. highways; they account for over 14 percent of all fatalities.

The research team developed numerous statistical models with and without motorcycle AADT using crash and traffic records from three states (Florida, Pennsylvania, and Virginia). Ultimately, it was found that while accurate counts for motorcycle AADT are preferred, in many cases, it is appropriate to use the total AADT as a surrogate. This finding will be valuable for State and local safety analysts who want to better understand the scope of motorcycle safety risks and explore options to reduce the number of motorcycle crashes and fatalities on their roads.

Monique R. Evans
Director, Office of Safety
Research and Development

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA-HRT-16-054	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle Investigating the Impact of Lack of Motorcycle Annual Average Daily Traffic Data in Crash Modeling and the Estimation of Crash Modification Factors	5. Report Date October 2016 6. Performing Organization Code:
7. Author(s) Craig Lyon, Bhagwant Persaud, Robert Scopatz, Scott Himes, Matt Albee, and Thanh Le	8. Performing Organization Report No.
9. Performing Organization Name and Address VHB 8300 Boone Blvd., Ste. 700 Vienna, VA 22182-2626 Persaud Lyon, Inc. 87 Elmcrest Road Toronto, Ontario M9C 3R7	10. Work Unit No. 11. Contract or Grant No. DTFH61-13-D-00001
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Highway Administration 1200 New Jersey Avenue, SE Washington, DC 20590	13. Type of Report and Period Covered Safety Evaluation 14. Sponsoring Agency Code Federal Highway Administration

15. Supplementary Notes
The Federal Highway Administration (FHWA) Office of Safety Research and Development managed this study. The FHWA Office of Safety Research and Development Contract Task Order Manager was Craig Thor.

16. Abstract
The development of safety performance functions (SPFs) and crash modification factors (CMFs) requires data on traffic exposure. The analysis of motorcycle crashes can be especially challenging in this regard because few jurisdictions collect motorcycle traffic volume data systematically. To address this challenge, the project team conducted several analyses to explore (1) how much predictive power for an SPF is lost when motorcycle volumes are unknown and how this lack of information may affect the development of CMFs for motorcycle crashes, and (2) alternative methods for deriving accurate predictions of motorcycle crashes or motorcycle volumes. The results of the analyses show that when motorcycle volumes are not known, using total average annual daily traffic (AADT) on its own is sufficient for developing SPFs and CMFs. The potential bias due to missing motorcycle-specific AADT is sufficiently negligible where it exists so as not to preclude SPF and CMF development. The project team also concluded that attempting to predict motorcycle volumes is not possible using typically available roadway and county-level data. Improvement could possibly be found in trip generation type modeling at a disaggregate scale, although given the success of SPF development using total AADT, such an effort may not be worthwhile. A more significant issue in developing motorcycle crash SPFs and CMFs is working with relatively rare crash types. In the analyses undertaken, SPFs could not be developed for all motorcycle crash types or site types. More evidently, in the estimation of CMFs using simulated data, the CMF value varied significantly between simulation runs due to the low frequency of motorcycle crashes. In terms of research gaps, a database is needed that includes implemented countermeasures expected to affect motorcycle crashes along with the location, date of treatment, and treatment description. This information would aid researchers in identifying treatments that are feasible for study. The report also identifies several research gaps related to analytical methods, related gaps, and data limitations.

17. Key Words Motorcycle crashes, safety performance functions, crash modification factors, safety evaluations, average annual daily traffic, AADT	18. Distribution Statement No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161. http://www.ntis.gov

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 146	22. Price N/A

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

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AADT	annual average daily traffic
ADT	average daily traffic
BAC	blood alcohol concentration
CMF	crash modification factor
CURE	cumulative residuals
EB	Empirical Bayes
FARS	Fatal Accident Reporting System
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration
GIS	geographic information system
GLM	generalized linear modeling
HSM	*Highway Safety Manual*
HSIP	Highway Safety Improvement Program
IR	infrared
KABCO	killed, A injury, B injury, C injury, property damage only
LRS	linear referencing system
MAD	mean absolute derivation
MMUCC	model minimum uniform crash criteria
PAR	police accident report
PDO	property damage only
RLC	red light camera
RR	relative risk
SE	standard error
SPF	safety performance function
STD	standard deviation
TIRTL	The Infra-Red Traffic Logger
TRIS	Transportation Research Information Service
VDOT	Virginia Department of Transportation
VMT	vehicle-miles traveled

EXECUTIVE SUMMARY

The objective of this project was to investigate and describe the effect of the lack of motorcycle annual average daily traffic (AADT) data on the performance of motorcycle safety evaluations aimed at developing motorcycle crash-specific crash modification factors (CMFs) and safety performance functions (SPFs). Motorcycle volumes would intuitively be required for modeling motorcycle crashes and studying the safety effectiveness of countermeasures due to the strong relationship between crashes and exposure. However, few jurisdictions collect motorcycle traffic volume data systematically. A second purpose of the research was to investigate and demonstrate methods and provide the mathematical models required for jurisdictions that lack motorcycle volumes when undertaking the evaluation of motorcycle related safety countermeasures.

The project scope included the following tasks:

- Identify current practices for statistical modeling of motorcycle crashes and motorcycle countermeasure evaluations.
- Identify the availability, applicability, quality, and other related features of motorcycle safety data resources.
- Perform a quantitative analysis of the effect of the lack of motorcycle AADT on safety evaluations.
- Determine data limitations, gaps, and future data needs to support future motorcycle safety evaluations.
- Identify opportunities to promote the findings and support follow up research through a marketing, communications, and outreach plan.

A literature review found many studies concerned with motorcycle safety, but very few focused on the prediction of crash frequency. The limited international research does suggest that motorcycle crash frequency models can be developed based only on total AADT for all vehicle types.

Guided by lessons learned from the literature review, especially the promise in seeking alternatives to directly including motorcycle volumes in motorcycle crash frequency prediction models, a number of analytical approaches were developed and undertaken. The intention of the models were to serve two broad purposes, namely the following:

- To explore how much predictive power for an SPF is lost when motorcycle volumes are unknown and how this lack of information may affect a study of motorcycle countermeasures estimating a CMF.
- To explore alternative methods for deriving accurate predictions of motorcycle crashes or motorcycle volume data.

The project team investigated two groups, or avenues, of methods. The methods for avenue A focused on investigating (1) the difference in predictive performance for motorcycle SPFs calibrated with motorcycle AADT versus total AADT, (2) the relation of total crash SPFs and motorcycle crash SPFs so jurisdictions without motorcycle volumes could predict motorcycle crashes using total crash SPFs, and (3) methods to predict segment-level motorcycle AADT.

The methods for avenue B focused on the differences in CMF estimates found when using motorcycle AADT versus total AADT when applying before-after or cross-sectional regression CMF estimation methods on simulated crash count data.

For developing the avenue A models, data were collected from Florida and Pennsylvania. Both States had a large number of locations with an estimated motorcycle AADT, which could provide linkable roadway inventory, traffic, and crash data. Virginia also provided data for the purpose of validating the models developed. The avenue B analyses used the roadway inventory, total AADT, and motorcycle AADT collected for the avenue A methods in Florida and Pennsylvania. For motorcycle crashes, SPFs developed in the avenue A models simulated crash counts as the initial starting point.

The findings of both the avenue A and avenue B modeling indicate that when motorcycle volumes are unknown, using total AADT on its own is sufficient for developing SPFs and CMFs. The potential bias due to missing motorcycle-specific AADT is sufficiently negligible, where it exists so as not to preclude SPF and CMF development. However, in the analysis undertaken, SPFs could not be developed for all motorcycle crash or site types. This is a significant issue in working with relatively rare crash types.

The findings also conclude that attempting to predict motorcycle volumes is not possible using typically available roadway and county-level data. Improvements could possibly be made in trip generation-type modeling at a disaggregate scale, although, given the success of the SPFs using total AADT, such an effort may not be worthwhile.

The research identified a number of data limitations and motorcycle SPF and CMF research gaps through the assessment of available data sources, analytic methods, and evaluation results. Data limitations identified relate to traffic volumes (AADT) addressing technology requirements in particular; crash data focusing on quality issues; and roadway inventory data emphasizing roadway class/ownership and missing data issues. For research gaps with respect to motorcycle safety and CMFs, very little information is known on the effects of roadway geometric and traffic control features on motorcycle crash frequency and severity. The reasons for this gap are likely twofold: motorcycle crashes are not usually the focus of safety-related countermeasures, and the rarity of motorcycle crashes combined with scarcity of treatment locations would result in a small sample size for study. With respect to SPFs for application in network screening and other safety management tasks, few SPFs at the segment level or intersection level exist. The SPFs developed in this project may contribute to filling this void, but there remains work to be done in terms of evaluating site types for which no SPF was developed and ensuring that SPFs exist that calibrate well in all jurisdictions.

CHAPTER 1. INTRODUCTION

The objective of the project was to investigate and describe the effect of the lack of motorcycle annual average daily traffic (AADT) data on the performance of motorcycle safety evaluations aimed at developing motorcycle crash-specific crash modification factors (CMFs) and safety performance functions (SPFs).

The project scope included the following tasks:

- Identify current practices for statistical modeling of motorcycle crashes and motorcycle countermeasure evaluations.
- Identify the availability, applicability, quality, and other related features of motorcycle safety data resources.
- Perform a quantitative analysis of the effect of the lack of motorcycle AADT on safety evaluations.
- Determine data limitations, gaps, and future data needs to support future motorcycle safety evaluations.
- Identify opportunities to promote the findings and support follow up research through a marketing, communications, and outreach plan.

This final project report organizes the findings of these tasks as follows:

- Chapter 2 discusses current practices for assessing motorcycle safety.
- Chapter 3 discusses the analysis methods.
- Chapter 4 discusses the data collection and a summary of the data.
- Chapter 5 discusses the analysis and results.
- Chapter 6 provides conclusions and recommendations.
- Chapter 7 discusses data and analysis limitations and future research needs.

CHAPTER 2. CURRENT PRACTICES FOR ASSESSING MOTORCYCLE SAFETY

A literature search identifying current practices in analyzing motorcycle crash data assessed the applied statistical methods and data used. The Transportation Research Information Service (TRIS) and International Transport Research Documentation databases were the primary resources for published research involving statistical analysis of motorcycle safety. TRIS includes the capability to search several databases, including the Highway Research Information Service database for domestic literature, the Highway Research in Progress database for ongoing research studies, and the International Road Research database for international literature. Additionally, the project team searched the national and international contacts for relevant research from local and State agencies that are not published or otherwise not widely available.

As expected, the review of literature found that, by far, the most widely used analytic methods applied to motorcycle safety data belong to the family of discrete outcome models. Such models are applied for estimating the impact of crash and/or behavioral characteristics on the type of crash that occurs. For example, a discrete choice model may predict the probability of a given level of severity, given that a crash has occurred, with predictor variables describing the rider, motorcycle, and roadway (e.g., use/non-use of helmet, engine size, type of roadway, or speed limit). It is important to note that such models do not provide an estimate of the expected crash frequency and cannot be directly applied for developing SPFs or CMFs. (They can, however, be used to develop severity distribution functions that can be applied to an SPF for total crashes, if such an SPF exists, to estimate crash frequency by severity.)

The review is divided into sections on crash frequency models, discrete choice "probabilistic models," and models that do not clearly fit either category. A summary of key findings appears at the end of this chapter.

CRASH FREQUENCY MODELS

Where motorcycle crash frequency models exist, they have the most direct relevance to the development of motorcycle-specific SPFs and CMFs. These models can be used to estimate expected crashes or to estimate CMFs through cross-section studies or before-after studies. The literature provides few examples of developing models to predict motorcycle crash frequency. The literature review identified four such studies regarding developing crash frequency models in recent years. A review of these studies follows.

Flask et al. applied a fully Bayesian multi-level fixed effects model to estimate expected multi-vehicle motorcycle crash frequencies on road segments in Ohio.(1) Three datasets were used in this study. The Ohio Department of Transportation provided the first dataset, which was composed of 32,289 interstate, U.S. route, and State route segments. This dataset included the following variables: pavement type, lane width, shoulder width, number of lanes, median presence, horizontal and vertical curve-related statistics, the overall vehicle average daily traffic (ADT), and segment length. In addition to the roadway segments, township information was available, including the number of lane-miles, area of the township, and the urban status of the township. All of these variables were considered as fixed effects parameters with the exception of the ADT and segment length, which were assumed to have a linear relationship with crashes.

U.S. Census data provided demographic information for the different regions of Ohio, such as the percentage of residents over the age of 65, percentage of residents under the poverty level, and the mean travel time to work. In addition to demographic information, the county population, number of motorcycle endorsements (motorcycle licenses), and number of registered motorcycles were used as measures of motorcycle and motor vehicle traffic and were compiled at a regional level. Spatial correlation through conditional autoregressive random effects were included in the model and were shown to reduce the model error by adding the prior knowledge of neighboring regions and segments, leading to better parameter estimates. The distance between neighbors was measured using the distance between two segments in any direction. Random effect terms may be used to reduce model error that is caused by unavailable or unrecorded data, such as motorcycle-specific vehicle-miles traveled (VMT) or motorcycle AADT. Variables found to affect motorcycle crashes in addition to total AADT included the number of lane miles in a township, urban versus rural area type, the number of motorcycle endorsements, county population, and mean travel time to work reported by county.

Including the spatial random effects in the Flask et al. model reduced model error due to the unavailability of motorcycle AADT data in the dataset, but the resulting model would not provide the same benefits if applied to another jurisdiction.(1) The consideration of spatial correlation will account for unobserved correlation in motorcycle volumes at nearby locations, but that is specific to the dataset the model is developed with.

French and Gumus studied the relationship between motorcycle fatalities and economic activity using Fatality Analysis Reporting System (FARS) data.(2) The U.S. Bureau of Economic Analysis and the U.S. Bureau of Labor Statistics provided data on real income per capita and unemployment rate, respectively. Other potential explanatory variables included motorcycle registrations, rural VMT, urban VMT, alcohol taxes, average temperature and precipitation, gasoline prices, and data on safety programs affecting motorcycle riders (e.g., introduction of a helmet law). Total fatal crashes and crashes disaggregated by crash type, day, time, and level of rider's blood alcohol concentration (BAC) were studied with fatalities per 100,000 population as the dependent variable. The project team used a generalized linear model with log-link function and included both year and State fixed effects. Each State-year observation was weighted by the square root of the State population. Among the findings, estimates suggested a 10-percent increase in real income per capita is associated with a 10.4-percent increase in motorcycle fatality rates.

Haque et al. sought to identify factors affecting motorcycle crashes at three- and four-legged signalized intersections in Singapore by developing Bayesian crash prediction models.(3) Explanatory variables included intersection geometry and total traffic volume. It is important to note that due to the high use of motorcycles in Singapore, it is likely that motorcycle volume is highly correlated to total traffic volume. Other variables besides traffic volumes affecting expected crash frequency included number of lanes, presence of wide median, uncontrolled left-turn lane, presence of exclusive right-turn lane, and presence of red-light cameras (RLCs).

Manan et al. developed a generalized linear model with negative binomial error structure to predict fatal motorcycle crashes on Malaysian primary roads.(4) Explanatory variables included motorcycle AADT and access points per hour (kilometer). Separate models using total AADT and motorcycle AADT were calibrated. The model using total AADT had slightly better overall goodness-of-fit measures, including an overdispersion parameter of 0.821 compared to 0.872 for the model using motorcycle AADT. However, the model using motorcycle AADT demonstrated some improvement in a cumulative residuals (CURE) plot versus access points per kilometer, indicating that the model was less biased with respect to access point density. The authors rightly point out that the model with motorcycle AADT would be more sensitive to modal shifts. According to the authors, a model with both motorcycle and non-motorcycle volumes was not attempted because the two measures were so closely correlated with a Pearson's correlation coefficient of 0.913.

DISCRETE CHOICE MODELS

By a large margin, most of the literature on modeling motorcycle crashes uses discrete choice models. This family of models predicts the probability of an outcome given that a crash has occurred based on the values of explanatory variables. Researchers focused on motorcycle safety typically apply discrete choice models to estimate the impact of both road user and roadway variables on the likelihood of a specified outcome severity given that a crash occurs. For example, researchers may examine the impact of helmet use or roadway curvature on the severity of crashes.

Within the family of discrete choice models, there are many different modeling approaches, but these are essentially based on two different types of models: unordered or ordered discrete outcome models.

For unordered models, logistic regression models (sometimes referred to as "logit models") are used to refer specifically to the problem in which the dependent variable is binary (only two possible outcomes), while problems with more than two categories are referred to as "multinomial logistic regression."

For ordered models, multiple categories are possible and are considered to be ordered in some logical way, such as severity data on the killed, A injury, B injury, C injury, and property damage only (PDO) (KABCO) scale. By considering crash severity as ordered, it is not assumed that the difference between an O and C crash is the same as between a B and K crash, for example. These models may be ordered logit or ordered probit, where the difference is the assumed distribution of the model error term and link function.

The following subsections illustrate examples of three discrete choice models applied to the study of motorcycle safety. The review is by no means comprehensive, given that the focus of the project (and the review) is on crash frequency modeling and the sheer volume of literature on crash severity modeling necessitated some selectivity.

Logistic Regression Models

Logistic regression models identify factors that affect the likelihood of an outcome—such as a crash resulting in a fatality—and can be used to predict the outcome of an event. Logistic models apply when only two outcomes are possible. Figure 1 displays a logistic model.

$$P(Y = 1) = \frac{exp(x\beta)}{1 + exp(x\beta)}$$

Figure 1. Equation. Logistic model.

Where:

$P(Y = 1)$ = The probability that the outcome was observed.
χ = The characteristics of the person, crash, etc.
β = The parameters of the model to be estimated.

In estimating the model parameters, figure 2 shows the LN of the odds (i.e., the logit).

$$LN \frac{P_i}{(1 - P_i)} = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n$$

Figure 2. Equation. Logistic model parameter estimation.

The *odds ratio* is defined as the probability of the outcome occurring divided by the probability of the alternate. For example, if the probability of a fatality in the event of a crash were $^1/_{10}$, then the odds ratio would be $(^1/_{10})/(^9/_{10}) = 0.11$. Taking the exponent of the estimated parameters reveals the amount by which the odds ratio increases or decreases as the independent variable changes by one unit.

Kim et al. developed a logistic regression model to explain the likelihood of alcohol impairment among crash involved motorcycle riders in police reported motorcycle crashes.(5) The basic logistic model is shown in figure 3.

$$LN \frac{\Pr(I)}{(1 - \Pr(I))} = a_0 + a_1 A + a_2 A^2 + a_3 W + a_4 N + a_5 O$$

Figure 3. Equation. Kim et al. logistic regression model.(5)

Where:

$\Pr(I)$ = The probability of impairment.
a = The parameters of the model to be estimated.
A = Age of rider.
W = Weekend occurrence.
N = Nighttime occurrence.
O = Non-resident status.

The developed models also attempted additional variables. Results indicated that impairment was more likely to be a factor for middle-aged riders, unlicensed riders, and riders who did not wear a helmet and that impairment-related crashes are more likely to occur at night, on weekends, and in rural areas.

Connor applied logistic regression to motorcycle fatality data where coroner and police reports were available and license endorsement status was known.(6) The goal was to find what characteristics increased the likelihood that a fatal motorcycle crash involved an unendorsed rider. These characteristics included single-vehicle crashes, younger drivers, and driver's license suspensions in the past 7 years.

Akaateba et al. applied logistic regression to roadside observations of helmet use at 12 randomly selected sites.(7) The authors estimated odds ratios, adjusted odds ratios, and 95-percent confidence intervals for variables associated with helmet use. Female riders as well as riding during the weekdays and morning periods and at locations within central business districts showed higher helmet wearing rates.

Gabauer developed logistic models to predict rider injury for motorcyclists impacting longitudinal barriers.(8) Rider characteristics such as helmet usage and alcohol involvement were found to have a larger influence on injury severity in comparison to associated roadway characteristics.

Theofilatos and Yannis investigated the relationship between stated attitudes and behaviors with respect to safety and crash involvement of motorcyclists in Europe based on a survey.(9) Principal component analysis of the 38 variables collected through the survey grouped variables that showed a similar variance and effect on the probability of having been in a crash. A logistic regression model was used to model the probability of having been in a crash as a function of the declared attitudes and behaviors, age and declared exposure.

Keall et al. used a case-control study design to quantify fatality risks for motorcyclists based on BAC.(10) The authors acquired case (i.e., crash) data from police reports and post-mortem data. Control data were collected roadside including BAC tests. The authors used a logistic regression to model the data. The results show a much higher risk of fatality, even at low levels of BAC. At a BAC of 0.03, the fatality risk was 3 times higher; at a BAC of 0.08, the fatality risk was 20 times higher.

Haworth et al. applied a case-control approach to collecting data for 222 motorcycle crashes (cases) and 1,195 non-crash involved (controls) motorcyclist trips past a crash site at the same time of day and day of week of crash occurrence.(11) Data collection included detailed information of each crash, a comparison of features of cases and controls, and motorcycle exposure information. The controls comprised three groups. One group included riders who did not stop. A second included riders who stopped and were interviewed roadside. A third group included riders who gave a roadside interview and a follow-up interview. Odds ratios were estimated through conditional logistic regression. The approach was termed “conditional” because cases were matched to their controls by day, time, and location, and other confounding variables such as age were included in the model as explanatory variables. Some of the factors found to increase crash risk included age under 25, never married, unlicensed, increased BAC, use of a sidecar, motorcycle engine over 750 cc, and rider not being the owner of the motorcycle.

Kim and Boski developed a logistic regression model for the probability of being at fault in a crash for motorcyclists and drivers using temporal, roadway, and environmental factors.(12)

Multinomial Logistic Models

Multinomial logistic (or logit) models apply when more than two outcomes are possible and there is no ordering to the outcomes. A multinomial logistic model is shown in figure 4.

$$P(Y = i) = \frac{exp(\beta_i x_i)}{\sum_i exp(\beta_i x_i)}$$

Figure 4. Equation. Multinomial logistic model.

Where:

P(*Y* = *i*) = The probability that the outcome, *i*, was observed given the family of possible outcomes, *I*.

χ = The characteristics of the person, crash, etc.

β = The parameters of the model to be estimated.

Each possible outcome, *i*, has its own set of explanatory variables and parameters, otherwise known as the "utility function."

Mannering and Grodsky developed a multinomial logit model to determine what factors significantly influence motorcyclists' estimates of their likelihood of becoming involved in an accident if they continue to ride for 10 more years.(13) A questionnaire was used to collect data characterized by four categories: rider characteristics (e.g., age), exposure (e.g., miles driven per year), experience (e.g., years of having a motorcycle license), and behavioral attributes (e.g., a stated preference for consistently exceeding the speed limit). The questionnaire responses on the riders' estimate of their likelihood to be involved in a crash in the next 10 years were grouped into low (0 to 20 percent), medium (30 to 70 percent), and high (80 to 100 percent).

The model predicted the likelihood of a response given the characteristics represented in the model. The model is shown in figure 5.

$$P_{ni} = \frac{e^{U_{ni}}}{\sum_i e^{U_{ni}}}$$

Figure 5. Equation. Multinomial logistic response model.

Where:

P_{ni} = The probability that rider n would categorize themselves as having a low, medium, or high risk of being in an accident in the next 10 years.

U_{ni} = Linear function of variables which determine the probability of a rider considering themselves in the low-, medium-, or high-risk group.

Among the studies' findings was that age, gender, and experience were significant determinants of the estimate of self-risk and that riders were generally aware of their relative crash risks.

In a follow-up study, Shankar and Mannering presented a multinomial logit model for rider injury severity in single-vehicle crashes, considering environmental roadway and vehicle factors.[(14)] Using data for Washington, relationships were found between crash severity and motorcycle displacement, rider age, alcohol impaired riding, rider ejection, speed, rider attention, pavement surface, and type of highway.

Building on the Shankar and Mannering research, Savolainen and Mannering built multinomial logit models of severity separately for single- and multi-vehicle motorcycle crashes in Indiana and found important differences for the two crash types.[(15)] In general, they revealed several factors leading to more severe injuries, including poor visibility (horizontal curvature, vertical curvature, and darkness), unsafe speed (citations for speeding), alcohol use, not wearing a helmet, right-angle and head-on collisions, and collisions with fixed objects. Wet pavement conditions, locations near intersections, and passengers on motorcycles were associated with severe crashes, suggesting motorcyclists may be managing risks.

Jung et al. examined factors associated with motorcyclist fatalities. The research found that a lack of or improper use of helmets, victim ejection, alcohol/drug effects, collisions (i.e., head-on, broadside, and hit-object), and truck involvement were more likely to result in fatal injuries regardless of age group.[(16)] Weekend and non-peak hour activity were found to have a strong effect on both the younger and older age groups. The authors determined that two factors—movement of running off the road preceding a collision and multi-vehicle involvement—were statistically significant factors in increasing motorcycle fatalities among drivers in the older age group. Use of street lights in the dark decreased the probability of severe injury for older motorcyclists. Being the driver (as opposed to passenger), being the at-fault driver, being on a local road, and committing a speed violation were significant factors in increasing the fatalities of younger motorcyclists. Road conditions and collision location factors were not statistically significant to motorcyclist fatalities.

Jones et al. applied multinomial logistic models to analyze factors affecting the injury severity outcome of motorcycle crashes.[(17)] The variables affecting motorcycle crashes were grouped by common characteristics into four categories: motorcyclist, crash, environment, and roadway. Crashes in the vicinity of large vehicles, around roadway curves, and in rural areas increased the likelihood of severe crash outcomes.

Geedipally et al. estimated multinomial logit models to identify differences in factors likely to affect the severity of crash injuries of motorcyclists.[(18)] Key findings showed that alcohol, gender, lighting, and presence of both horizontal and vertical curves played significant roles in injury outcomes of motorcyclist crashes in urban areas. Similar factors were found to have significantly affected the injury severity of motorcyclists in rural areas, but older riders (older than 55), single-vehicle crashes, angular crashes, and divided highways also affected injury severity outcomes in rural motorcycle crashes.

Ordered Probit and Logit Models

Ordered probit and logit models are similar to logistic models. As with unordered models, probit and logit models have different assumptions for the error term distribution and link functions. These models determine which factors significantly affect the probability of the outcome of an event and can be used to predict the likelihood of each possible outcome of an event but differ in that they differentiate unequal differences between ordinal categories in the dependent variable (e.g., it does not assume that the difference between no injury and minor injury is the same difference as between a severe injury and a fatality given a unit change in an explanatory variable).

In the multiple response case, figure 6 displays the ordered model.

$$P(y=i)=\Phi(\mu_i-\beta x)-\Phi(\mu_{i-1}-\beta x)$$

Figure 6. Equation. Ordered probit model.

Where:

y = The ordinal for the outcome data.

x = A vector of variables determining the discrete ordering for each observation.

β = A vector of estimable parameters.

μ_i and μ_{i-1} = The upper bound and lower bounds for injury severity i, respectively.

The modeling process estimates both the vector or parameters and the upper and lower bound limits.

Barrette et al. studied the impacts of changes to the Michigan universal helmet use law using an ordered probit model.(19) The degree of injury severity sustained by crash-involved motorcyclists before and after Michigan's transition from a universal to a partial helmet law was examined. The models controlled for a variety of rider, roadway, traffic, and weather characteristics and indicated that helmets reduced the probability of fatalities by more than 50 percent.

Ariannezhad et al. applied ordered logit models to study the factors contributing to crash severity of motorcycle crashes on suburban roads.(20) The results indicated that there are several factors that increase the severity of motorcycle crashes. Factors include weekends, winter and fall, dawn, foggy and clear weather, non-administrative areas, riders older than 60 years old, riders without a proper license, lack of helmet, motorcycle at-fault, speeding, overtaking, collisions with buses, and heavy vehicle, pedestrian, and single-vehicle crashes. Additional factors increasing severity included head-on crashes, fatigue and sleepiness, rules violation, road imperfection, and curvature.

Wang et al. applied ordered probit models to injury severity of single-vehicle motorcycle crashes on curved roadways.(21) Results indicated that curve radius is a significant factor influencing injury severity of single-motorcycle crashes along horizontal curved roadway segments. The authors estimated that an increase of 1,000 ft in curve radius decreases the likelihood of fatalities and serious injuries by 0.2 and 0.15 percent, respectively, in single-motorcycle crashes along a curved roadway section. The authors also found that speeding and hit-object increased the likelihood of higher severities.

Blackman and Haworth applied ordered probit models to compare the crash risk and crash severity of motorcycles, mopeds, and larger scooters.(22) Greater motorcycle crash severity was associated with higher (>50 mi/h (80 km/h)) speed zones, horizontal curves, weekend, single-vehicle, and nighttime crashes. Moped crashes were more severe at night and in speed zones of 56 mi/h (90 km/h) or faster. Larger scooter crashes were more severe in 43-mi/h (70-km/h) zones than in 37-mi/h (60-km/h) zones but not in higher speed zones, and they were less severe on weekends than on weekdays.

OTHER MODELS

The literature review revealed several other analysis methods, which were applied to motorcycle safety evaluation.

Haque et al. developed log-linear models of motorcycle crash risk.(23) Log-linear models can be used to identify conditions that increase crash risk or severity and to estimate odds multipliers that express the increased or decreased risk associated with a change in a variable or interactions of variables in the model. Conventionally, contingency tables, which record the number of responses for each combination of variable values, are used. However, when the number of variables is greater than two, the process can be arduous. A log-linear model predicts the frequency of crashes for each combination of levels of explanatory variables. The frequency predicted is the number of crashes meeting the levels of each variable out of all crashes observed. This should not be confused with predicting the expected crash frequency on the roadway. For example, one category may be male riders, aged 25 to 44, with a BAC over 0.08. The frequency of crashes meeting these criteria would be predicted. The authors used quasi-induced exposure to account for exposure. In this approach, it is assumed that the presence of not-at-fault riders in the crash data represent the general population. The relative exposure for not-at-fault riders within the population of two-vehicle crashes is their crash frequency divided by the total population crash frequency. The authors then calculated the relative risk (RR) by dividing the respective odds ratios by the odds ratios for exposure under the same conditions.

Chin and Haque investigated the effects of RLCs on motorcycle crashes.(24) Quasi-induced exposure was applied using not-at-fault riders involved in right-angle crashes. Results showed that an RLC reduced the relative crash vulnerability or crash-involved exposure of motorcycles at right-angle crashes. Log-linear models were also developed and indicated that light and heavy vehicles were more likely to experience a right-angle crash with a motorcycle at non-RLC locations than intersections with RLCs.

de Rome et al. studied the effectiveness of protective clothing by recruiting motorcyclists involved in crashes through hospitals and repair services and then conducting interviews.(25) Hospitalization and injury were modeled using Poisson regression with log-link function to estimate the RR controlling for confounding variables. RR can be interpreted as the likelihood of the outcome without the variable of interest present divided by the likelihood of the outcome with the variable present. For example, the RR of an injury while wearing a helmet would be the ratio of injury crashes to non-injury crashes for riders not wearing helmets to the ratio of injury to non-injury for riders wearing helmets. The use of motorcycle jackets, pants, and gloves reduces the likelihood of hospitalization, as well as risk of upper body injury, hands/wrists, and feet/ankles.

Huang and Lai applied survival analysis using Cox regression models to identify risk factors for time until death comparing single-vehicle crashes for both alcohol- and non-alcohol-related crashes.(26) Survival analysis essentially models the probability that if one has survived until time *t*, then they will succumb to the event (in this case death) in the next instant. Cox regression models account for the effects of various covariates on the likelihood of survival and the results. The results show the impact each covariate has on the risk (in this case, the risk of death). Factors increasing risk of death for motorcycle riders included older age, crashing into trees, nighttime riding, curved roads, and local roads.

Chung applied boosted regression trees to classify single-vehicle motorcycle crashes into fatal or non-fatal crashes.(27) The output of the analysis indicated which variables contributed most to correctly classifying crashes, thus indicating which had the greatest impact on crash severity.

SUMMARY OF KEY FINDINGS FROM THE LITERATURE REVIEW

As noted earlier, the vast majority of motorcycle crash research used probability models to identify factors associated with crash severity outcomes. These models are not directly relevant to the estimation of CMFs and SPFs; however, the research was selectively reviewed because of the potential for applying probability models to crash frequency models for total crashes to estimate crash frequency by severity type.

The limited international research does suggest that motorcycle crash frequency models can be developed based only on total AADT for all vehicle types. However, the project team researched jurisdictions where motorcycle volume constitutes a sizable proportion of and is strongly correlated with the total traffic volume. The one significant study from the United States that developed crash frequency models for road segments did not use motorcycle volume as a variable due to the unavailability of these data. Instead, the number of motorcycle endorsements (motorcycle licenses) and number of registered motorcycles were used as surrogate measures of motorcycle and motor vehicle traffic, but these could only be compiled and applied at a regional level. In addition, spatial random effects modeling reduced model error due to the unavailability of motorcycle AADT data. However, the fact that this benefit is specific to the Ohio dataset modeled suggests that the developed models will not be easily transferable to another jurisdiction. Nevertheless, the research does suggest that there is promise in seeking alternatives to directly including motorcycle volumes in motorcycle crash frequency prediction models.

The literature review found very few studies focused on the prediction of crash frequency but many studies concerned with motorcycle safety. Crash frequency models are required for developing SPFs and CMFs. Table 1 provides a brief summary of the methods reviewed, including the data required, outcomes, uses, strengths, and limitations.

Table 1. Analytic methods from literature review.

Method	Data Required	Outcomes	Uses	Strengths	Limitations
Discrete choice	Crash data and presence or absence of feature of interest	Estimates the increased likelihood of dependent variable being present, given a crash has occurred, as a function of explanatory variables	Useful for identifying risk factors and comparing the RR between two or more factors	Advanced statistical methods are available; can often be accomplished using only readily available crash record data	Does not provide an estimate of crash frequency
Count frequency	Crash data, exposure data, and geometric data	Provides an estimate of expected crash frequency; CMFs can be inferred from parameter estimates	Useful for estimating CMFs and developing predictive models for before-after studies and network screening	Advanced statistical methods are available	Modeling can be difficult for rare crash types; exposure is the biggest influence of expected crashes, but motorcycle volumes are often unavailable
Log-linear/contingency table	Crash data and presence or absence of feature of interest	Estimates the increased likelihood of dependent variable being present, given a crash has occurred, as a function of explanatory variables	Useful for identifying risk factors and comparing the RR between two or more factors	Advanced statistical methods are available; can often be accomplished using only readily available crash record data	Does not provide an estimate of crash frequency
Quasi-induced exposure	Crash data	Estimates the relative presence of units (i.e. motorcyclists) in a population	Can be used as an estimate of motorcycle exposure in the absence of volume data	Requires only crash data to apply	Cannot be used as an estimate of motorcycle exposure at specific locations

Method	Data Required	Outcomes	Uses	Strengths	Limitations
RR/Poisson model	Crash data and presence or absence of feature of interest	Estimates increased risk of event without a variable of interest present compared to when it is present	Useful for identifying risk factors and comparing the RR between two or more factors	Requires only crash data to apply	Does not provide estimate of crash frequency
Survival analysis/Cox regression	Crash data and presence or absence of feature of interest	Estimates the effect a variable has on the risk of an event occurring	Useful for identifying risk factors and comparing the RR between two or more factors	Requires only crash data to apply	Does not provide estimate of crash frequency
Boosted regression trees	Crash data and presence or absence of feature of interest	Classifies crashes and indicates which variables influence this classification	Useful for identifying risk factors and determining which most affect the outcome of interest	Requires only crash data to apply	Does not provide estimate of crash frequency

CHAPTER 3. ANALYSIS METHODS

The project team developed and undertook a number of analytical approaches in order to do the following:

- Explore how much predictive power for an SPF is lost when motorcycle volumes are unknown and how this lack of information may affect a study of motorcycle countermeasures estimating a CMF.

- Explore alternative methods for deriving accurate predictions of motorcycle crashes or motorcycle volume data.

This chapter provides an overview of each analysis method and a discussion of which research goal(s) were met by its application. The project team pursued modeling of motorcycle crash data on several fronts that are helpful in casting light on the two research questions.

The methods applied may be broadly classified into two groups: avenue A and avenue B. The methods for avenue A focus on investigating the difference in predictive performance for motorcycle SPFs calibrated with motorcycle AADT versus total AADT only. The methods for avenue B focus on the difference in CMF estimates found when using motorcycle AADT versus total AADT only.

AVENUE A METHODS

The three methods applied in avenue A focus on the impact of the lack of motorcycle AADT on modeling motorcycle crashes and the development of tools for those jurisdictions lacking these data. The methods make use of statewide databases for States that have motorcycle AADT estimates available. The following section discusses the three methods, all of which involve the estimation of predictive models. Table 2 summarizes the details of these methods for quick and easy reference throughout this report.

Table 2. Summary of avenue A methods.

Model Type and Intended Function	Basic Purpose	SPFs Developed	Approach
A1: Provide a direct measure of how the predictive power of a model is affected by either including or excluding motorcycle volumes.	Explore how much predictive power is lost when motorcycle volumes are not known.	• A1.1. Motorcycle crashes versus total AADT and other independent variables. • A1.2. Motorcycle crashes versus motorcycle AADT and other independent variables.	1. Assess goodness-of-fit of two model sets and compare. 2. Assess how well each model set predicts motorcycle crashes at high crash locations. 3. Using the results from the steps above, assess predictive ability of SPF A1.1. 4. Consider SPF A1.1 for application to any jurisdiction if successful. 5. Use FHWA SPF calibration tool to assess application of SPF A1.1 to selected jurisdictions.
A2: Allow jurisdictions without motorcycle volumes to predict motorcycle crashes based on SPFs for total crashes.	Develop a relationship between predicted motorcycle crash frequency and predicted total crash frequency.	• A2.1. Motorcycle crashes versus motorcycle AADT. • A2.2. All crashes versus total AADT. • A2.3 Motorcycle crashes versus predicted total crashes and other variables.	1. Develop and assess a model that relates predictions from SPF A2.1 to predictions from SPF A2.2. 2. Consider SPF A2.3 for application to any jurisdiction if successful.
A3: Allow jurisdictions to directly estimate motorcycle volumes.	Develop models to estimate motorcycle traffic volumes based on roadway characteristics and other variables that may influence motorcycle trip generation.	A3. Motorcycle AADT versus variables related to roadway, motorcycle registrations, licensing, and sociodemographic characteristics.	1. Assess/include variables that cause motorcycle AADT to vary. 2. Consider model A3 for estimating AADT in any jurisdiction where causal variables available if successful.

FHWA = Federal Highway Administration.

All models were calibrated using generalized linear modeling (GLM) using the *R* software. GLM allows the specification of various error structures. The negative binomial error structure is recognized as an appropriate form for modeling crash data. The negative binomial overdispersion parameter (the inverse of the shape parameter provided by *R*), which is estimated in the modeling process, can be used in the comparative assessment of models fit to the same data in that a smaller value indicates a better fit model. Crash counts at sample sites are used as estimates of the dependent variable, which is the expected number of crashes per year while corresponding road characteristics and traffic data are used as estimates of the independent variables.

Model Type A1

The purpose of model type A1 is to explore how much predictive power is lost when motorcycle volumes are unknown and how this lack of information would affect an evaluation of motorcycle countermeasures. The approach is to first develop SPFs for several road classes for motorcycle crashes with and without motorcycle volumes. The performance of each SPF pair, with and without motorcycle volumes, is then evaluated across the range of motorcycle AADTs to assess the overall goodness-of-fit to the data, as well as ranges of motorcycle AADTs and any other variables to identify circumstances where the lack of motorcycle AADT may cause the model to perform poorly. Assessment measures, in addition to the overdispersion parameters, include mean absolute deviation (MAD), adjusted R^2 values, and CURE plots described in subsequent paragraphs.

MAD gives a measure of the average magnitude of variability of prediction. Smaller values are preferred to larger values. MAD is the sum of the absolute value of predicted minus observed crashes divided by the number of sites, as shown in figure 7.

$$MAD = \frac{\sum_{i=1}^{n} \left| \hat{Y}_i - Y_i \right|}{n}$$

Figure 7. Equation. MAD.

Where:

n = Validation data sample size.

Fridstrom et al. introduced a modified R^2 value.(28) This goodness-of-fit measurement subtracts the normal amount of random variation that would be expected even with a perfectly specified model. As a result, the amount of systematic variation explained by the model is measured. Larger values indicate a better fit to the data. Values greater than 1 indicate that the model is overfit, and some of the expected random variation is incorrectly explained as the systematic variation. Figure 8 shows the calculation.

$$R^2 = \frac{\sum_i (y_i - \bar{y})^2 - \sum_i \widehat{\mu_i}^2}{\sum_i (y_i - \bar{y})^2 - \sum_i \widehat{y_i}}$$

Figure 8. Equation. Modified R^2 goodness-of-fit measure.

Where:

y_i = Observed counts.
$\widehat{y}_i$ = Predicted values from the SPF.
$\bar{y}$ = Sample average.
$\widehat{\mu}_i = y_i - \widehat{y}_i$.

For an SPF to produce useful estimates, it must be good for all values of every variable. An alternative tool to describe goodness-of-fit is the CURE plot. A CURE plot is a graph of the cumulative residuals (observed minus predicted crashes) against a variable of interest sorted in ascending order. Long trends (increasing or decreasing) indicate regions of bias that should be rectified through model improvement either by the addition of new variables or by a change of functional form. Large vertical changes in the CURE plot invite the examination of outliers. The CURE plot is useful in determining whether an SPF is acceptable and in comparing multiple SPFs. The following steps are used to construct a CURE plot:

- Step 1: Sort sites in ascending order of the variable of interest so that N is the number of sites, n is an integer between 1 and N, and $S(n)$ is the cumulative sum of residuals from 1 to n.

- Step 2: For each site calculate the residuals, *res*, as the observed minus predicted crashes.

- Step 3: For each site calculate the CURE, $S(n)$, which is the sum of residuals from 1 to n.

- Step 4: For each site calculate the squared residuals, res^2.

- Step 5: For each site calculate the cumulative squared residuals, $\sigma^2(n)$, which is the sum of squared residuals from 1 to n.

- Step 6: Sum the cumulative squared residuals over all sites, $\sigma^2(N)$.

- Step 7: For each site, estimate the variance of the random walk using the equation in figure 9.

$$\sigma^2 = \sigma^2(n)[1 - \frac{\sigma^2(n)}{\sigma^2(N)}]$$

Figure 9. Equation. CURE plot variance estimate.

- Step 8: For each site, calculate the 95-percent confidence limits using the equations in figure 10 and figure 11.

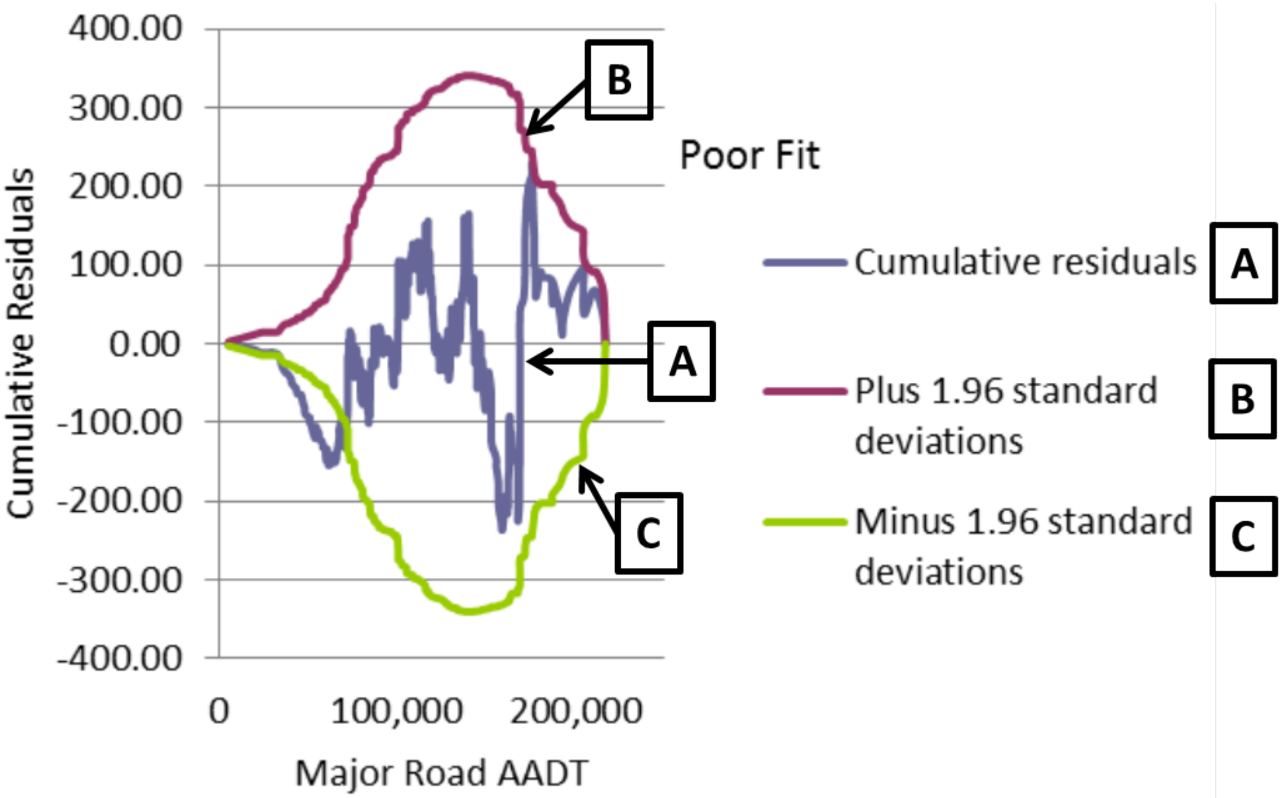
$$Lower\ Limit = -1.96\sqrt{\sigma^2}$$

Figure 10. Equation. Lower limit of 95-percent confidence interval.

$$Upper\ Limit = +1.96\sqrt{\sigma^2}$$

Figure 11. Equation. Upper limit of 95-percent confidence interval.

- Step 9: Plot the CURE $S(n)$ and the 95-percent confidence limits on the y-axis against the explanatory variable of interest on the x-axis. Figure 12 displays an example CURE plot for the variable major road AADT. In this example, the model is performing relatively well, as the CURE remain within the 95-percent confidence limits over most of the range, only crossing outside the boundary limits for a short range of lower AADT. The areas outside the boundary limits indicate a poor fit as indicated in the figure. If the graph of CURE was in that area frequently, then this would indicate the presence of model bias.



©Persaud and Lyon.

Figure 12. Graph. Example CURE plot.

The project team used two comparative measures to assess the CURE plots for competing models. One is the maximum CURE deviation, which has a value of 237 at an AADT of 159,000 in the figure 1 example. The other is percent CURE deviation, which is defined as the percent of the range of the x-axis variable for which the CURE plot is outside the 95-percent confidence limits. In figure 12, this occurs between AADTs of approximately 35,000 and 70,000; about 16 percent of the individual data points lie outside the 95 percent confidence limits.

It is also of particular interest to assess how well the models predict motorcycle crashes for high-crash locations, as these sites are the ones typically of interest in treatment applications that would form the basis for future CMF development. To do this, sites are ranked first by the crash counts per mile in one period; the Empirical Bayes (EB) estimates based on the calibrated SPFs and crash counts in that period for the highest ranked locations are then compared to crash counts for these locations in a subsequent period.

Model A1 was applied because it provides a direct measure of how the predictive power of a model is affected by either including or excluding motorcycle volumes. By using the same data for both sets of models, the impact of motorcycle volumes on predictive power was assessed directly without any biases. The use of CURE plots also allows this assessment to be broken down by ranges of motorcycle volumes.

Model Type A2

The purpose of model type A2 is to develop a relationship between motorcycle crash frequency and total crash frequency as a function of traffic volumes for the two vehicle categories. Models were calibrated for both motorcycle and total crashes using traffic volumes for motorcycles and all vehicles, respectively. Then, a relationship between the SPFs for the two crash types can be inferred from the motorcycle SPF prediction using the total crash SPF prediction as an explanatory variable. If successful, this relationship could then be applied to the SPF for total crashes for another State to infer an SPF for motorcycle crashes for that State. In turn, that SPF can be used in the evaluation of retrospective and prospective before-after evaluations of the effects on motorcycle crashes of infrastructure countermeasures. The assessment of success uses similar measures (MAD, etc.) as for assessing model type A1.

Model A2 was applied because, if successful, it would allow jurisdictions without motorcycle volumes to predict motorcycle crashes based on SPFs for total crashes. The limitation of the approach is that the models may not transfer well between jurisdictions that differ in terms of factors influencing motorcycle VMT and riding patterns.

Model Type A3

The purpose of investigating model type A3 was to attempt the development of models to estimate motorcycle traffic volumes based on roadway characteristics and other variables that may influence motorcycle trip generation. For the latter, information on motorcycle registrations, licensing, and sociodemographic variables was collected. If successful, these models could be used to estimate motorcycle volumes in similar jurisdictions. Due to differences in weather, motorcycle riding culture, commuter versus recreational riding, and other factors affecting motorcycle volumes between jurisdictions, it will be important to identify which of these factors need consideration when determining the applicability of a model. The assessment of success used similar measures as for assessing model type A1.

The project team applied this approach because, if successful, it would allow jurisdictions to directly estimate motorcycle volumes. If the important factors influencing motorcycle VMT can be identified and included in the models developed, this approach should provide for better model transferability than model type A2.

AVENUE B METHODS

The methods applied in avenue B focus on the impact of the lack of motorcycle AADT on the estimation of CMFs. These methods make use of simulated data. Simulating data creates a database with many locations and with assumed relationships between roadway geometry or other countermeasures and motorcycle crashes. The ability to accurately measure this "true" relationship is then tested when motorcycle volumes are and are not used in the process. The fixed relationships affecting motorcycle crashes were determined considering a likely range of values based on existing safety knowledge.

To investigate the impact of the lack of motorcycle AADTs on the estimation of CMFs, two CMF estimation approaches were investigated: B1, the EB before-after approach, and B2, cross-sectional generalized linear models.

For the EB before-after approach, one or more countermeasures were assumed with a known value of its CMF. The simulated database was divided into two time periods, and the after period expected crash means for each location is adjusted by the value of the CMF. The new after period counts are then generated from the Poisson distribution. The EB approach was then applied to the data for these treated sites, using the remaining sites as a reference group. This was done once using the motorcycle AADTs and once for total AADT. A comparison was then made to see how the lack of motorcycle AADT affected the estimate of the CMF and its variance. This entire process, beginning with the simulated data, was performed multiple times and with multiple sample sizes so that conclusions could be made with confidence and have broad applicability.

For the cross-sectional regression model approach, an assumed relationship, based on logical considerations and related research, was defined between one or more geometric variables and added to the SPFs developed in model A1. The relationship was defined in terms of a CMF. This modified SPF was then used to simulate the data as described above. GLM was then used to re-estimate the SPF, including the fictional variable, using motorcycle AADT and then using total AADT only. A comparison was then made to see how the lack of motorcycle AADT affected the estimate of the CMF and its variance.

This entire process, beginning with the simulated data, was performed multiple times so that conclusions could be made with confidence.

This approach for avenue B was applied because it provided a direct measure of how the lack of motorcycle AADT affects CMF estimation by replicating the process of estimating CMFs. The use of simulated data provides a realistic and unbiased dataset for making this assessment. The limitation is that the true relationships between motorcycle AADT and crashes and geometric features and crashes need to be assumed. However, the knowledge gained in avenue A using real data informed these decisions.

CHAPTER 4. DATA COLLECTION AND SUMMARY

This chapter describes the data collection procedures undertaken to support the analyses outlined in chapter 3. In general terms, these data are primarily roadway inventory, crash data, total traffic volumes, and motorcycle traffic volumes. Other data collection includes further information on motorcycle licenses, registrations, and other sociodemographic data at the county level.

To identify which States could provide a substantial sample of motorcycle volume counts, the project team asked various contacts for this information, primarily through the following channels:

- A request to the Transportation Research Board committee on motorcycles and mopeds, ANF30.
- Listserv communication with State highway safety engineers.
- Requests to members of the FHWA Development of CMFs Project Technical Advisory Committee.

These inquiries quickly confirmed that few States have substantial motorcycle volume estimates. For many States, the only available motorcycle counts come from permanent counting stations, where full class counts are performed and the interest is focused on truck traffic, not motorcycles. No attempt is made to estimate motorcycle counts for nearby segments in these instances.

There were, however, a number of States reported by these contacts as having a large number of classification count locations from permanent and short-term count programs. These States focused on for determining which States could provide the best datasets for the analyses.

The prioritization of States for data collection considered several factors, including the following:

- Availability of electronically linked and stored data on the roadway geometry, traffic volumes (including motorcycle volumes), and crash data for 5 years.
- For crash data, the necessary ability to identify motorcycle-involved crashes.
- For traffic data, classification counts, including motorcycles (a distinct class) and the total count for each site.
- Availability of data on variables useful for predicting motorcycle AADT, such as motorcycle licensure and socio-demographics.
- The ability to provide seasonal estimates of motorcycle volumes. These are known to fluctuate significantly, and it would be desirable to reflect this in the models developed.
- Geographic diversification by including States from northern and southern regions.

For each method applied, this section describes the source of data, variables acquired, and steps taken to assemble the data into an appropriate format for analysis.

AVENUE A DATABASES

For developing the avenue A models, data were collected from Florida and Pennsylvania. The project team selected these States because they had a large number of locations with an estimated motorcycle AADT; were able to provide linkable roadway inventory, traffic, and crash data; and expressed an interest in providing data in a timely manner. They also provided a degree of geographical diversity. The project team acquired data from Virginia to validate the models developed.

The following sections provide further details on the data acquired and data manipulation.

Florida

Florida provided statewide roadway inventory, traffic volume, and crash data. The Florida Department of Transportation (FDOT) provided the roadway inventory and traffic volume data on the 2012 Florida Transportation Information DVD. Roadway and traffic volume data variables included the following:

- Segment location and length.
- AADT estimates for current and past years.
- Motorcycle AADT estimates for current and past years for those segments where an estimate exists.
- Number of lanes.
- Posted speed.
- Surface width.
- Functional class.
- Divided versus undivided.
- Inside and outside shoulder widths.
- Median width.
- Median type.
- Degree of curvature.

Crash data were provided for 2008 to 2012, including the following:

- Location.
- Date.
- Severity.
- First and second harmful events.
- Location type.
- Crash type.
- Vehicle types involved.
- Number of vehicles involved.

The query included total, motorcycle, multi-vehicle motorcycle, and single-vehicle motorcycle crashes.

Roadway identification numbers and mileposts linked all roadway, traffic, and crash data. Segments were defined so that all variables were homogeneous for the entire length.

Further data collected included the number of motorcycle licenses and registrations by county from 2008 to 2012. These data were available from the Florida Department of Highway Safety and Motor Vehicles and were linked to the other data by county.

Sociodemographic data were obtained from the U.S. Census Bureau at the county level. These variables included the following:

- Population by age.
- Population by sex.
- Percentage of population with high school or college degree.
- Mean travel time to work.
- Housing units.
- Home ownership rate.
- Persons per household.
- Median household income.
- Percentage of persons below poverty level.
- Non-employer establishments.
- Number of firms.
- Land area in square miles.
- Population per square mile.

These data were also linked to the other data by county. Segments were initially grouped into the following six categories for analysis:

- Rural freeway (type 1).
- Urban freeway (type 2).
- Rural arterial (type 3).
- Urban arterial (type 4).

- Rural collector/local (type 5).
- Urban collector/local (type 6).

Table 3 provides the variable definitions of the database created. Table 4 shows the total length of roadway and total number of crashes by crash type for the six site types in Florida. Table 5 through table 8 provide summary statistics for other variables included in the Florida dataset by site type.

Table 3. Florida data variable definitions.

Variable	Definition
AVGMOTO	Motorcycle AADT
AVGAADT	Total vehicle AADT
NOLANES	Number of lanes
SPDLIMT	Posted speed limit (mi/h)
SURFWIDTH	Surface width in ft
DIVUND	Divided versus undivided
OUTSHLDWID	Outside shoulder width in ft
MEDWIDTH	Median width in ft
INSHLDWID	Inside shoulder width in ft
CURVLENGTH	Length of horizontal curves in segment in mi
LENGTH	Total segment length in mi
TOT	Total crash frequency
MOTO	Motorcycle crash frequency
MOTOSINGLE	Single-vehicle motorcycle crash frequency
MOTOMULTI	Multi-vehicle motorcycle crash frequency
VEHREG	Number of registered motorcycles in county
LIC	Number of motorcycle-licensed individuals in county
PST045213	Population, 2013 estimate
AGE135213	Persons under 5 years, percent, 2013
AGE295213	Persons under 18 years, percent, 2013
AGE775213	Persons 65 years and over, percent, 2013
SEX255213	Female persons, percent, 2013
EDU635213	High school graduate or higher, percent of persons age 25+, 2009–2013
EDU685213	Bachelor’s degree or higher, percent of persons age 25+, 2009–2013
LFE305213	Mean travel time to work (minutes), workers age 16+, 2009–2013
HSG010213	Housing units, 2013
HSG445213	Homeownership rate, 2009–2013
HSD310213	Persons per household, 2009–2013
INC110213	Median household income, 2009–2013
PVY020213	Persons below poverty level, percent, 2009–2013
NES010212	Non-employer establishments, 2012
SBO001207	Total number of firms, 2007
LND110210	Land area in mi^2, 2010
POP060210	Population per mi^2, 2010

Table 4. Total length and crash frequency for Florida data.

Type	LENGTH (mi)	TOT	MOTO	MOTOSINGLE	MOTOMULTI
1	494.2	10,611	174	111	63
2	552.6	49,453	921	385	536
3	3,068.2	21,400	1,031	449	582
4	3,425.8	269,689	9,539	2,341	7,198
5	640.8	698	33	20	13
6	1,041.2	2,800	152	57	95

1 mi = 1.6 km.

Table 5. Summary statistics for Florida data.

Type	Statistic	AVGMOTO	AVGAADT (mi)	NOLANES	SPDLIMT (mi/h)	SURFWIDTH (ft)	DIVUND	OUTSHLDWID (ft)
1	No. Segments	482	482	482	482	482	482	482
1	MIN	18.9	16,050.0	4.0	65.0	48.0	2.0	6.0
1	MAX	339.9	98,700.0	8.0	70.0	96.0	2.0	21.0
1	MEAN	105.6	34,845.0	4.5	69.9	53.4	2.0	10.0
1	STD	72.2	19,303.0	0.9	0.8	10.3	0.0	1.2
2	No. Segments	952	952	952	952	952	952	952
2	MIN	18.9	6,120.0	2.0	30.0	24.0	0.0	2.0
2	MAX	39,488.9	316,000.0	10.0	70.0	128.0	2.0	35.0
2	MEAN	862.1	84,174.3	5.7	64.8	68.5	2.0	10.0
2	STD	3,795.4	56,359.9	1.6	6.8	20.0	0.1	3.0
3	No. Segments	3,243	3,243	3,243	3,243	3,243	3,243	3,243
3	MIN	1.4	450.0	2.0	25.0	18.0	0.0	2.0
3	MAX	507.8	43,000.0	6.0	70.0	72.0	2.0	26.0
3	MEAN	51.0	8,044.0	2.6	54.4	30.7	0.9	5.3
3	STD	52.5	6,820.3	0.9	7.6	11.0	1.0	2.2
4	No. Segments	8,721	8,721	8,721	8,721	8,721	8,721	8,721
4	MIN	0.6	170.0	2.0	20.0	18.0	0.0	1.0
4	MAX	2,416.4	92,785.0	10.0	65.0	120.0	2.0	32.0
4	MEAN	171.5	26,198.8	4.0	44.4	47.4	1.7	4.6
4	STD	141.2	15,797.2	1.6	6.9	18.3	0.7	2.9
5	No. Segments	622	622	622	622	622	622	622
5	MIN	1.4	100.0	1.0	25.0	12.0	0.0	1.0
5	MAX	274.4	31,000.0	4.0	60.0	52.0	2.0	12.0
5	MEAN	32.3	4,416.6	2.1	47.5	23.7	0.3	6.1
5	STD	37.9	4,211.4	0.3	8.6	4.6	0.7	3.1
6	No. Segments	2,345	2,345	2,345	2,345	2,345	2,345	2,345
6	MIN	1.9	170.0	2.0	15.0	14.0	0.0	1.0
6	MAX	362.1	55,000.0	6.0	55.0	78.0	2.0	25.0
6	MEAN	63.8	8,498.3	2.4	37.9	27.3	0.8	6.3
6	STD	51.1	6,679.8	0.9	7.6	10.1	1.0	3.9

STD = Standard deviation.

Table 6. Summary statistics for Florida data continued—roadway geometry.

Type	Statistic	MEDWIDTH (ft)	INSHLDWID (ft)	CURVLENGTH (mi)	LENGTH (mi)	TOT	MOTO	MOTOSINGLE	MOTOMULTI
1	No. Segments	482	482	482	482	482	482	482	482
1	MIN	3.0	3.0	0.0	0.1	0.0	0.0	0.0	0.0
1	MAX	961.0	24.0	1.1	23.4	232.0	5.0	4.0	3.0
1	MEAN	147.2	6.5	0.1	1.0	22.0	0.4	0.2	0.13
1	STD	159.2	3.6	0.2	1.8	31.4	0.8	0.6	0.39
2	No. Segments	952	932	952	952	952	952	952	952
2	MIN	0.0	2.0	0.0	0.1	0.0	0.0	0.0	0.0
2	MAX	989.0	51.0	2.0	7.8	707.0	13.0	5.0	9.0
2	MEAN	84.8	9.3	0.1	0.6	51.9	1.0	0.4	0.56
2	STD	98.3	4.3	0.1	0.8	76.2	1.6	0.8	1.08
3	No. Segments	3,243	496	3,243	3,243	3,243	3,243	3,243	3,243
3	MIN	0.0	1.0	0.0	0.1	0.0	0.0	0.0	0.0
3	MAX	480.0	15.0	1.2	16.4	155.0	11.0	7.0	7.0
3	MEAN	13.8	2.9	0.0	0.9	6.6	0.3	0.1	0.18
3	STD	21.9	1.6	0.1	1.6	11.4	0.8	0.5	0.54
4	No. Segments	8,721	3,459	8,721	8,721	8,721	8,721	8,721	8,721
4	MIN	0.0	1.0	0.0	0.1	0.0	0.0	0.0	0.0
4	MAX	377.0	27.0	2.0	6.1	1,578.0	77.0	16.0	69.0
4	MEAN	20.8	2.3	0.0	0.4	30.9	1.1	0.3	0.83
4	STD	18.5	1.4	0.1	0.4	60.2	2.3	0.7	1.91
5	No. Segments	622	16	622	622	622	622	622	622
5	MIN	0.0	2.0	0.0	0.1	0.0	0.0	0.0	0.0
5	MAX	170.0	7.0	5.3	14.6	25.0	4.0	1.0	3.0
5	MEAN	2.5	2.4	0.0	1.0	1.1	0.1	0.0	0.02
5	STD	10.1	1.3	0.2	1.8	2.9	0.3	0.2	0.17
6	No. Segments	2,345	222	2,345	2,345	2,345	2,345	2,345	2,345
6	MIN	0.0	1.0	0.0	0.1	0.0	0.0	0.0	0.0
6	MAX	154.0	25.0	0.5	5.4	95.0	5.0	3.0	3.0
6	MEAN	7.0	2.3	0.0	0.4	1.2	0.1	0.0	0.0
6	STD	12.5	2.1	0.0	0.5	5.2	0.3	0.2	0.2

Table 7. Summary statistics for Florida data continued—model estimates for variables VEHREG, LIC, and *a* through *i*.

Type	Statistic	VEHREG	LIC	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*
1	No. Segments	482	482	482	482	482	482	482	482	482	482	482
1	MIN	0.03	0.06	14,194.00	2.10	7.70	10.90	42.10	75.10	9.40	19.90	5,663.00
1	MAX	4.00	6.94	1,838,844.00	6.80	25.10	51.60	52.50	91.90	44.20	30.00	812,565.00
1	MEAN	0.77	1.45	232,353.61	5.22	19.57	21.67	49.71	84.83	21.65	25.57	115,870.27
1	STD	0.84	1.51	272,384.41	1.06	3.39	8.60	2.03	4.86	7.02	2.56	125,062.91
2	No. Segments	952	952	952	952	952	952	952	952	952	952	952
2	MIN	0.11	0.22	48,922.00	3.20	13.30	10.50	45.00	78.00	13.30	19.90	20,715.00
2	MAX	4.21	6.94	2,617,176.00	6.80	25.40	37.00	52.50	91.90	44.20	30.60	993,993.00
2	MEAN	2.43	4.26	1,034,493.61	5.62	20.89	17.54	51.27	86.72	27.49	25.63	450,529.30
2	STD	1.23	1.97	736047.64	0.76	2.23.00	5.84	0.70	3.27	5.66	2.40	286,103.34
3	No. Segments	3,243	3,243	3,243	3,243	3,243	3,243	3,243	3,243	3,243	3,243	3,243
3	MIN	0.01	0.03	8,349.00	2.10	7.70	10.50	35.30	64.20	7.80	18.80	3,282.00
3	MAX	4.21	6.94	2,617,176.00	7.90	28.30	51.60	52.50	93.20	44.20	32.10	993,993.00
3	MEAN	0.81	1.47	272,039.92	5.28	19.96	20.86	48.93	82.95	19.81	25.58	126,694.96
3	STD	0.94	1.61	426,048.86	0.94	3.11	7.30	3.30	6.67	8.10	2.81	176,082.45
4	No. Segments	8,721	8,721	8,721	8,721	8,721	8,721	8,721	8,721	8,721	8,721	8,721
4	MIN	0.04	0.09	22,857.00	2.10	7.70	10.50	43.60	64.20	8.80	18.80	9,516.00
4	MAX	4.21	6.94	2,617,176.00	7.90	28.30	51.60	52.50	93.20	44.20	32.10	993,993.00
4	MEAN	2.13	3.77	847,443.97	5.42	20.40	19.87	51.14	86.43	25.92	25.47	382,889.66
4	STD	1.25	2.06	694,603.48	0.81	2.55	6.35	1.07	4.09	6.24	2.53	281,445.36
5	No. Segments	622	622	622	622	622	622	622	622	622	622	622
5	MIN	0.01	0.03	8,349.00	2.10	7.70	10.90	35.30	71.30	7.80	18.80	3,298.00
5	MAX	4.21	6.17	2,617,176.00	6.70	25.40	51.60	52.50	91.80	44.20	32.10	993,993.00
5	MEAN	0.92	1.74	329,206.54	5.14	19.63	20.78	48.63	83.03	19.41	26.18	156,257.94
5	STD	0.99	1.80	426,773.89	0.83	2.97	7.78	3.64	5.12	7.38	2.79	191,564.69
6	No. Segments	2,345	2,345	2,345	2,345	2,345	2,345	2,345	2,345	2,345	2,345	2,345
6	MIN	0.07	0.14	26,850.00	2.10	7.70	10.50	43.70	64.40	9.20	18.80	10,837.00
6	MAX	4.21	6.94	2,617,176.00	7.90	28.30	51.60	52.50	91.90	44.20	30.60	993,993.00

Type	Statistic	VEHREG	LIC	*a*	*b*	*c*	*d*	*e*	*f*	*g*	*h*	*i*
6	MEAN	1.76	3.23	620,362.09	5.14	19.73	23.25	51.23	86.51	24.22	25.32	295,750.35
6	STD	1.03	1.73	504,284.65	0.83	2.71	6.64	0.89	3.15	5.76	2.40	216,888.03

a = PST045213.
b = AGE135213.
c = AGE295213.
d = AGE775213.
e = SEX255213.
f = EDU635213.
g = EDU685213.
h = LFE305213.
i = HSG010213.

Table 8. Summary statistics for Florida data continued—model estimates for variables *J* through *r*.

Type	Statistic	*J*	*k*	*l*	*m*	*N*	*o*	*p*	*q*	*r*
1	No. Segments	482	482	482	482	482	482	482	482	482
1	MIN	5,663.00	53.80	4,657.00	33,833.00	12.00	527.00	816.00	472.54	24.70
1	MAX	812,565.00	90.20	663,458.00	57,703.00	26.50	215,377.00	237,524.00	1,998.32	1,444.90
1	MEAN	115,870.27	73.38	88,345.35	44,336.88	17.38	179,14.67	21,492.63	838.31	260.78
1	STD	125,062.91	5.89	100,444.19	6,209.94	4.41	25,988.27	29,450.56	377.32	258.27
2	No. Segments	952	952	952	952	952	952	952	952	952
2	MIN	20,715.00	53.80	16,244.00	36,809.00	11.30	2,642.00	3,106.00	273.80	54.20
2	MAX	993,993.00	80.20	828,031.00	58,175.00	23.20	385,593.00	40,3672.00	1,998.32	3,347.50
2	MEAN	450,529.30	65.13	369,088.93	47,390.83	16.60	107,506.72	119,041.09	1,017.32	1,049.15
2	STD	286,103.34	6.99	240,329.25	3,870.59	2.60	109,657.00	114,337.80	486.79	764.17
3	No. Segments	3,243	3,243	3,243	3,243	3,243	3,243	3,243	3,243	3,243
3	MIN	3,282.00	53.80	2,305.00	32,497.00	9.60	318.00	247.00	243.56	10.00
3	MAX	993,993.00	90.20	828,031.00	64,876.00	29.60	385,593.00	403,672.00	1,998.32	1,444.90
3	MEAN	126,694.96	72.58	98,822.28	43,205.24	18.48	24,274.26	27,818.92	905.12	255.82
3	STD	176,082.45	6.72	143,902.59	6,734.04	4.79	55,716.44	58,978.97	419.83	289.70
4	No. Segments	8,721	8,721	8,721	8,721	8,721	8,721	8,721	8,721	8,721
4	MIN	9,516.00	53.80	7,463.00	32,497.00	9.60	976.00	1,363.00	273.80	21.60
4	MAX	993,993.00	90.20	828,031.00	64,876.00	29.60	385,593.00	403,672.00	1,998.32	3,347.50
4	MEAN	382,889.66	68.15	308,052.76	47,074.35	16.14	85,823.48	96,295.40	1,050.18	870.48
4	STD	281,445.36	6.85	234,944.41	4,691.20	3.01	97,161.36	102,438.99	515.71	789.18
5	No. Segments	622	622	622	622	622	622	622	622	622
5	MIN	3,298.00	53.80	2,305.00	32,780.00	9.80	411.00	354.00	243.56	10.00
5	MAX	993,993.00	90.20	828,031.00	59,482.00	29.60	385,593.00	403,672.00	1,998.32	1,315.50
5	MEAN	156,257.94	73.17	120,845.27	42,810.66	18.18	27,168.37	31,470.86	894.82	311.57
5	STD	191,564.69	6.68	152,005.26	6,072.52	3.98	48,519.98	52,539.47	441.17	316.27
6	No. Segments	2,345	2,345	2,345	2,345	2,345	2,345	2,345	2,345	2,345
6	MIN	10,837.00	53.80	8,857.00	34,963.00	11.30	1,229.00	1,899.00	273.80	34.00
6	MAX	993,993.00	90.20	828,031.00	58,175.00	29.60	385,593.00	403,672.00	1,998.32	3,347.50

Type	Statistic	*J*	*k*	*l*	*m*	*N*	*o*	*p*	*q*	*r*
6	MEAN	295,750.35	71.55	231,854.99	46,281.02	15.71	55,193.08	63,522.63	1,089.80	609.38
6	STD	216,888.03	5.71	178,603.72	4,641.58	2.49	63,611.04	68,193.67	529.46	578.72

j = HSG010213.
k = HSG445213.
l = HSD410213.
m = INC110213.
n = PVY020213.
o = NES010212.
p = SBO001207.
q = LND110210.
r = POP060210.

Pennsylvania

Pennsylvania provided statewide roadway inventory, traffic volume, and crash data. Roadway and traffic volume data variables included the following:

- Segment location and length.
- AADT estimates for current and past years.
- Motorcycle AADT estimates for current and past years for those segments where an estimate exists.
- Divided versus undivided.
- Surface width.
- Lane width.
- Posted speed.
- Number of lanes.
- Right shoulder type.
- Right shoulder width.
- Left shoulder type.
- Left shoulder width.

Crash data were provided for 2009–2013, including the following:

- Location.
- Date.
- Severity.
- Location type.
- Crash type.
- Vehicle types involved.
- Number of vehicles involved.

The query included total, motorcycle, multi-vehicle motorcycle, and single-vehicle motorcycle crashes. County, route number, segment number, and offset linked all roadway, traffic, and crash data. Segments were defined so that all variables were homogeneous for the entire length. The data for freeways are for one direction of travel only.

Further data collected included the number of motorcycle licenses and registrations by county from 2009–2013. These data were available from the Department of Motor Vehicles and linked to the other data by county.

The project team obtained sociodemographic data from the U.S. Census Bureau at the county level. These variables included the following:

- Population by age.
- Population by sex.
- Percent of population with high school or college degree.
- Mean travel time to work.
- Housing units.

- Home ownership rate.
- Persons per household.
- Median household income.
- Percent of persons below poverty level.
- Non-employer establishments.
- Number of firms.
- Lane area in mi^2.
- Population per mi^2.

These data were also linked to the other data by county. Segments were initially grouped into the following four categories for analysis:

- Rural freeway (type 1).
- Urban freeway (type 2).
- Rural non-freeway (type 3).
- Urban non-freeway (type 4).

Table 9 provides the variables' definitions. Table 10 shows the total length of roadway and total number of crashes by crash type for the six site types in Pennsylvania. Table 11 to table 14 provide summary statistics for other variables included in the Pennsylvania dataset by site type.

Table 9. Pennsylvania data variable definitions.

Variable	Definition
AVGMOTO	Motorcycle AADT
AVGAADT	Total vehicle AADT
NOLANES	Number of lanes
SPDLIMT	Posted speed limit (mi/h)
SURFWIDTH	Surface width in ft
DIVUND	Divided versus undivided
LSHLDWID	Left side shoulder width in ft
MEDWIDTH	Median width in ft
RSHLDWID	Right side shoulder width in ft
WIDTH	Total lane widths in ft
LENGTH	Total segment length in mi
TOT	Total crash frequency
MOTO	Motorcycle crash frequency
MOTOSINGLE	Single-vehicle motorcycle crash frequency
MOTOMULTI	Multi-vehicle motorcycle crash frequency
VEHREG	Number of registered motorcycles in county
LIC	Number of motorcycle licenced individuals in county
PST045213	Population, 2013 estimate
AGE135213	Persons under 5 years, percent, 2013
AGE295213	Persons under 18 years, percent, 2013
AGE775213	Persons 65 years and over, percent, 2013
SEX255213	Female persons, percent, 2013
EDU635213	High school graduate or higher, percent of persons age 25+, 2009–2013
EDU685213	Bachelor’s degree or higher, percent of persons age 25+, 2009–2013
LFE305213	Mean travel time to work (minutes), workers age 16+, 2009–2013
HSG010213	Housing units, 2013
HSG445213	Homeownership rate, 2009–2013
HSD310213	Persons per household, 2009–2013
INC110213	Median household income, 2009–2013
PVY020213	Persons below poverty level, percent, 2009–2013
NES010212	Non-employer establishments, 2012
SBO001207	Total number of firms, 2007
LND110210	Land area in mi^2, 2010
POP060210	Population per mi^2, 2010

Table 10. Summary of length and crash frequency for Pennsylvania data.

Type	LENGTH (mi)	TOT	MOTO	MOTOSINGLE	MOTOMULTI
1	2,136	15,603	234	184	50
2	1,666	31,853	615	382	233
3	20,559	82,666	3,749	2,432	1,317
4	6,252	125,642	3,893	1,530	2,363

Note: Type 1 and type 2 data are for one direction of travel only.
1 mi = 1.6 km.

Table 11. Summary statistics for Pennsylvania data.

Type	Statistic	SURFWID (ft)	WID (ft)	SPDLIMT (mi/h)	NOLANES	LSHLDWID (ft)	RSHLDWID (ft)	AVGAADT (mi)	AVGMOTO (mi)	DIVUND	LENGTH (mi)
1	No. Segments	4,350	4,350	4,350	4,350	4,350	4,350	4,350	4,350	4,350	4,350
1	MIN	52.0	12.0	35.0	1.0	0.0	0.0	390.7	0.1	1.0	0.1
1	MAX	98.0	52.0	70.0	4.0	15.0	14.0	27,673.3	735.0	1.0	1.1
1	MEAN	64.4	24.5	63.5	2.0	7.0	6.8	4,515.0	72.0	1.0	0.5
1	STD	5.2	2.7	3.8	0.1	3.1	3.1	4,010.6	81.6	0.0	0.1
2	No. Segments	3,459	3,459	3,459	3,459	3,459	3,459	3,459	3,459	3,459	3,459
2	MIN	52.0	12.0	35.0	1.0	0.0	0.0	347.3	1.1	1.0	0.1
2	MAX	98.0	72.0	65.0	5.0	24.0	20.0	118,193.2	817.0	1.0	1.2
2	MEAN	65.6	26.6	57.0	2.1	6.4	6.7	10,524.0	125.3	1.0	0.5
2	STD	6.8	6.8	6.0	0.4	3.5	3.5	9,871.5	127.4	0.0	0.1
3	No. Segments	43,914	43,914	43,914	43,914	43,914	43,914	4,3908	43,914	43,914	43,914
3	MIN	20.0	8.0	0.0	1.0	0.0	0.0	14.8	0.2	0.0	0.1
3	MAX	99.0	70.0	55.0	3.0	14.0	14.0	27,958.8	727.0	0.0	1.3
3	MEAN	57.2	20.9	45.2	2.0	1.5	1.5	1,753.0	22.1	0.0	0.5
3	STD	7.1	3.4	8.2	0.1	2.1	2.2	2,345.1	36.2	0.0	0.1
4	No. Segments	14,438	14,438	14,438	14,438	14,438	14,438	14,438	14,438	14,438	14,438
4	MIN	40.0	10.0	0.0	1.0	0.0	0.0	71.2	0.8	0.0	0.1
4	MAX	99.0	80.0	60.0	5.0	15.0	22.0	36,010.6	2,078.0	0.0	0.8
4	MEAN	59.0	26.0	37.8	2.0	2.1	2.2	6,599.0	64.8	0.0	0.4
4	STD	8.1	7.7	7.6	0.2	2.6	2.7	4,969.9	71.5	0.0	0.2

Table 12. Summary statistics for Pennsylvania data continued—roadway geometry.

Type	Statistic	TOT	MOTO	MOTOSINGLE	MOTOMULTI
1	No. Segments	4,350	4,350	4,350	4,350
1	MIN	0.0	0.0	0.0	0.0
1	MAX	34.0	3.0	2.0	2.0
1	MEAN	3.6	0.1	0.0	0.0
1	STD	3.0	0.2	0.2	0.1
2	No. Segments	3,459	3,459	3,459	3,459
2	MIN	0.0	0.0	0.0	0.0
2	MAX	125.0	6.0	6.0	3.0
2	MEAN	9.2	0.2	0.1	0.1
2	STD	12.2	0.5	0.4	0.3
3	No. Segments	43,914	43,914	43,914	43,914
3	MIN	0.0	0.0	0.0	0.0
3	MAX	83.0	9.0	9.0	3.0
3	MEAN	1.9	0.1	0.1	0.0
3	STD	3.1	0.3	0.3	0.2
4	No. Segments	14,438	14,438	14,438	14,438
4	MIN	0.0	0.0	0.0	0.0
4	MAX	129.0	9.0	7.0	6.0
4	MEAN	8.7	0.3	0.1	0.2
4	STD	10.8	0.6	0.4	0.5

Table 13. Summary statistics for Pennsylvania data continued—model estimates for variables VEHREG, LIC, and *A* through *H*.

Type	Statistic	VEHREG	LIC	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
1	No. Segments	5,817	5,817	5,817	5,817	5,817	5,817	5,817	5,817	5,817	5,817
1	MIN	574.4	1,388.6	14,670.0	3.9	15.5	12.3	44.7	81.5	11.5	18.6
1	MAX	27,476.4	61,992.6	1,231,527.0	6.6	24.3	20.2	52.1	93.3	40.4	41.5
1	MEAN	5,287.3	10,998.6	135,337.1	5.2	20.5	17.9	50.4	88.1	20.5	23.9
1	STD	3,602.3	7,474.1	105,230.9	0.6	1.6	1.8	1.1	2.4	6.3	4.4
2	No. Segments	4,340	4,340	4,340	4,340	4,340	4,340	4,340	4,340	4,340	4,340
2	MIN	16,55.6	3,090.8	37,838.0	4.1	15.5	12.3	44.7	81.2	11.5	18.6
2	MAX	27,476.4	61,992.6	1,553,165.0	7.0	24.3	20.2	52.7	93.5	48.5	38.9
2	MEAN	11,897.6	25,884.1	459,256.6	5.4	20.9	16.7	51.0	89.5	28.3	24.5
2	STD	7,327.3	16,611.3	415,431.1	0.6	1.8	2.0	0.9	3.4	9.2	3.6
3	No. Segments	58,095	58,095	58,095	58,095	58,095	58,095	58,095	58,095	58,095	58,095
3	MIN	306.6	655.6	4,886.0	1.7	8.8	12.3	33.1	81.5	7.8	15.2
3	MAX	27,476.4	61,992.6	1,231,527.0	6.6	24.3	26.5	52.1	93.5	48.5	41.5
3	MEAN	5,525.3	11,481.7	141,715.9	5.1	20.4	18.2	50.0	87.8	19.6	24.5
3	STD	5,092.4	10,551.0	152,302.6	0.6	2.0	2.0	1.9	2.4	6.6	4.0
4	No. Segments	19,074	19,074	19,074	19,074	19,074	19,074	19,074	19,074	19,074	19,074
4	MIN	731.0	1,633.0	18,541.0	4.0	16.0	12.0	45.0	81.0	12.0	17.0
4	MAX	27,476.0	61,993.0	1,553,165.0	7.0	24.0	20.0	53.0	94.0	49.0	39.0
4	MEAN	12,526.0	27,127.0	444,862.0	5.0	21.0	17.0	51.0	90.0	28.0	25.0
4	STD	7,315.0	16,543.0	359,173.0	1.0	2.0	2.0	1.0	3.0	10.0	3.0

A = PST045213.
B = AGE135213.
C = AGE295213.
D = AGE775213.
E = SEX255213.
F = EDU635213.
G = EDU685213.
H = LFE305213.

Table 14. Summary statistics for Pennsylvania data continued—model estimates for variables *I* through *Q*.

Type	Statistic	*I*	*J*	*K*	*L*	*M*	*N*	*O*	*P*	*Q*
1	No. Segments	5,817	5,817	5,817	5,817	5,817	5,817	5,817	5,817	5,817
1	MIN	70,71.0	59.6	5,965.0	39,115.0	5.4	941.0	1,238.0	130.2	33.9
1	MAX	587,831.0	84.2	526,004.0	76,555.0	20.5	73,833.0	95,698.0	1,228.6	1,675.6
1	MEAN	60,026.7	73.4	53,304.0	47,603.9	13.6	7,672.9	10,309.8	731.6	196.3
1	STD	43,747.3	4.5	40,928.5	6,013.2	2.8	6,441.7	8,289.7	233.1	147.2
2	No. Segments	4,340	4,340	4,340	4,340	4,340	4,340	4,340	4,340	4,340
2	MIN	16,057.0	53.3	14,397.0	37,192.0	5.4	1,541.0	2,203.0	134.1	44.2
2	MAX	667,571.0	80.0	580,017.0	86,050.0	26.5	77,675.0	95,698.0	1,228.6	11,379.5
2	MEAN	199,906.7	70.6	180,965.5	54,722.2	12.6	28,434.4	36,364.3	695.1	1,108.6
2	STD	189,067.9	5.9	167,709.4	12,955.0	4.7	25,000.5	31,022.7	245.8	2,312.7
3	No. Segments	58,095	58,095	58,095	58,095	58,095	58,095	58,095	58,095	58,095
3	MIN	4,382.0	59.6	2,001.0	36,556.0	5.4	250.0	0.0	130.2	12.8
3	MAX	587,831.0	84.2	526,004.0	86,050.0	20.5	73,833.0	95,698.0	1,228.6	1,675.6
3	MEAN	62,360.2	74.5	55,271.9	48,178.5	13.2	8,516.2	11,287.3	769.3	198.4
3	STD	62,676.0	4.1	58,856.9	7,852.3	2.8	10,328.0	13,173.8	250.1	220.8
4	No. Segments	19,074	19,074	19,074	19,074	19,074	19,074	19,074	19,074	19,074
4	MIN	8,003.0	53.0	7,233.0	37,192.0	5.0	1,042.0	1,353.0	130.0	39.0
4	MAX	667,571.0	80.0	580,017.0	86,050.0	27.0	77,675.0	95,698.0	1,229.0	11,380.0
4	MEAN	190,998.0	72.0	174,372.0	56,612.0	12.0	28,809.0	37,125.0	678.0	905.0
4	STD	163,509.0	5.0	146,785.0	13,132.0	4.0	23,597.0	29,774.0	243.0	1,517.0

I = HSG010213.
J = HSG445213.
K = HSD410213.
L = INC110213
M = PVY020213.
N = NES010212.
O = SBO001207.
P = LND110210.
Q = POP060210.

Virginia

Virginia provided statewide roadway inventory, traffic volume, and crash data. The roadway inventory and total AADT data were provided from the Virginia Department of Transportation (VDOT) EYROAD data files. Motorcycle AADTs, where available, were queried specifically for the project and provided by VDOT staff. Crash data were downloaded from the VDOT Web site.

Roadway and traffic volume data variables included the following:

- Segment location and length.
- Functional class.
- AADT estimates for current and past years.
- Motorcycle AADT estimates for the most current year available.
- Surface width.
- Pavement width.
- Number of lanes.
- Shoulder width.
- Minimum and maximum median widths within the segment.

Crash data were provided for 2010–2014, including the following:

- Location.
- Date.
- Severity.
- Location type.
- Crash type.
- Vehicle types involved.
- Number of vehicles involved.

Total, motorcycle, multi-vehicle motorcycle, and single-vehicle motorcycle crashes were queried. All roadway, traffic, and crash data were linked together by a route name variable unique to each segment and milepost. Segments were defined so that all variables were homogeneous for the entire length.

Segments were grouped into the following four categories for validation:

- Rural freeway (type 1).
- Urban freeway (type 2).
- Rural arterial (type 3).
- Urban arterial (type 4).

Table 15 provides variable definitions in the database. Table 16 shows the total length of roadway and total number of crashes by crash type for the four site types in Virginia. Due to the low numbers of motorcycle crashes for rural and urban freeways (18 and 90, respectively), the data for freeways were not used for validation of the avenue A models. Table 17 and table 18 provide summary statistics for other variables included in the Virginia dataset for rural arterials (type 3) and urban arterials (type 4).

Table 15. Virginia data variable definitions.

Variable	Definition
AVGMOTO	Motorcycle AADT
AVGAADT	Total vehicle AADT
NOLANES	Number of lanes
SURFWIDTH	Surface width in ft
PAVEMENTWIDTH	Pavement width in ft
OUTSHLDWID	Outside shoulder width in ft
MINMEDWIDTH	Minimum median width in ft
MAXMEDWIDTH	Maximum median width in ft
INSHLDWID	Inside shoulder width in ft
LENGTH	Total segment length in mi
TOT	Total crash frequency
MOTO	Motorcycle crash frequency
MOTOSINGLE	Single-vehicle motorcycle crash frequency
MOTOMULTI	Multi-vehicle motorcycle crash frequency

Table 16. Total length and crash frequency for Virginia data.

Type	LENGTH (mi)	TOT	MOTO	MOTOSINGLE	MOTOMULTI
1	22.30	246	18	16	2
2	278.25	4,700	90	50	40
3	4,441.53	33,325	799	410	388
4	3,323.39	125,313	2,705	836	1,865

1 mi = 1.6 km.

Table 17. Summary statistics for Virginia data.

Type	Statistic	AVGMOTO (mi)	AVGAADT (mi)	NOLANES	SURFWIDTH (ft)	PAVEMENTWIDTH (ft)	OUTSHLDWID (ft)
3	No. Segments	12,574	12,574	12,574	12,574	12,574	12,574
3	MIN	1.00	305.00	1.00	12.00	0.00	0.00
3	MAX	245.00	58,715.40	7.00	86.00	90.00	33.00
3	MEAN	34.72	8,146.82	2.80	32.65	32.38	5.29
3	STD	30.69	6,816.04	0.99	12.86	15.40	2.94
4	No. Segments	25,390	25,390	25,390	25,390	25,390	25,390
4	MIN	1.00	131.00	1.00	0.00	0.00	0.00
4	MAX	509.00	130,076.50	9.00	108.00	138.00	33.50
4	MEAN	54.27	17,565.15	3.28	40.14	37.13	2.19
4	STD	55.20	14,331.02	1.28	15.54	24.13	3.12

Table 18. Summary statistics for Virginia data continued.

Type	Statistic	MINMEDWIDTH (ft)	MAXMEDWIDTH (ft)	INSHLDWID (ft)	LENGTH (ft)	TOT	MOTO	MOTOSINGLE	MOTOMULTI
3	No. Segments	12,574	12,574	12,574	12,574	12,574	12,574	12,574	12,574
3	MIN	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
3	MAX	120.00	460.00	12.00	8.59	143.00	19.00	14.00	5.00
3	MEAN	11.66	20.66	1.16	0.35	2.65	0.06	0.03	0.03
3	STD	19.76	45.31	2.02	0.47	4.72	0.35	0.25	0.19
4	Segments	25,390	25,390	25,390	25,390	25,390	25,390	25,390	25,390
4	MIN	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
4	MAX	100.00	460.00	16.50	4.40	155.00	7.00	3.00	6.00
4	MEAN	4.79	9.34	0.41	0.13	4.94	0.11	0.03	0.07
4	STD	10.11	20.58	1.31	0.17	8.53	0.37	0.19	0.30

AVENUE B DATABASES

The databases used for the avenue B analyses are a combination of actual and simulated data. The roadway inventory used total AADT and motorcycle AADT collected for the avenue A methods in Florida and Pennsylvania. This ensured that realistic combinations of roadway geometry and traffic volumes were represented. For motorcycle crashes, the crash counts were simulated using the SPFs developed in the A1 models as a starting point. Further details on this approach follow for both the B1 and B2 models.

Model B1 Approach

The model B1 approach was to develop CMFs using the EB before-after method. Using the simulated data, a countermeasure was assumed with a known value of its CMF. The CMF was applied to the simulated after period crashes for a selection of sites. Then, the CMF was estimated twice, once using the motorcycle AADTs and once for total AADT. A comparison was then made to see how the lack of motorcycle AADT affected the estimate of the CMF and its variance. The following steps discuss this process in greater detail:

1. Estimate the existing roadway geometry and traffic volume file for a given class of road (e.g., rural freeway). The appropriate SPF from the A1 modeling, including the overdispersion parameter k, is used to estimate the mean motorcycle crash frequency for each site, m_u. This represents the expected mean value of all sites with the same road geometry and traffic volumes. The SPFs used are documented in chapter 5.

2. Estimate the site-specific mean crash frequency m_i by simulating a multiplier, r, to be applied to m_u. The multiplier r is generated from a gamma distribution with the shape and rate parameters equal to the overdispersion parameter, k. This is done using the equation: $m_i = r \times m_u$.

3. Simulate the observed crash count at a site. The count in year j is assumed to follow a Poisson distribution about its mean, m_i. In this step, the number of motorcycle crashes for each site i in year j , X_{ij}, are simulated using the Poisson distribution and the site mean, m_i. Counts for 6 years are simulated.

4. Select a subset of the sites as the treatment group, with the remainder to be used as a reference group. For the treatment sites, a CMF is assumed for a fictitious treatment with a known value. The first 3 years are used as a before period and the second 3 years as an after period. The crash counts for the after period are simulated by first applying the CMF value to the site-specific mean, CMF x m_i. Then, simulate the after period crashes at the treated sites using this new expected mean as was done for all sites in step 4. For the treated sites, these new crash counts in the after period replace those simulated in step 4.

5. Use the reference group sites to estimate two SPFs to predict motorcycle crashes. The first uses motorcycle AADT as an exposure variable, and the second uses total AADT. The geometric and volume data adopted for the B1 analysis came from Florida. The A1 models developed for Florida did not always support both motorcycle and non-motorcycle volumes in the same model, and where they were both included, the model performance was roughly

the same as the model including only motorcycle AADT. For this reason, the B1 analysis did not apply models including both motorcycle and non-motorcycle AADT.

6. Use the SPFs from step 6 with the treatment site data to estimate the expected number of crashes in the after period and estimate the CMF and its variance using the EB before-after study methodology.

7. Compare the estimated CMFs and standard errors (SEs) of these estimates from step 7 to each other as well as to the assumed CMF value to determine whether using total AADT SPFs is less accurate than using SPFs with motorcycle AADTs and, if so, the magnitude of this bias.

8. Repeat steps 2 through 8 multiple times so that conclusions can be made with confidence and have broad applicability.

Model B2 Approach

The model B2 approach was to develop CMFs using cross-sectional regression modeling. In this approach, an assumed relationship was defined between one or more geometric variables and added to the SPFs developed in model A1. The relationship was defined in terms of a CMF. This modified SPF was then used to simulate the motorcycle crash counts. Then, a GLM was used to re-estimate the SPF using motorcycle AADT and then using total AADT only. A comparison was then made to see how the lack of motorcycle AADT affected the estimate of the CMF and its variance. The following steps discuss this process in greater detail:

1. Adopt the existing roadway geometry and traffic volume file for a given class of road (e.g., rural freeway). Re-estimate the appropriate SPF from the A1 modeling, including the overdispersion parameter *k*, after first assuming a CMF and including this CMF in the SPF as an offset in the model formulation. This ensures that the predicted motorcycle frequencies are realistic, given the assumed CMF value.

2. Use the new SPF from step 1 to estimate the mean motorcycle crash frequency for each site, m_u. This represents the expected mean value of all sites with the same road geometry and traffic volumes.

3. Estimate the site-specific mean crash frequency m_i by simulating a multiplier, r, and applied to m_u. The multiplier r is generated from a gamma distribution with the shape and rate parameters equal to the overdispersion parameter, k. This is done using the equation: $m_i = r \times m_u$.

4. Simulate the observed crash count at a site. The count is assumed to follow a Poisson distribution about its mean, m_i. In this step, use the Poisson distribution and site mean m_i to simulate the number of motorcycle crashes for each site for a 5-year period.

5. Use the data with simulated motorcycle crash counts to develop SPFs with the same model form as from step 1, first using motorcycle AADT as an exposure variable and then using total AADT.

6. Compare the CMFs and SEs of these estimates derived from the parameter estimates of the SPFs developed in step 6 to each other as well as to the assumed CMF value to determine whether using total AADT SPFs is less accurate than using SPFs with motorcycle AADTs and, if so, the magnitude of this bias. Repeat steps 3 through 7 so that conclusions can be made with confidence and have broad applicability.

CHAPTER 5. DATA ANALYSIS AND RESULTS

As outlined in chapter 4, the project team investigated two groups, or avenues. The methods for avenue A focus on investigating (1) the difference in predictive performance for motorcycle SPFs calibrated with motorcycle AADT versus total AADT, (2) the relation of total crash SPFs to motorcycle crash SPFs so jurisdictions without motorcycle volumes could predict motorcycle crashes using total crash SPFs, and (3) methods to predict segment-level motorcycle AADT. The methods for avenue B focus on the differences in CMF estimates found when using motorcycle AADT versus total AADT when applying before-after or cross-sectional regression CMF estimation methods.

The avenue A models were developed with data collected from Florida and Pennsylvania, both of which had a large number of locations with an estimated motorcycle AADT and that could provide linkable roadway inventory, traffic, and crash data, Data acquired from Virginia were used for validation of the models developed.

AVENUE A MODEL TYPE A1

The purpose of model type A1 was to explore how much predictive power is lost when motorcycle volumes are unknown and how this lack of information would affect an evaluation of motorcycle countermeasures. The project team attempted models for all motorcycle crashes (MOTO), single-vehicle motorcycle crashes (MOTOSINGLE), and multi-vehicle motorcycle crashes (MOTOMULTI).

For each dependent variable, the project team attempted three separate models, with motorcycle AADT, total AADT, and motorcycle and non-motorcycle AADT as separate terms. In some cases, mainly due to limited samples, it was not possible to develop robust models for each of the three dependent variables and for each of the three independent AADT variable specifications.

Additional variables were included where possible, but it should be noted that the models developed were done to derive the most accurate models for predicting crash frequency, as opposed to the objectives of causal models. With this goal, it is possible to include variables that indicate an opposite relationship to crashes than what is expected or to treat a logically categorical variable, such as speed, as continuous. If a variable correlates with other variables that affect crash risk, then the counterintuitive relationship frequently occurs.

To assess a variable for inclusion in the model, it was added to the basic SPF including the AADT terms and the model estimated. Then, an assessment was made of the improvement in fit with the additional variable as measured by the overdispersion parameter and the statistical significance of the parameter estimate.

The project team developed the A1 models using both Florida and Pennsylvania data and using data from Virginia for validation. The first three sections of this chapter detail the analysis and results based on the data from these three States. The final section presents the results of the assessment of how well the four sets of models predict motorcycle crashes for high-crash locations.

Florida

Models were successfully calibrated for site types 1–4. The data for site types 5 and 6 (rural collector/local and urban collector/local) did not allow for models to be developed. As table 4 shows, these site types had very few motorcycle crashes.

Note that the posted speed limit variable, SPDLIMT, which appears in the types 2–4 models, has been modeled as a continuous variable. Often, posted speed is considered as a categorical variable since limits are typically set in multiples of 5, (i.e. 45 mi/h (72 km/h), 55 mi/h (89 km/h), etc.). During model development, however, treating posted speed as a categorical variable showed inconsistent results and no logical groupings of posted speed into similar categories.

The following sections detail the calibrated models and goodness-of-fit assessments for each site type.

Florida Type 1—Rural Freeways

For rural freeways, models were successfully calibrated for MOTO and MOTOSINGLE crashes but not for multi-vehicle motorcycle crashes. The models are of two forms, dependent on the exposure measure used, as seen in figure 13 and figure 14.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}\,AVGMOTO^{c}$$

Figure 13. Equation. Florida rural freeway crashes per year using motorcycle traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}\,AVGAADT^{c}$$

Figure 14. Equation. Florida rural freeway crashes per year using total traffic counts.

Attempts to estimate models with separate terms for motorcycle and non-motorcycle volumes were not successful because illogical parameter estimates resulted in decreasing crash frequencies with motorcycle volumes. Such illogical results can occur when highly correlated variables are included in the same model. Table 19 provides the parameter estimates for these models with the SE provided in parenthesis after the parameter estimates. Table 20 provides overall goodness-of-fit statistics. Table 21 provides goodness-of-fit statistics from CURE plots for variables included in the models.

Table 19. A1 models for Florida type 1 sites.

Model Type	Parameter	MOTO (SE)	MOTOSINGLE (SE)
Motorcycle AADT	Intercept	-5.5368 (0.6033)	-5.1341 (0.7741)
Motorcycle AADT	*b*	0.8239 (0.0801)	0.8939 (0.1102)
Motorcycle AADT	*c*	0.6622 (0.1274)	0.4755 (0.1663)
Motorcycle AADT	Dispersion	0.4353 (0.2160)	1.1346 (0.4327)
Total AADT	Intercept	-15.2081 (1.7870)	-14.3297 (2.4202)
Total AADT	*b*	0.8289 (0.0780)	0.9101 (0.1099)
Total AADT	*c*	1.2118 (0.1684)	1.0847 (0.2292)
Total AADT	Dispersion	0.2886 (0.1920)	0.9233 (0.3863)

Table 20. Goodness-of-fit statistics for A1 models for Florida type 1 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	174	0.43	0.65	0.44
MOTO	Total AADT	174	0.41	0.77	0.29
MOTOSINGLE	Motorcycle AADT	111	0.32	0.34	1.14
MOTOSINGLE	Total AADT	111	0.31	0.44	0.92

Table 21. CURE plot statistics for A1 models for Florida type 1 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)
MOTO	Motorcycle AADT	15.83	11	7.42	3
MOTO	Total AADT	12.55	5	8.35	3
MOTOSINGLE	Motorcycle AADT	10.42	5	7.80	3
MOTOSINGLE	Total AADT	10.83	5	6.83	3

The goodness-of-fit statistics indicate that the models using motorcycle AADT performed very similarly to those using total AADT for both MOTO and MOTOSINGLE crashes. In fact, there are some indications that the models using total AADT may be slightly better when considering the modified R^2 and dispersion parameter.

Florida Type 2—Urban Freeways

For urban freeways, the project team successfully calibrated models for MOTO, MOTOSINGLE, and MOTOMULTI crashes. The models are of two forms, depending on the exposure measure used, as shown in figure 15 and figure 16.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}\, AVGMOTO^{c}\, exp^{(d*CURVE + e*SPDLIMT + f*SURFWIDTH\ )}$$

Figure 15. Equation. Florida urban freeway crashes per year using motorcycle traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}\, AVGAADT^{c}\, exp^{(d*CURVE + e*SPDLIMT + f*SURFWIDTH\ )}$$

Figure 16. Equation. Florida urban freeway crashes per year using total traffic counts.

Where:

CURVE = 1 if a horizontal curve is present in the segment and 0 if not.

Table 3 provides definitions for the other variables. As with the results for type 1, illogical parameter estimates resulted when terms for motorcycle and non-motorcycle volumes were included in the same model. Table 22 provides the parameter estimates for these models, and table 23 provides overall goodness-of-fit statistics. Table 24 provides goodness-of-fit statistics from CURE plots for variables included in the models. Table 25 provides calibration factors for each level of the non-continuous variables that were included in at least one of the models for that crash type.

Table 22. A1 models for Florida type 2 sites.

Model Type	Parameter	MOTO (SE)	MOTOSINGLE (SE)	MOTOMULTI (SE)
Motorcycle AADT	Intercept	-0.0067 (0.4894)	-1.1643 (0.6538)	-0.2295 (0.5733)
Motorcycle AADT	*b*	0.8107 (0.0480)	0.7942 (0.1191)	0.8121 (0.0579)
Motorcycle AADT	*c*	0.1628 (0.0447)	0.1191 (0.0564)	0.1693 (0.0513)
Motorcycle AADT	*d*	0.2459 (0.0987)	N/A	0.3874 (0.1236)
Motorcycle AADT	*e*	-0.0532 (0.0067)	-0.0354 (0.0091)	-0.0654 (0.0080)
Motorcycle AADT	*f*	0.0173 (0.0023)	0.0116 (0.0031)	0.0212 (0.0027)
Motorcycle AADT	Dispersion	0.4300 (0.0757)	0.2592 (0.1224)	0.3766 (0.0993)
Total AADT	Intercept	-10.3484 (0.8758)	-8.4768 (1.2158)	-13.1836 (1.1053)
Total AADT	*b*	0.8080 (0.0442)	0.7853 (0.0580)	0.8194 (0.0547)
Total AADT	*c*	1.0563 (0.0662)	0.7429 (0.0885)	1.2948 (0.0842)
Total AADT	*d*	0.2294 (0.0913)	N/A	0.3752 (0.1153)
Total AADT	*e*	-0.0447 (0.0064)	-0.0290 (0.0091)	-0.0537 (0.0078)
Total AADT	*f*	N/A	N/A	N/A
Total AADT	dispersion	0.2285 (0.0570)	0.1309 (0.1061)	0.1670 (0.0707)

N/A = Not applicable.

Table 23. Goodness-of-fit statistics for A1 models for Florida type 2 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	921	0.86	0.56	0.43
MOTO	Total AADT	921	0.77	0.73	0.23
MOTOSINGLE	Motorcycle AADT	385	0.47	0.66	0.26
MOTOSINGLE	Total AADT	385	0.46	0.82	0.13
MOTOMULTI	Motorcycle AADT	536	0.58	0.59	0.38
MOTOMULTI	Total AADT	536	0.50	0.77	0.17

Table 24. CURE plot statistics for A1 models for Florida type 2 sites.

Crash Type	Exposure Measure	Max Curve Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)	Max CURE Deviation for SPDLIMT	CURE Deviation for SPDLIMT (Percent)	Max CURE Deviation for SURFWIDTH	CURE Deviation for SURFWIDTH (Percent)
MOTO	Motorcycle AADT	54.69	35	22.11	0	65.75	46	67.7	28
MOTO	Total AADT	27.39	1	21.31	0	25.59	0	41.21	9
MOTOSINGLE	Motorcycle AADT	26.21	16	11.53	0	19.10	4	17.08	10
MOTOSINGLE	Total AADT	15.81	0	11.83	0	18.40	3	16.58	0
MOTOMULTI	Motorcycle AADT	37.34	41	18.12	16	37.07	19	42.43	30
MOTOMULTI	Total AADT	14.23	2	15.88	11	15.99	1	36.81	31

Table 25. Calibration factors for A1 models for Florida type 2 sites.

Crash SPF Conditions	Observed	Observed/Predicted
MOTO crash SPF using motorcycle AADT with no curvature	706	1.02
MOTO crash SPF using motorcycle AADT with curvature	215	0.95
MOTO crash SPF with total AADT with no curvature	706	1.01
MOTO crash SPF with total AADT with curvature	215	0.97
MOTOMULTI crash SPF using motorcycle AADT with no curvature	423	1.01
MOTOMULTI crash SPF using motorcycle AADT with curvature	113	0.95
MOTOMULTI crash SPF with total AADT with no curvature	423	1.01
MOTOMULTI crash SPF with total AADT with curvature	113	0.98

As was the case for type 1 sites (rural freeways), the goodness-of-fit statistics indicate that the models using total AADT not only performed similarly to those with motorcycle AADT but may be slightly better, especially when considering the modified R^2 and dispersion parameter. The calibration factors indicate that both sets of AADT predictor models were just as successful when applied to segments with or without horizontal curves.

Florida Type 3—Rural Arterials

For rural arterials, the project team successfully calibrated models for MOTO, MOTOSINGLE, and MOTOMULTI crashes.

The models are of three forms, one of which uses estimates of motorcycle AADT and other AADT (AVGOTHER = AVGAADT – AVGMOTO), depending on the exposure measure used, as shown in figure 17 through figure 19.

$$Crashes/year = exp^{Inter\ cept}\, LENGTH^b AVGMOTO^c exp^{(e*SPDLIMT + f*DIVUND + g*OUTSHLDWID + h*SURFWIDTH + i*MEDWIDTH)}$$

Figure 17. Equation. Florida rural arterial crashes per year using motorcycle traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^b AVGAADT^c exp^{(e*SPDLIMT + f*DIVUND + g*OUTSHLDWID + h*SURFWIDTH + i*MEDWIDTH)}$$

Figure 18. Equation. Florida rural arterial crashes per year using total traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}\, AVGMOTO^{c}\, AVGOTHER^{d}\, exp^{(e*SPDLIMT + f*DIVUND + g*OUTSHLDWID + h*SURFWIDTH + i*MEDWIDTH)}$$

Figure 19. Equation. Florida rural arterial crashes per year using separate traffic counts.

Where:

DIVUND = 1 if road is undivided; 0 if road is divided. (Note: medwidth is 0 for an undivided road.)

AVGOTHER = Non-motorcycle AADT = AVGAADT – AVGMOTO.

Table 26 provides the parameter estimates for these models, and table 27 provides overall goodness-of-fit statistics. Table 28 provides goodness-of-fit statistics from CURE plots for variables included in the models. Table 29 through table 31 provide calibration factors for each level of the non-continuous variables that were included in at least one of the models for that crash type.

Table 26. A1 models for Florida type 3 sites.

Model Type	Parameter	MOTO (SE)	MOTOSINGLE (SE)	MOTOMULTI (SE)
Motorcycle AADT	Intercept	-1.8929 (0.4016)	-3.5059 (0.5816)	-2.0657 (0.5044)
Motorcycle AADT	*b*	0.8158 (0.0380)	0.8112 (0.0517)	0.8105 (0.0485)
Motorcycle AADT	*c*	0.4904 (0.0442)	0.3901 (0.0611)	0.5806 (0.0568)
Motorcycle AADT	*d*	N/A	N/A	N/A
Motorcycle AADT	*e*	-0.0323 (0.0062)	-0.0206 (0.0090)	-0.0418 (0.0078)
Motorcycle AADT	*f*	-0.6195 (0.0875)	N/A	-0.6275 (0.1100)
Motorcycle AADT	*g*	-0.0595 (0.0211)	N/A	-0.1043 (0.0295)
Motorcycle AADT	*h*	N/A	N/A	N/A
Motorcycle AADT	*i*	N/A	0.0053 (0.0024)	N/A
Motorcycle AADT	Dispersion	0.7653 (0.1091)	0.9094 (0.2160)	0.8522 (0.1708)
Total AADT	Intercept	-5.8092 (0.6436)	-5.6814 (0.8963)	-7.4544 (0.8249)
Total AADT	*b*	0.8299 (0.0392)	0.8260 (0.0518)	0.8080 (0.0482)
Total AADT	*c*	0.7014 (0.0649)	0.4310 (0.0836)	0.8506 (0.0786)
Total AADT	*d*	N/A	N/A	N/A
Total AADT	*e*	-0.0332 (0.0062)	-0.0215 (0.0088)	-0.0421 (0.0077)
Total AADT	*f*	-0.5221 (0.1262)	N/A	N/A
Total AADT	*g*	-0.0881 (0.0217)	N/A	-0.1364 (0.0296)
Total AADT	*h*	N/A	N/A	N/A
Total AADT	Dispersion	0.7699 (0.1093)	1.0111 (0.2286)	0.8010 (0.1671)
Motorcycle AADT and other AADT	Intercept	-4.5690 (0.6860)	-4.4664 (0.9559)	-6.1709 (0.8906)
Motorcycle AADT and other AADT	*b*	0.8158 (0.0377)	0.8254 (0.0516)	0.8103 (0.0480)
Motorcycle AADT and other AADT	*c*	0.3042 (0.0581)	0.3203 (0.0800)	0.2960 (0.0747)
Motorcycle AADT and other AADT	*d*	0.3820 (0.0798)	0.1493 (0.1090)	0.5771 (0.1040)

Model Type	Parameter	MOTO (SE)	MOTOSINGLE (SE)	MOTOMULTI (SE)
Motorcycle AADT and other AADT	*e*	-0.0333 (0.0062)	-0.0196 (0.0089)	-0.0422 (0.0077)
Motorcycle AADT and other AADT	*f*	-0.4142 (0.0967)	-0.5140 (0.1375)	-0.3190 (0.1218)
Motorcycle AADT and other AADT	*g*	-0.0714 (0.0214)	N/A	-0.1214 (0.0295)
Motorcycle AADT and other AADT	*h*	N/A	N/A	N/A
Motorcycle AADT and other AADT	Dispersion	0.7116 (0.1055)	0.9179 (0.2178)	0.7421 (0.1611)

N/A =Not applicable.

Table 27. Goodness-of-fit statistics for A1 models for Florida type 3 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	1,031	0.40	0.44	0.77
MOTO	Total AADT	1,031	0.40	0.44	0.77
MOTO	Motorcycle and other AADT	1,031	0.39	0.48	0.71
MOTOSINGLE	Motorcycle AADT	449	0.21	0.43	0.91
MOTOSINGLE	Total AADT	449	0.21	0.39	1.01
MOTOSINGLE	Motorcycle and other AADT	449	0.21	0.43	0.92
MOTOMULTI	Motorcycle AADT	582	0.26	0.37	0.85
MOTOMULTI	Total AADT	582	0.26	0.43	0.80
MOTOMULTI	Motorcycle and other AADT	582	0.26	0.46	0.74

Table 28. CURE plot statistics for A1 models for Florida type 3 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)	Max CURE Deviation for SPDLIMT	CURE Deviation for SPDLIMT (Percent)	Max CURE Deviation for OUTSHLDWID	CURE Deviation for OUTSHLDWID (Percent)
MOTO	Motorcycle AADT	37.40	0	21.32	0	50.60	7	32.36	1
MOTO	Total AADT	88.41	70	18.85	0	60.83	29	26.11	1
MOTO	Motorcycle and other AADT	38.67	1	17.85	0	37.79	3	29.70	3
MOTOSINGLE	Motorcycle AADT	24.33	1	8.21	0	16.28	3	15.25	0
MOTOSINGLE	Total AADT	47.99	56	9.61	0	24.74	8	10.86	0
MOTOSINGLE	Motorcycle and other AADT	25.94	2	9.01	0	15.20	3	13.29	0
MOTOMULTI	Motorcycle AADT	19.71	1	20.19	6	26.67	3	18.19	2
MOTOMULTI	Total AADT	44.70	47	14.53	6	35.54	23	27.60	4
MOTOMULTI	Motorcycle and other AADT	20.14	1	15.73	6	24.40	3	23.13	3

Table 29. Calibration factors for A1 MOTO models for Florida type 3 sites.

Crash SPF Conditions	Observed	Observed/Predicted
MOTO crash SPF using motorcycle AADT for divided roads	529	1.00
MOTO crash SPF using motorcycle AADT for undivided roads	502	1.00
MOTO crash SPF with total AADT for divided roads	529	1.00
MOTO crash SPF with total AADT for undivided roads	502	1.00
MOTO crash SPF using motorcycle AADT and non-motorcycle AADT for divided roads	529	1.00
MOTO crash SPF using motorcycle AADT and non-motorcycle AADT for undivided roads	502	1.00
MOTO crash SPF using motorcycle AADT with no curvature	706	1.02
MOTO crash SPF using motorcycle AADT with curvature	215	0.95
MOTO crash SPF with total AADT with no curvature	706	1.02
MOTO crash SPF with total AADT with curvature	215	0.95

Table 30. Calibration factors for A1 MOTOSINGLE models for Florida type 3 sites.

Crash SPF Conditions	Observed	Observed/Predicted
MOTOSINGLE crash SPF using motorcycle AADT for divided roads	224	1.00
MOTOSINGLE crash SPF using motorcycle AADT for undivided roads	225	1.00
MOTOSINGLE crash SPF with total AADT for divided roads	224	1.01
MOTOSINGLE crash SPF with total AADT for undivided roads	225	0.99
MOTOSINGLE crash SPF using motorcycle AADT and non-motorcycle AADT for divided roads	224	1.01
MOTOSINGLE crash SPF using motorcycle AADT and non-motorcycle AADT for undivided roads	225	0.99

Table 31. Calibration factors for A1 MOTOMULIT models for Florida type 3 sites.

Crash SPF Conditions	Observed	Observed/Predicted
MOTOMULTI crash SPF using motorcycle AADT for divided roads	305	0.99
MOTOMULTI crash SPF using motorcycle AADT for undivided roads	277	1.01
MOTOMULTI crash SPF with total AADT for divided roads	305	0.99
MOTOMULTI crash SPF with total AADT for undivided roads	277	1.01
MOTOMULTI crash SPF using motorcycle AADT and non-motorcycle AADT for divided roads	305	0.99
MOTOMULTI crash SPF using motorcycle AADT and non-motorcycle AADT for undivided roads	277	1.01
MOTOMULTI crash SPF using motorcycle AADT with no curvature	423	1.01
MOTOMULTI crash SPF using motorcycle AADT with curvature	113	0.95
MOTOMULTI crash SPF with total AADT with no curvature	423	1.01
MOTOMULTI crash SPF with total AADT with curvature	113	0.95

The goodness-of-fit results in table 27 indicate overall that models with total AADT performed as well as those with motorcycle AADT in terms of MAD, modified R^2, and dispersion parameter. However, the CURE plot statistics indicate that the models with motorcycle AADT outperformed those with total AADT for these measures. As expected, some measures (MAD, modified R^2, and dispersion parameter) indicate that some models (single- and multi-vehicle crashes) that include both motorcycle AADT and non-motorcycle AADT can outperform models with only motorcycle AADT. Finally, the calibration factors in table 29 indicate that all three sets of AADT predictor models were just as successful when applied to segments with the two levels of the two indicator variables.

Florida Type 4—Urban Arterials

For urban arterials, the project team successfully calibrated models for MOTO, MOTOSINGLE, and MOTOMULTI crashes. The models are of three forms, depending on the exposure measure used, as shown in figure 20 through figure 22.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}\, AVGMOTO^{c}\, exp^{(e*SPDLIMT\,+f*DIVUND\,+g*OUTSHLDWID\,+h*SURFWIDTH\,)}$$

Figure 20. Equation. Florida urban arterial crashes per year using motorcycle traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}AVGAADT^{c}exp^{(e*SPDLIMT+f*DIVUND+g*OUTSHLDWID+h*SURFWIDTH)}$$

Figure 21. Equation. Florida urban arterial crashes per year using total traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}AVGMOTO^{c}AVGOTHER^{d}exp^{(e*SPDLIMT+f*DIVUND+g*OUTSHLDWID+h*SURFWIDTH)}$$

Figure 22. Equation. Florida urban arterial crashes per year using separate traffic counts.

Where:

DIVUND = 1 if road is undivided and 0 if road is divided.
AVGOTHER = The non-motorcycle AADT = AVGAADT – AVGMOTO.

Table 32 provides the parameter estimates for these models, and table 33 provides overall goodness-of-fit statistics. Table 34 provides goodness-of-fit statistics from CURE plots for variables included in the models. Table 35 through table 37 provide calibration factors for each level of the non-continuous variables that were included in at least one of the models for that crash type.

Table 32. A1 models for Florida type 4 sites.

Model Type	Parameter	MOTO (SE)	MOTOSINGLE (SE)	MOTOMULTI (SE)
Motorcycle AADT	Intercept	-2.4685 (0.1821)	-4.5196 (0.2039)	-3.2128 (0.1911)
Motorcycle AADT	*b*	0.7885 (0.0219)	0.7349 (0.0297)	0.7895 (0.0240)
Motorcycle AADT	*c*	0.6274 (0.0253)	0.4006 (0.0414)	0.7543 (0.0274)
Motorcycle AADT	*d*	N/A	N/A	N/A
Motorcycle AADT	*e*	-0.0238 (0.0026)	N/A	-0.0278 (0.0029)
Motorcycle AADT	*f*	-0.6686 (0.0595)	-0.4541 (0.1035)	N/A
Motorcycle AADT	*g*	-0.0695 (0.0070)	-0.0361 (0.0106)	-0.0905 (0.0078)
Motorcycle AADT	*h*	N/A	0.0105 (0.0018)	N/A
Motorcycle AADT	Dispersion	1.1890 (0.0433)	1.1191 (0.0940)	1.3242 (0.0536)
Total AADT	Intercept	-9.0362 (0.3447)	-9.0785 (0.4910)	-10.4836 (0.3447)
Total AADT	*b*	0.7936 (0.0213)	0.7408 (0.0296)	0.8154 (0.0233)
Total AADT	*c*	0.9657 (0.0317)	0.7026 (0.0472)	1.1176 (0.0325)
Total AADT	*d*	N/A	N/A	N/A
Total AADT	*e*	-0.0277 (0.0025)	N/A	-0.0361 (0.0028)
Total AADT	*f*	-0.2442 (0.0625)	-0.3694 (0.1022)	N/A
Total AADT	*g*	-0.0563 (0.0069)	-0.0375 (0.0105)	-0.0653 (0.0076)
Total AADT	*h*	N/A	N/A	N/A
Total AADT	Dispersion	1.0505 (0.0403)	1.0857 (0.0931)	1.1042 (0.0478)
Motorcycle AADT and other AADT	Intercept	-8.3451 (0.3585)	-8.2796 (0.5137)	-9.8429 (0.3615)
Motorcycle AADT and other AADT	*b*	0.7923 (0.0212)	0.7423 (0.0296)	0.8138 (0.0232)

Model Type	Parameter	MOTO (SE)	MOTOSINGLE (SE)	MOTOMULTI (SE)
Motorcycle AADT and other AADT	*c*	0.2018 (0.0327)	0.2307 (0.0495)	0.1862 (0.0360)
Motorcycle AADT and other AADT	*d*	0.7935 (0.0416)	0.5094 (0.0619)	0.9587 (0.0440)
Motorcycle AADT and other AADT	*e*	-0.0267 (0.0025)	N/A	-0.0352 (0.0028)
Motorcycle AADT and other AADT	*f*	-0.2536 (0.0624)	-0.3815 (0.1021)	N/A
Motorcycle AADT and other AADT	*g*	-0.0570 (0.0069)	-0.0368 (0.0105)	-0.0663 (0.0076)
Motorcycle AADT and other AADT	*h*	N/A	N/A	N/A
Motorcycle AADT and other AADT	Dispersion	1.0370 (0.0400)	1.0688 (0.0922)	1.0924 (0.0475)

N/A = Not applicable.

Table 33. Goodness-of-fit statistics for A1 models for Florida type 4 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	9,539	1.11	0.28	1.19
MOTO	Total AADT	9,539	1.06	0.33	1.05
MOTO	Motorcycle and other AADT	9,539	1.06	0.34	1.04
MOTOSINGLE	Motorcycle AADT	2,341	0.38	0.31	1.12
MOTOSINGLE	Total AADT	2,341	0.38	0.33	1.09
MOTOSINGLE	Motorcycle and other AADT	2,341	0.38	0.34	1.07
MOTOMULTI	Motorcycle AADT	7,198	0.91	0.25	1.32
MOTOMULTI	Total AADT	7,198	0.86	0.32	1.10
MOTOMULTI	Motorcycle and other AADT	7,198	0.86	0.33	1.09

Table 34. CURE plot statistics for A1 models for Florida type 4 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)	Max CURE Deviation for SPDLIMT	CURE Deviation for SPDLIMT (Percent)	Max CURE Deviation for OUTSHLDWID	CURE Deviation for OUTSHLDWID (Percent)
MOTO	Motorcycle AADT	242.23	14	120.12	0	392.26	24	210.64	6
MOTO	Total AADT	468.42	84	143.90	1	315.80	30	158.88	6
MOTO	Motorcycle and other AADT	219.81	10	145.09	1	203.80	6	143.64	6
MOTOSINGLE	Motorcycle AADT	52.04	2	51.84	0	65.67	15	72.11	13
MOTOSINGLE	Total AADT	121.49	69	58.40	1	77.36	36	81.69	34
MOTOSINGLE	Motorcycle and other AADT	51.58	1	58.64	1	77.66	42	84.46	39
MOTOMULTI	Motorcycle AADT	178.29	8	120.03	1	223.54	3	157.08	1
MOTOMULTI	Total AADT	324.78	80	119.832	1	167.83	8	98.79	1
MOTOMULTI	Motorcycle and other AADT	154.31	3	125.87	1	135.86	3	101.34	5

Table 35. Calibration factors for A1 MOTO models for Florida type 4 sites.

Crash SPF Conditions	Observed	Observed/Predicted
MOTO crash SPF using motorcycle AADT for divided roads	623	1.09
MOTO crash SPF using motorcycle AADT for undivided roads	8,916	0.99
MOTO crash SPF with total AADT for divided roads	623	1.10
MOTO crash SPF with total AADT for undivided roads	8,916	0.99
MOTO crash SPF using motorcycle AADT and non-motorcycle AADT for divided roads	623	1.10
MOTO crash SPF using motorcycle AADT and non-motorcycle AADT for undivided Roads	8,916	0.99

Table 36. Calibration factors for A1 MOTOSINGLE models for Florida type 4 sites.

Crash SPF Conditions	Observed	Observed/Predicted
MOTOSINGLE crash SPF using motorcycle AADT for divided roads	151	1.00
MOTOSINGLE crash SPF using motorcycle AADT for undivided roads	2,190	1.00
MOTOSINGLE crash SPF with total AADT for divided roads	151	1.00
MOTOSINGLE crash SPF with total AADT for undivided roads	2,190	1.00
MOTOSINGLE crash SPF using motorcycle AADT and non-motorcycle AADT for divided roads	151	1.00
MOTOSINGLE crash SPF using motorcycle AADT and non-motorcycle AADT for undivided roads	2,190	1.00

Table 37. Calibration factors for A1 MOTOMULIT models for Florida type 4 sites.

Crash SPF Conditions	Observed	Observed/Predicted
MOTOMULTI crash SPF using motorcycle AADT for divided roads	472	0.66
MOTOMULTI crash SPF using motorcycle AADT for undivided roads	6,726	1.04
MOTOMULTI crash SPF with total AADT for divided roads	472	1.00
MOTOMULTI crash SPF with total AADT for undivided roads	6,726	1.00
MOTOMULTI crash SPF using motorcycle AADT and non-motorcycle AADT for divided roads	472	0.99
MOTOMULTI crash SPF using motorcycle AADT and non-motorcycle AADT for undivided roads	6,726	1.00

The goodness-of-fit results in table 33 indicate overall that models with total AADT performed at least as well as those with motorcycle AADT in terms of MAD, modified R^2, and dispersion parameter. In fact, models with total AADT were somewhat better by these measures for MOTO and MOTOMULTI. However, the CURE plot statistics for the key AADT variable in table 34 indicate the opposite—the models with motorcycle AADT outperformed those with total AADT for the CURE measures. Not surprisingly, some measures (modified R^2, dispersion parameter, and CURE plot measures) indicate that the three sets of models that include both motorcycle AADT and non-motorcycle AADT can outperform models with only motorcycle AADT. Finally, the calibration factors in table 35 through table 37 indicate that all three sets of AADT predictor models were just as successful when applied to segments with the two levels of the two indicator variables, with the exception of multi-vehicle motorcycle crashes on divided roads that are overpredicted.

Pennsylvania

Using Pennsylvania data, the project team could not calibrate separate models by area type (urban versus rural), so data were combined by area type to develop the SPFs with the use of an area type indicator variable.

Pennsylvania Type 1 and 2—Rural and Urban Freeways

For Pennsylvania freeways, the data and models pertain to one direction of travel only. For freeways, the project team successfully calibrated models for MOTO and MOTOMULTI crashes but not for single-vehicle motorcycle crashes, likely because of the small sample. The models are of three forms, depending on the exposure measure used, as shown in figure 23 through figure 25.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b} AVGMOTO^{c} exp^{(e*URBRUR+f*WIDTH+g*LSHLDWID)}$$

Figure 23. Equation. Pennsylvania rural and urban freeway crashes per year using motorcycle traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b} AVGAADT^{c} exp^{(e*URBRUR+f*WIDTH+g*LSHLDWID)}$$

Figure 24. Equation. Pennsylvania rural and urban freeway crashes per year using total traffic counts.

$$Crashes/year \\ = exp^{Intercept}\ LENGTH^{b} AVGMOTO^{c} AVGOTHER^{d} exp^{(e*URBRUR+f*WIDTH+g*LSHLDWID)}$$

Figure 25. Equation. Pennsylvania rural and urban freeway crashes per year using separate traffic counts.

Where:

$URBRUR$ = 1 if rural and 0 otherwise.
$AVGOTHER$ = The non-motorcycle AADT = AVGAADT – AVGMOTO.

Other variables as defined in table 9.

Table 38 provides the parameter estimates for these models, and table 39 provides overall goodness-of-fit statistics. Table 40 provides goodness-of-fit statistics from CURE plots for variables included in the models. Table 41 provides calibration factors for the categorical variables included in the models.

Table 38. A1 models for Pennsylvania type 1 and 2 sites.

Model Type	Parameter	MOTO (SE)	MOTOMULTI (SE)
Motorcycle AADT	Intercept	-5.0846 (0.2753)	-5.7901 (0.3879)
Motorcycle AADT	*b*	0.7159 (0.1689)	0.6652 (0.2717)
Motorcycle AADT	*c*	0.2309 (0.0399)	0.4313 (0.0686)
Motorcycle AADT	*d*	N/A	N/A
Motorcycle AADT	*e*	-0.8489 (0.0873)	-1.5526 (0.1657)
Motorcycle AADT	*f*	0.0505 (0.0048)	N/A
Motorcycle AADT	*g*	-0.0333 (0.0111)	N/A
Motorcycle AADT	Dispersion	1.2724 (0.2056)	3.2825 (0.7948)
Total AADT	Intercept	-7.4047 (0.4663)	-10.9414 (0.7837)
Total AADT	*b*	0.6908 (0.1621)	0.6018 (0.2491)
Total AADT	*c*	0.3776 (0.0468)	0.7716 (0.0803)
Total AADT	*d*	N/A	N/A
Total AADT	*e*	-0.6855 (0.0921)	-1.1491 (0.1752)
Total AADT	*f*	0.0467 (0.0047)	N/A
Total AADT	*g*	-0.0293 (0.0110)	N/A
Total AADT	Dispersion	1.1204 (0.1918)	2.5169 (0.6548)
Motorcycle and Other AADT	Intercept	-7.3820 (0.4615)	
Motorcycle and Other AADT	*b*	0.6891 (0.1628)	
Motorcycle and Other AADT	*c*	0.1256 (0.0420)	
Motorcycle and Other AADT	*d*	0.3151 (0.0501)	
Motorcycle and Other AADT	*e*	-0.6651 (0.0924)	

Model Type	Parameter	MOTO (SE)	MOTOMULTI (SE)
Motorcycle and Other AADT	*f*	0.0460 (0.0047)	
Motorcycle and Other AADT	*g*	-0.0291 (0.0110)	
Motorcycle and Other AADT	Dispersion	1.1196 (0.1912)	

N/A = Parameter was not included in the model.
Blank cell = SPF was not calibrated for that crash type.

Table 39. Goodness-of-fit statistics for A1 models for Pennsylvania type 1 and 2 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	849	0.18	0.26	1.27
MOTO	Total AADT	849	0.18	0.31	1.12
MOTO	Motorcycle and other AADT	849	0.18	0.31	1.12
MOTOMULTI	Motorcycle AADT	283	0.07	0.21	3.28
MOTOMULTI	Total AADT	283	0.07	0.21	2.52

Table 40. CURE plot statistics for A1 models for Pennsylvania type 1 and 2 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)	Max CURE Deviation for WIDTH	CURE Deviation for WIDTH	Max CURE Deviation for LSHLDWID	CURE Deviation for LSHLDWID (Percent)
MOTO	Motorcycle AADT	33.56	3	41.11	27	31.95	3	38.33	25
MOTO	Total AADT	38.67	31	31.14	5	26.73	1	35.78	22
MOTO	Motorcycle and other AADT	38.30	9	32.32	7	26.11	1	37.42	23
MOTOMULTI	Motorcycle AADT	24.52	16	30.00	45	70.57	73	29.44	39
MOTOMULTI	Total AADT	34.42	51	23.89	10	66.04	77	22.62	15

Table 41. Calibration factors for A1 models for Pennsylvania type 1 and 2 sites.

Crash Type	Observed	Observed/Predicted
MOTO with AVGMOTO urban	420	1.07
MOTO with AVGMOTO rural	359	0.93
MOTO with AVGAADT urban	420	1.07
MOTO with AVGAADT rural	359	0.92
MOTO with AVGAADT and OTHER urban	420	1.06
MOTO with AVGAADT and OTHER rural	359	0.94

The goodness-of-fit results in table 39 indicate the opposite result from the previous models—the models with motorcycle AADT outperformed those with total AADT for the CURE measures. Not surprisingly, some measures (modified R^2 and dispersion parameter) indicate that the MOTO model that includes both motorcycle AADT and non-motorcycle AADT can outperform the model with only motorcycle AADT. However, its performance by these measures is similar to that of the MOTO model
with only total AADT. Finally, the calibration factors in table 41 indicate that all AADT predictor models were just as successful when applied to segments with the two levels of the indicator variables.

Pennsylvania Type 3 and 4—Rural and Urban Non-Freeways

For non-freeways, the project team successfully calibrated models only for MOTO crashes and not for single- or multi-vehicle motorcycle crashes. The models are of three forms, depending on the exposure measure used, as shown in figure 26 through figure 28.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}AVGMOTO^{c}exp^{(e*URBRUR+f*WIDTH+g*AVGSHLDWID)}$$

Figure 26. Equation. Pennsylvania rural and urban non-freeway crashes per year using motorcycle traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}AVGAADT^{c}exp^{(e*URBRUR+f*WIDTH+g*AVGSHLDWID)}$$

Figure 27. Equation. Pennsylvania rural and urban non-freeway crashes per year using total traffic counts.

$$Crashes/year = exp^{Intercept}\ LENGTH^{b}AVGMOTO^{c}AVGOTHER^{d}exp^{(e*URBRUR+f*WIDTH+g*AVGSHLDWID)}$$

Figure 28. Equation. Pennsylvania rural and urban non-freeway crashes per year using separated traffic counts.

Where:

URBRUR = 1 if rural and 0 otherwise.
AVGSHLDWID = The average shoulder width on both sides of roadway.
AVGOTHER = The non-motorcycle AADT = AVGAADT – AVGMOTO.

Table 42 provides the parameter estimates for these models, and table 43 provides overall goodness-of-fit statistics. The project team calibrated the models using the full dataset, but the goodness-of-fit statistics pertain to a random sample of 6,000 road segments. The project team used a subset to make the calculations and CURE plots more manageable. Table 44 provides goodness-of-fit statistics from CURE plots for variables included in the models, and table 45 provides calibration factors for the categorical variables included in the models.

Table 42. A1 models for Pennsylvania type 3 and 4 sites.

Model Type	Parameter	MOTO
Motorcycle AADT	Intercept	-4.7635 (0.0667)
Motorcycle AADT	*b*	0.7377 (0.0352)
Motorcycle AADT	*c*	0.4585 (0.0123)
Motorcycle AADT	*d*	
Motorcycle AADT	*e*	-0.4861 (0.0295)
Motorcycle AADT	*f*	0.0263 (0.0020)
Motorcycle AADT	*g*	-0.0153 (0.0054)
Motorcycle AADT	Dispersion	0.8197 (0.0528)
Total AADT	Intercept	-8.0110 (0.1180)
Total AADT	*b*	0.7225 (0.0353)
Total AADT	*c*	0.6232 (0.0146)
Total AADT	*d*	
Total AADT	*e*	-0.1730 (0.0315)
Total AADT	*f*	0.0150 (0.0021)
Total AADT	*g*	-0.0433 (0.0055)
Total AADT	Dispersion	0.7239 (0.0497)
Motorcycle and other AADT	Intercept	-7.4806 (0.1329)
Motorcycle and other AADT	*b*	0.7221 (0.0353)
Motorcycle and other AADT	*c*	0.1414 (0.0172)

Model Type	Parameter	MOTO
Motorcycle and other AADT	*d*	0.5026 (0.0205)
Motorcycle and other AADT	*e*	-0.1874 (0.0315)
Motorcycle and other AADT	*f*	0.0139 (0.0021)
Motorcycle and other AADT	*g*	-0.0438 (0.0055)
Motorcycle and other AADT	Dispersion	0.7120 (0.0493)

Note: Blank cell indicates parameter was not included in the model.

Table 43. Goodness-of-fit statistics for A1 models for Pennsylvania type 3 and 4 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	779	0.21	0.42	0.84
MOTO	Total AADT	779	0.20	0.46	0.74
MOTO	Motorcycle and other AADT	779	0.20	0.46	0.73

Table 44. CURE plot statistics for A1 models for Pennsylvania type 3 and 4 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)	Max CURE Deviation for WIDTH	CURE Deviation. for WIDTH (Percent)	Max CURE Deviation for AVGSHLDWID	CURE Deviation for AVGSHLDWID (Percent)
MOTO	Motor AADT	20.78	4	28.39	3	34.92	42	35.76	18
MOTO	Total AADT	33.93	7	28.63	3	34.81	11	41.93	41
MOTO	Motor and other AADT	20.04	0	29.42	4	31.87	4	41.24	32

Table 45. Calibration factors for A1 models for Pennsylvania type 3 and 4 sites.

Crash Type	Observed	Observed/Predicted
MOTO with AVGMOTO urban	420	1.07
MOTO with AVGMOTO rural	359	0.93
MOTO with AVGAADT urban	420	1.07
MOTO with AVGAADT rural	359	0.92
MOTO with AVGAADT and other urban	420	1.06
MOTO with AVGAADT and other rural	359	0.94

The goodness-of-fit results in table 43 indicate overall that MOTO model with total AADT performs slightly better than that with motorcycle AADT in terms of the modified R^2 and dispersion parameter. However, the CURE plot statistics for the key motorcycle AADT variable in table 45 indicate that all AADT predictor models were just as successful when applied to segments with the two levels of the indicator variables.

Virginia

The project team used Virginia data to validate the A1 models developed. The goal was to assess whether the motorcycle crash SPFs transferred well to a new jurisdiction and to assess whether the models using motorcycle AADT or total AADT transferred better (or worse). Table 46 through table 51 provide a comparison of goodness-of-fit statistics for the Florida and Pennsylvania A1 models applied to Virginia data.

Because the Florida and Pennsylvania A1 models require input variables that are unavailable in Virginia, mean values of these variables in the calibration data were used to reduce the models to segment length, AADT of interest, and, when necessary, a rural/urban indicator. Mean values were determined from the descriptive statistics by individual roadway type.

Table 46. Validation of A1 models for MOTO crashes.

State	Site Type	AADT Type in Model	Observed Crashes	MAD	Modified R^2	Dispersion
Florida	1	Total	18	0.25	0	29.52
Florida	1	Motorcycle	18	0.26	0	17.12
Florida	2	Total	90	0.15	0	3.09
Florida	2	Motorcycle	90	0.15	0.15	2.16
Florida	3	Total	799	0.11	0.04	4.16
Florida	3	Motorcycle	799	0.11	0.11	3.07
Florida	3	Other	799	0.11	0.06	3.65
Florida	4	Total	2,705	0.18	0.14	1.12
Florida	4	Motorcycle	2,705	0.18	0.23	1.03
Florida	4	Other	2,705	0.18	0.13	1.11
Pennsylvania	1 and 2	Total	108	0.16	0.02	5.00
Pennsylvania	1 and 2	Motorcycle	108	0.16	0.04	4.59
Pennsylvania	1 and 2	Other	108	0.16	0.03	5.17
Pennsylvania	3 and 4	Total	3,504	0.16	0.16	1.53
Pennsylvania	3 and 4	Motorcycle	3,504	0.16	0.20	1.40
Pennsylvania	3 and4	Other	3,504	0.16	0.17	1.47

Table 47. CURE statistics for the validation of A1 models for MOTO crashes.

State	Type	AADT Type	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)
Florida	1	Total	10.87	64	7.87	41
Florida	1	Motorcycle	13.57	76	5.08	31
Florida	2	Total	25.12	91	18.96	41
Florida	2	Motorcycle	8.84	0	20.20	46
Florida	3	Total	159.31	97	77.86	68
Florida	3	Motorcycle	72.69	73	68.06	74
Florida	3	Other	91.73	84	76.42	69
Florida	4	Total	362.81	99	186.12	96
Florida	4	Motorcycle	225.83	99	188.34	98
Florida	4	Other	277.37	100	207.65	98
Pennsylvania	1 and 2	Total	23.23	68	16.77	16
Pennsylvania	1 and 2	Motorcycle	19.13	33	18.67	22
Pennsylvania	1 and 2	Other	21.79	60	17.79	17
Pennsylvania	3 and 4	Total	96.99	28	190.65	94
Pennsylvania	3 and 4	Motorcycle	112.28	46	253.52	96
Pennsylvania	3 and 4	Other	77.42	27	206.63	95

The results for MOTO crashes were not clear on whether the Florida or Pennsylvania models calibrated better to the Virginia data. Whichever one was better depends on the specific measure and site type considered. What was consistent, however, is that the motorcycle AADT-based models showed a better goodness-of-fit as measured by the modified R^2 and dispersion. While it is not true that the models using motorcycle AADT always are preferred in terms of maximum CURE deviation and percent CURE deviation, this is mostly true. However, all of the models showed a significant amount of bias as measured by the percent CURE deviation.

Table 48. Validation of A1 models for MOTOSINGLE crashes in Florida.

Site Type	AADT Type in Model	Observed Crashes	MAD	Modified R^2	Dispersion
1	Total	16	0.23	0	30.27
1	Motorcycle	16	0.23	0	16.77
2	Total	50	0.09	0.11	2.84
2	Motorcycle	50	0.09	0.26	1.71
3	Total	410	0.06	0.05	6.12
3	Motorcycle	410	0.06	0.08	4.93
3	Other	410	0.06	0.07	5.38
4	Total	836	0.06	0.43	0.70
4	Motorcycle	836	0.06	0.44	0.68
4	Other	836	0.06	0.43	0.70

Table 49. CURE statistics for the validation of A1 models for MOTOSINGLE crashes in Florida.

Type	AADT Type	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)
1	Total	12.51	77	4.60	28
1	Motorcycle	9.29	56	6.66	39
2	Total	14.53	88	11.60	30
2	Motorcycle	7.78	8	12.17	32
3	Total	93.62	95	65.26	27
3	Motorcycle	42.98	73	62.86	25
3	Other	47.21	82	63.58	25
4	Total	56.50	78	39.07	16
4	Motorcycle	26.23	11	34.83	10
4	Other	65.58	87	34.21	12

For MOTOSINGLE crashes, the project team did not calibrate models from Pennsylvania. The Florida models using motorcycle AADT showed better goodness-of-fit statistics in general. These models, however, showed significant bias as measured by the percent CURE deviation.

Table 50. Validation of A1 models for MOTOMULTI crashes.

State	Site Type	AADT Type in Model	Observed Crashes	MAD	Modified R^2	Dispersion
Florida	2	Total	40	0.07	0	3.83
Florida	2	Motorcycle	40	0.07	0.01	3.27
Florida	3	Total	388	0.06	0.03	3.30
Florida	3	Motorcycle	388	0.06	0.12	2.48
Florida	3	Other	388	0.06	0.05	3.01
Florida	4	Total	1,865	0.13	0	1.61
Florida	4	Motorcycle	1,865	0.13	0.10	1.43
Florida	4	Other	1,865	0.13	0	1.60
Pennsylvania	1 and 2	Total	42	0.07	0.12	2.69
Pennsylvania	1 and 2	Motorcycle	42	0.07	0.18	2.20

Table 51. CURE statistics for the validation of A1 models for MOTOMULTI crashes in Pennsylvania.

Site Type	AADT Type	Maximum CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Maximum CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)
2	Total	11.29	48	7.91	18
2	Motorcycle	6.40	29	8.55	26
3	Total	58.27	82	54.77	78
3	Motorcycle	28.23	26	62.30	86
3	Other	42.09	70	56.82	83
4	Total	332.44	98	191.56	97
4	Motorcycle	222.02	99	187.75	97
4	Other	236.70	97	210.38	98
1/2	Total	7.98	29	4.43	5
1/2	Motorcycle	4.67	7	7.14	9

The results for MOTOMULTI crashes indicate that the Pennsylvania models may transfer better than the Florida models to the Virginia data. The motorcycle AADT-based models in general showed better goodness-of-fit measures. However, almost all of the models showed a significant amount of bias as measured by the percent CURE deviation.

Assessment of A1 Models for Predictions at High Crash Locations

The project team conducted an assessment of how well the A1 models predict motorcycle crashes for high-crash locations. Such locations are typically of interest in treatment applications that form the basis for future CMF development. To do this, sites of a specific type were first ranked in descending order of the crash counts per mile in one period (denoted as a before period). EB estimates were then obtained for the top-ranked sites based on the calibrated SPFs using motorcycle AADT and total AADT and crash counts in that period. The sum of these EB estimates for the top-ranked sites were then compared to the sum of the crash counts for these locations in a subsequent period (denoted as an after period) after adjusting the EB estimates for the difference in years between the two periods. For Florida, the top 10 percent of sites were ranked based on a 3-year before period from 2008–2010; the after period was a 2-year period from 2011 to 2012.

For Pennsylvania, sites with at least one motorcycle crash, which constituted less than 10 percent of all sites, were top ranked based on a 3-year before period from 2009–2011; the after period was a 2-year period from 2012 to 2013. Table 52 shows the results of this assessment.

Table 52. Assessment of type 1 models for predictions at high crash locations.

State	Site Type	Number of Top-Ranked Sites	Motorcycle Crash Count (2-Year Before Period)	Motorcycle Crash Count (2-Year After Period)	EB Expected Crashes in 2-Year After Period Using Model with Motorcycle AADT	EB Expected Crashes in 2-Year After Period Using Model with Total AADT	MAD for Comparison to after Period Crashes with EB Estimate from Model with Motorcycle AADT	MAD for Comparison to after Period Crashes with EB Estimate from Model with Total AADT
Florida	1	48	45.3	13	10.28	11.13	0.363	0.364
Florida	2	96	136.7	56	43.45	47.44	0.568	0.575
Florida	3	325	311.3	109	79.43	107.99	0.408	0.433
Florida	4	872	1,584.7	926	650.37	675.17	0.849	0.852
Pennsylvania	1 and 2	434	327.3	70	26.3	29.7	0.025	0.028
Pennsylvania	3 and 4	4,348	3,126.0	632	423.5	456.6	0.212	0.215

The results in table 52 show that, although MAD tended to be slightly higher for the total AADT models, the EB estimates from the total AADT models of crashes at the top-ranked sites were marginally closer to the actual counts than those based on the motorcycle AADT models.

It should be noted in passing that before period counts (adjusted for a 2-year period) were much higher than the 2-year after period counts because of a substantial regression to the mean effect. The EB estimates were much closer, but precise matching should not be expected because trends in crash occurrence and traffic volume differences between the before and after periods were not considered as they would be in a rigorous EB before-after study. In the case of type 4 Florida sites and the Pennsylvania sites, an additional factor is that the top-ranked sites are likely heavily biased towards the ranges of variables such as traffic volume and segment length where the SPF may be underpredicting. This result would indicate that there is room for model improvement but should not be interpreted as saying these models and methods should not be applied for site-based analyses.

AVENUE A MODEL TYPE A2

The purpose of model type A2 is to develop a relationship between motorcycle crash frequency and total crash frequency. The project team developed models for both motorcycle crashes and total crashes using traffic volumes for motorcycles and all vehicles, respectively. Other geometric variables from table 3 and table 9 were not included in these models because the goal is for these models to be transferable to any jurisdiction.

The goal was to infer a relationship between the models for the two crash types. This relationship can then be applied to the model for total crashes for another State to infer a model for motorcycle crashes for that State. In turn, that model can be used in the evaluation of retrospective and prospective before-after evaluations of the effects of infrastructure countermeasures on motorcycle crashes.

Because the goal is to develop models that may be transferable between jurisdictions, only length and traffic volume variables were included since other jurisdictions may not have the same data available.

The modeling is a three-step process. The first is to develop a model for total crashes, as shown in figure 29.

$$TOT/year = exp^{Intercept}\ LENGTH^{b} AVGAADT^{c}$$

Figure 29. Equation. Type A2 total crashes model.

The predicted values from this model are given the name PREDTOT. The second step is to develop a model for the motorcycle crashes of interest. For example, for MOTO crashes, this would be the model shown in figure 30.

$$MOTO/year = exp^{Intercept}\ LENGTH^{b} AVGMOTO^{c}$$

Figure 30. Equation. Type A2 motorcycle crashes model.

The predicted values from this model are given the name PREDMOTO. The final step is to develop a model that predicts the value of PREDMOTO/year. In developing this model, for each site, the value of PREDMOTO is the dependent variable, and the value of PREDTOT is considered as one of the possible explanatory variables in the model.

The project team pursued the A2 models using both data from Florida and Pennsylvania and using data from Virginia for validation. The remainder of this section reports on the results of model development and validation.

Florida

An exponential model and a linear model were attempted. The project team found that although the two models performed similarly over most sites, the exponential model overpredicts crashes significantly when the total crash frequency is high. For this reason, the project team adopted the linear model form as shown in figure 31 through figure 33. The error distribution for these models was assumed as a gamma distribution.

$$PREDMOTO\ Crashes/year = Intercept + b * PREDTOT + c * \frac{AADT}{10{,}000}$$

Figure 31. Equation. Florida type A2 motorcycle crash model.

$$PREDMOTOSINGLE\ Crashes/year = Intercept + b * PREDTOT + c * \frac{AADT}{10{,}000}$$

Figure 32. Equation. Florida type A2 motorcycle single-vehicle crash model.

$$PREDMOTOMULTI\ Crashes/year = Intercept + b * PREDTOT + c * \frac{AADT}{10{,}000}$$

Figure 33. Equation. Florida type A2 motorcycle multi-vehicle crash model.

Florida Type 1—Rural Freeways

Table 53 presents the parameter estimates for the type 1 sites. Attempts to estimate a full set of corresponding models for MOTOMULTI were unsuccessful. There are three models for each category of motorcycle crashes: the model for total crashes, the model for the motorcycle crashes of interest, and the model that predicts the value of the motorcycle crash model based on the predicted value of total crashes and total AADT. Table 54 and table 55 provide goodness-of-fit statistics for the A2 model predictions and compare them to the predictions from the A1 models using motorcycle AADT as an explanatory variable. The A2 model predictions use the third model (i.e., the model that predicts the estimate from a motorcycle crash SPF based on the prediction from a total crash SPF).

Table 53. A2 models for Florida type 1 sites.

Model	Parameter	MOTO	MOTOSINGLE
Total crashes	Intercept	-10.6602 (0.7239)	-10.6602 (0.7239)
Total crashes	*b*	0.7995 (0.0300)	0.7995 (0.0300)
Total crashes	*c*	1.1700 (0.0697)	1.1700 (0.0697)
Total crashes	Dispersion	0.4541 (0.0403)	0.4541 (0.0403)
Motorcycle crashes	Intercept	-5.5368 (0.6033)	-5.1341 (0.7741)
Motorcycle crashes	*b*	0.8239 (0.0801)	0.8939 (0.1102)
Motorcycle crashes	*c*	0.6622 (0.1274)	0.4755 (0.1663)
Motorcycle crashes	Dispersion	0.4353 (0.2160)	1.1346 (0.4327)
Predicted motorcycle crashes	Intercept	0.0555 (0.0037)	0.0436 (0.0020)
Predicted motorcycle crashes	*b*	0.0187 (0.0004)	0.0125 (0.0003)
Predicted motorcycle crashes	*c*	-0.0217 (0.0016)	-0.0191 (0.0007)

Table 54. Goodness-of-fit statistics for A2 models for Florida type 1 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	174	0.43	0.65	0.44
MOTO	A2 model	174	0.41	0.78	0.29
MOTOSINGLE	Motorcycle AADT	111	0.32	0.34	1.14
MOTOSINGLE	A2 model	111	0.31	0.46	0.93

Table 55. CURE plot statistics for A2 models for Florida type 1 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)
MOTO	Motorcycle AADT	15.83	11	7.42	3
MOTO	A2 model	13.08	3	7.31	3
MOTOSINGLE	Motorcycle AADT	10.42	5	7.80	3
MOTOSINGLE	A2 model	10.96	6	9.17	5

The results indicate (especially for the modified R^2 and dispersion parameter measures) that both A2 models, which predict motorcycle crashes from predictions from a total crash versus total AADT model, outperformed the corresponding models that predict motorcycle crashes from motorcycle AADT.

Florida Type 2—Urban Freeways

Table 56 presents the parameter estimates for the type 2 sites. There are three models for each category of motorcycle crashes: the model for total crashes, the model for the motorcycle crashes of interest, and the model that predicts the value of the motorcycle crash model based on the predicted value of total crashes and total AADT. Table 57 and table 58 provide goodness-of-fit statistics for the A2 model predictions and compare them to the predictions from the A1 models using motorcycle AADT as an explanatory variable. The A2 model predictions use the third model (i.e., the model that predicts the estimate from a motorcycle crash SPF based on the prediction from a total crash SPF).

Table 56. A2 models for Florida type 2 sites.

Model	Parameter	MOTO	MOTOSINGLE	MOTOMULTI
Total crashes	Intercept	-12.2924 (0.4305)	-12.2924 (0.4305)	-12.2924 (0.4305)
Total crashes	*b*	0.7559 (0.0268)	0.7559 (0.0268)	0.7559 (0.0268)
Total crashes	*c*	1.3304 (0.0384)	1.3304 (0.0384)	1.3304 (0.0384)
Total crashes	Dispersion	0.5179 (0.0264)	0.5179 (0.0264)	0.5179 (0.0264)
Motorcycle crashes	Intercept	-3.1295 (0.2465)	-3.3297 (0.2688)	-3.8930 (0.2930)
Motorcycle crashes	*B*	0.7184 (0.0488)	0.7283 (0.0575)	0.6943 (0.0591)
Motorcycle crashes	*C*	0.3553 (0.0433)	0.2371 (0.0469)	0.3910 (0.0509)
Motorcycle crashes	Dispersion	0.6978 (0.0951)	0.3559 (0.1347)	0.8087 (0.1392)
Predicted motorcycle crashes	Intercept	0.4952 (0.0158)	0.2362 (0.0053)	0.2762 (0.0101)
Predicted motorcycle crashes	*b*	0.0196 (0.0007)	0.0079 (0.0002)	0.0111 (0.0004)
Predicted motorcycle crashes	*c*	-0.0520 (0.0026)	-0.0250 (0.0008)	-0.0269 (0.0017)

Table 57. Goodness-of-fit statistics for A2 models for Florida type 2 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	921	0.86	0.56	0.43
MOTO	A2 model	921	0.84	0.65	0.35
MOTOSINGLE	Motorcycle AADT	385	0.47	0.66	0.26
MOTOSINGLE	A2 model	385	0.47	0.75	0.18
MOTOMULTI	Motorcycle AADT	536	0.58	0.59	0.38
MOTOMULTI	A2 model	536	0.58	0.63	0.34

Table 58. CURE plot statistics for A2 models for Florida type 2 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)
MOTO	Motorcycle AADT	54.69	35	22.11	0
MOTO	A2 model	90.73	80	24.3	5
MOTOSINGLE	Motorcycle AADT	26.21	16	11.53	0
MOTOSINGLE	A2 model	24.26	11	17.94	0
MOTOMULTI	Motorcycle AADT	37.34	41	18.12	16
MOTOMULTI	A2 model	71.80	83	20.62	31

The results indicate, especially for the modified R^2 and dispersion parameter measures, that the three A2 models, which predict motorcycle crashes from predictions from a total-crash versus total AADT model, outperformed corresponding models that predict motorcycle crashes from motorcycle AADT. However, the CURE plot statistics for the key AADT variable, AVGMOTO, indicate the opposite for the MOTO and MOTOMULTI models.

Florida Type 3—Rural Arterials

Table 59 presents the parameter estimates for the type 3 sites. There are three models for each category of motorcycle crashes: the model for total crashes, the model for the motorcycle crashes of interest, and the model that predicts the value of the motorcycle crash model based on the predicted value of total crashes and total AADT.

Table 60 and table 61 provide goodness-of-fit statistics for the A2 model predictions and compare them to the predictions from the A1 models using motorcycle AADT as an explanatory variable. The A2 model predictions use the third model (i.e., the model that predicts the estimate from a motorcycle crash SPF based on the prediction from a total crash SPF).

Table 59. A2 models for Florida type 3 sites.

Model	Parameter	MOTO (SE)	MOTOSINGLE (SE)	MOTOMULTI (SE)
Total crashes	Intercept	-6.9586 (0.2236)	-6.9586 (0.2236)	-6.9586 (0.2236)
Total crashes	*b*	0.7145 (0.0163)	0.7145 (0.0163)	0.7145 (0.0163)
Total crashes	*c*	0.8411 (0.0256)	0.8411 (0.0256)	0.8411 (0.0256)
Total crashes	Dispersion	0.8672 (0.0312)	0.8672 (0.0312)	0.8672 (0.0312)
Motorcycle crashes	Intercept	-4.6303 (0.1684)	-5.1015 (0.2302)	-5.5552 (0.2199)
Motorcycle crashes	*B*	0.6780 (0.0315)	0.7210 (0.0436)	0.6494 (0.0397)
Motorcycle crashes	*C*	0.5671 (0.0431)	0.4712 (0.0591)	0.6592 (0.0553)
Motorcycle crashes	Dispersion	0.8912 (0.1194)	1.0532 (0.2348)	1.0096 (0.1892)
Predicted motorcycle crashes	Intercept	0.0499 (0.0018)	0.0235 (0.0007)	0.0232 (0.0011)
Predicted motorcycle crashes	*b*	0.0534 (0.0007)	0.0239 (0.0003)	0.0294 (0.0005)
Predicted motorcycle crashes	*c*	-0.0720 (0.0028)	-0.0404 (0.0009)	-0.0267 (0.0020)

Table 60. Goodness-of-fit statistics for A2 models for Florida type 3 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	1,031	0.40	0.44	0.77
MOTO	A2 model	1,031	0.41	0.39	0.89
MOTOSINGLE	Motorcycle AADT	449	0.21	0.43	0.91
MOTOSINGLE	A2 model	449	0.21	0.35	1.10
MOTOMULTI	Motorcycle AADT	582	0.26	0.37	0.85
MOTOMULTI	A2 model	582	0.27	0.37	0.99

Table 61. CURE plot statistics for A2 models for Florida type 3 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)
MOTO	Motorcycle AADT	37.40	0	21.32	0
MOTO	A2 model	106.65	76	27.08	0
MOTOSINGLE	Motorcycle AADT	24.33	1	8.21	0
MOTOSINGLE	A2 model	41.21	25	19.82	0
MOTOMULTI	Motorcycle AADT	19.71	1	20.19	6
MOTOMULTI	A2 model	69.92	89	25.57	7

All of the results consistently indicate that the three models that predict motorcycle crashes from motorcycle AADT outperformed the corresponding A2 models, which predict motorcycle crashes from predictions from a total-crash vs total AADT model. This is somewhat contrary to the findings for type 1 and type 2 sites.

Florida Type 4Urban Arterials

Table 62 presents the parameter estimates for the type 4 sites. There are three models for each category of motorcycle crashes: the model for total crashes, the model for the motorcycle crashes of interest, and the model that predicts the value of the motorcycle crash model based on the predicted value of total crashes and total AADT. Table 63 and table 64 provide goodness-of-fit statistics for the A2 model predictions and compare them to the predictions from the A1 models using motorcycle AADT as an explanatory variable. The A2 model predictions use the third model (i.e., the model that predicts the estimate from a motorcycle crash SPF based on the prediction from a total crash SPF).

Table 62. A2 models for Florida type 4 sites.

Model	Parameter	MOTO (SE)	MOTOSINGLE (SE)	MOTOMULTI (SE)
Total crashes	Intercept	-11.6281 (0.2286)	-11.6281 (0.2286)	-11.6281 (0.2286)
Total crashes	*b*	0.5894 (0.0204)	0.5894 (0.0204)	0.5894 (0.0204)
Total crashes	*c*	1.3751 (0.0227)	1.3751 (0.0227)	1.3751 (0.0227)
Total crashes	Dispersion	2.0547 (0.0336)	2.0547 (0.0336)	2.0547 (0.0336)
Motorcycle crashes	Intercept	-4.5339 (0.1272)	-5.2472 (0.1857)	-5.0429 (0.1427)
Motorcycle crashes	*b*	0.7419 (0.0218)	0.7375 (0.0301)	0.7488 (0.0241)
Motorcycle crashes	*c*	0.7469 (0.0248)	0.6129 (0.0360)	0.7918 (0.0278)
Motorcycle crashes	Dispersion	1.3244 (0.0465)	1.2330 (0.0093)	1.4792 (0.0574)
Predicted motorcycle crashes	Intercept	0.5031 (0.0067)	0.1470 (0.0015)	0.3572 (0.0051)
Predicted motorcycle crashes	*b*	0.0479 (0.0006)	0.0114 (0.0001)	0.0367 (0.0005)
Predicted motorcycle crashes	*c*	-0.2972 (0.0054)	-0.0791 (0.0012)	-0.2204 (0.0043)

Table 63. Goodness-of-fit statistics for A2 models for Florida type 4 sites.

Crash Type	Exposure Measure	Total Observed	MAD	Modified R^2	Dispersion
MOTO	Motorcycle AADT	9,539	1.11	0.28	1.19
MOTO	A2 model	9,539	1.10	0.30	1.20
MOTOSINGLE	Motorcycle AADT	2,341	0.38	0.31	1.12
MOTOSINGLE	A2 model	2,341	0.38	0.32	1.16
MOTOMULTI	Motorcycle AADT	7,198	0.91	0.25	1.32
MOTOMULTI	A2 model	7,198	0.90	0.27	1.32

Table 64. CURE plot statistics for A2 models for Florida type 4 sites.

Crash Type	Exposure Measure	Max CURE Deviation for AVGMOTO	CURE Deviation for AVGMOTO (Percent)	Max CURE Deviation for LENGTH	CURE Deviation for LENGTH (Percent)
MOTO	Motorcycle AADT	242.23	14	120.12	0
MOTO	A2 model	288.17	45	644.54	95
MOTOSINGLE	Motorcycle AADT	52.04	2	51.84	0
MOTOSINGLE	A2 model	123.19	76	105.03	21
MOTOMULTI	Motorcycle AADT	178.29	8	120.03	1
MOTOMULTI	A2 model	528.91	96	223.50	34

The results indicate that the performance of the three A2 models, which predict motorcycle crashes from predictions from a total-crash versus total AADT model, was similar to that of the corresponding models that predict motorcycle crashes from motorcycle AADT. However, the CURE plot statistics for the key AADT variable, AVGMOTO, indicate that models that predict motorcycle crashes from motorcycle AADT were better in terms of the CURE measures than the corresponding A2 models. These results for urban arterials, overall, were consistent with findings for rural arterials.

Pennsylvania

The linear model form that worked for the Florida dataset failed to converge for the Pennsylvania dataset. While the exponential model form did converge for the Pennsylvania dataset, the sum of the predictions was significantly higher than the observed number of crashes. Similar to the Florida exponential model, this overprediction occurred for sites where the total crash frequency was high. For this reason, the project team concluded that the A2 models using the Pennsylvania data were not successful.

Virginia

To validate the A2 models developed, the project team split the data for Virginia into calibration and validation datasets for rural and urban arterials. The data for freeways did not include enough crashes for validation of the A2 models. The project team used the calibration datasets to develop the total crash SPF required for application of the A2 models. The project team also used the calibration data to develop motorcycle crash SPFs using total AADT to compare to the A2 models.

The project team used the validation datasets to evaluate the performance of the A2 models and compare that to the performance of a total AADT SPF for motorcycle crashes. To perform this validation, the project team calibrated the predictions from both the A2 model process and the Virginia motorcycle SPFs to the calibration data and goodness-of-fit statistics determined. Table 65 provides the summary data for split validation and calibration datasets for rural and urban arterials.

Table 65. Length and crash frequency for Virginia model type A2 validation analysis.

Type	LENGTH (mi)	TOT	MOTO	MOTOSINGLE	MOTOMULTI
3A calibration	2,195.4	1,6653	373	183	189
3B validation	2,246.1	1,6672	426	227	199
4A calibration	1,665.0	6,4651	1,414	432	981
4B validation	1,658.4	6,0662	1,291	404	884

1 mi = 1.6 km.

Table 66 provides summary data for motorcycle AADT, total AADT, and number of lanes. Comparisons between calibration and validation datasets shows consistency in minimum, maximum, and average values.

Table 66. Summary statistics for Virginia model type A2 validation analysis.

Type	Statistic	AVGMOTO	AVGAADT	NOLANES
3A	No. Segments	6,287	6,287	6,287
3A	MIN	1.0	305.0	1.0
3A	MAX	245.0	58,715.0	7.0
3A	MEAN	34.8	8,187.0	2.8
3A	STD	31.1	6,782.6	1.0
3B	No. Segments	6,287	6,287	6,287
3B	MIN	1.0	305.0	1.0
3B	MAX	245.0	58,715.0	6.0
3B	MEAN	34.6	8,106.6	2.8
3B	STD	30.3	6,849.6	1.0
4A	No. Segments	12,695	12,695	12,695
4A	MIN	1.0	131.0	1.0
4A	MAX	509.0	130,077.0	9.0
4A	MEAN	54.8	17,705.1	3.3
4A	STD	54.7	14,380.2	1.3
4B	No. Segments	12,695	12,695	12,695
4B	MIN	1.0	270.0	1.0
4B	MAX	509.0	130,077.0	9.0
4B	MEAN	53.8	17,425.2	3.3
4B	STD	55.7	14,280.9	1.3

For rural arterials, the project team developed SPFs for total, motorcycle, and multi-vehicle motorcycle crashes. Total AADT and segment length were used as explanatory variables for all SPFs. Table 67 shows the parameter estimates for the total crash SPF, the motorcycle crash SPF, and the A2 model applied. Using the Virginia data, an SPF for predicting single-vehicle motorcycle crashes did not converge.

Table 67. SPFs for validating model types A2 for Virginia type 3 (rural arterial) sites.

Model	Parameter	MOTO (SE)	MOTOMULTI (SE)
Total crash SPF	Intercept	-5.3321 (0.1575)	-5.3321 (0.1575)
Total crash SPF	*b*	0.6160 (0.0122)	0.6160 (0.0122)
Total crash SPF	*c*	0.6224 (0.0181)	0.6224 (0.0181)
Total crash SPF	Dispersion	0.7529 (0.0253)	0.7529 (0.0253)
Motorcycle crash SPF using total AADT	Intercept	-5.3223 (0.5848)	-8.1803 (0.8208)
Motorcycle crash SPF using total AADT	*b*	0.6302 (0.0519)	0.5152 (0.0658)
Motorcycle crash SPF using total AADT	*c*	0.1957 (0.0679)	0.4306 (0.0927)
Motorcycle crash SPF using total AADT	Dispersion	3.0688 (0.5901)	1.4954 (0.7830)
Florida A2 models	Intercept	0.0499 (0.0018)	0.0232 (0.0011)
Florida A2 models	*b*	0.0534 (0.0007)	0.0294 (0.0005)
Florida A2 models	*c*	-0.0720 (0.0028)	-0.0267 (0.0020)

For urban arterials, the project team developed SPFs for total, motorcycle, single-vehicle motorcycle, and multi-vehicle motorcycle crashes in Virginia. Total AADT and segment length were used as explanatory variables for all SPFs. Table 68 shows the parameter estimates for the total crash SPF, the motorcycle crash SPF, and the A2 model applied.

Table 68. SPFs for validating A2 models for Virginia type 4 (urban arterial) sites.

Model	Parameter	MOTO (SE)	MOTOSINGLE (SE)	MOTOMULTI (SE)
Total crash SPF	Intercept	-6.3949 (0.1343)	-6.3949 (0.1343)	-6.3949 (0.1343)
Total crash SPF	*b*	0.6202 (0.0109)	0.6202 (0.0109)	0.6202 (0.0109)
Total crash SPF	*c*	0.7988 (0.0137)	0.7988 (0.0137)	0.7988 (0.0137)
Total crash SPF	Dispersion	1.0363 (0.0183)	1.0363 (0.0183)	1.0363 (0.0183)
Motorcycle crash SPF using total AADT	Intercept	-7.9779 (0.3824)	-8.0193 (0.6428)	-8.8642 (0.4583)
Motorcycle crash SPF using total AADT	*b*	0.6594 (0.0309)	0.8446 (0.0534)	0.5772 (0.0364)
Motorcycle crash SPF using total AADT	*c*	0.5808 (0.0384)	0.4985 (0.0650)	0.6176 (0.0459)
Motorcycle crash SPF using total AADT	Dispersion	0.9252 (0.1389)	0.9436 (0.3897)	1.1309 (0.2078)
Florida A2 models	Intercept	0.5031 (0.0067)	0.1470 (0.0015)	0.3572 (0.0051)
Florida A2 models	*b*	0.0479 (0.0006)	0.0114 (0.0001)	0.0367 (0.0005)
Florida A2 models	*c*	-0.2972 (0.0054)	-0.0791 (0.0012)	-0.2204 (0.0043)

When applying the A2 models to the validation data, the project team observed that the predictions produced some negative values for expected crashes when the total crash SPF prediction was low. This illogical result of negative crash prediction shows that the A2 models do not transfer well to another jurisdiction if the expected total crash frequency is markedly different. The project team measured the success of the A2 in how well they can predict crashes for a new jurisdiction that does not have motorcycle AADT estimates but can develop an SPF for total crashes using total AADT. For this reason, the results concluded that the A2 modeling was not successful.

The possibility of negative crash predictions was made possible by the linear model form of the A2 models, shown in figure 34, where the parameter *c* was estimated to be negative.

$$PREDMOTO\ Crashes/year = Intercept + b * PREDTOT + c * \frac{AADT}{10{,}000}$$

Figure 34. Equation. A2 linear model form.

When developing the A2 models, other model forms that would not allow such negative predictions were not successful in that they significantly overpredicted crashes at high AADTs.

AVENUE A MODEL TYPE A3

The purpose of investigating model type 3 was to attempt the development of models to estimate motorcycle traffic volumes based on roadway characteristics and other variables that may influence motorcycle trip generation. If successful, such models could be used to estimate motorcycle volumes in similar jurisdictions, and these volumes could be used for any study design where motorcycle volumes are desired.

The A3 models were estimated using simple linear regression, which assumes a normal distribution for the error term. The form was adopted from and is consistent with the forms used for trip generation models in the transportation planning field. Figure 35 shows an example of an estimated model form.

$$MOTOAADT = Intercept + b * LANES + c * SPDLIMT + d * POP060210 + e * INC110213 + f * LFE305213 + g * SEX255213$$

Figure 35. Equation. Type A3 model form example.

Where:

MOTOAADT = The motorcycle AADT on a roadway segment.
LANES = 0 if the roadway has two lanes and 1 if the roadway has more than two lanes.
SPDLIMT = The posted speed limit in mi/h.
POP060210 = The population per mi^2.
INC110213 = The median household income.
LFE305213 = The mean travel time to work in min for workers age 16+.
SEX255213 = The percentage of the population that is female.

The project team attempted models for site types 1 through 4 in both Florida and Pennsylvania. For both datasets, the project team could not consider the models developed successful. Although the models did include parameter estimates that were statistically significant at the 95-percent confidence limit, the explanatory power of the models was very low and only marginally better than simply assuming the average level of motorcycle AADT for all sites. This is likely because variables that are most related to motorcycle AADT may not have been available for inclusion.

The lack of explanatory power provided by the models was evident from examining the R^2 coefficient. The R^2 coefficient is a statistical measure of how much variation in the data is being explained by the model. In equation form, this is expressed as shown in figure 36 and is always between 0 and 100 percent.

$$R^2 = \frac{Explained\ variation}{Total\ variation}$$

Figure 36. Equation. R^2 coefficient.

The models developed for Florida and Pennsylvania showed R^2 values typically between 5 and 15 percent, indicating that they are not explaining much of the variation in motorcycle AADT between road segments and are thus not very useful. For this reason, the project team did not consider the A3 modeling a success.

AVENUE B MODEL TYPE B1

The model type B1 analyses were conducted using the data for types 1 (rural freeway) and 3 (rural arterial) in Florida. The approach used for simulating data is described in chapter 4 in the subsection titled Avenue B Databases. The open-source software R was used for simulating the data and model estimation. Several parameters were changed in running several simulations to examine their effects on the results. These included the number of treatment sites, the assumed CMF, and how the project team selected the treatment sites.

Florida Type 1

Table 69 shows the results for simulation 1 using the Florida type 1 data. In simulation 1, the treated sites were selected as those with the highest crash frequency per mi in the before period. For simulation 1, there were 100 treatment sites and 382 reference sites. The assumed CMF was 0.70, and there were 3 years of before and after data. There were 10 separate trials indicated by trial ID in the table. For each trial, the project team simulated the observed number of crashes for all sites so that the data were different for each trial. For each trial, the table provides the estimated CMF and its SE using the SPF with motorcycle AADT as an explanatory variable, the estimated CMF and its SE using the SPF with total AADT as an explanatory variable, and the absolute value of the difference between the two estimated CMFs and the SE of this difference. The average values are provided in the last row.

Table 69. Model type B2 results for simulation 1 using type 1 Florida data.

Trial	CMF MOTO	SE MOTO	CMF AADT	SE AADT	Difference	SE Difference
1	1.13	0.22	1.13	0.22	0.00	0.31
2	1.28	0.24	1.28	0.24	0.00	0.34
3	1.03	0.20	1.09	0.21	0.06	0.29
4	0.78	0.15	0.85	0.17	0.07	0.23
5	0.82	0.17	0.88	0.18	0.06	0.25
6	0.93	0.19	0.82	0.17	0.11	0.25
7	1.76	0.30	1.74	0.30	0.02	0.42
8	0.57	0.13	0.59	0.14	0.02	0.19
9	1.19	0.21	1.27	0.23	0.08	0.31
10	0.57	0.14	0.54	0.13	0.03	0.19
Average	1.01	0.20	1.02	0.20	0.05	0.28

The results from simulation 1 showed that the estimated CMF could vary a lot between trials, and the SEs of the CMF estimates were high. Very few of the CMF estimates were statistically different from a value of 1.0.

A value of 1.0 for a CMF indicated that there was no effect on crashes. This illustrates the difficulty in estimating accurate CMFs using motorcycle crashes, which are rare. While this is illustrative of this fact, the prime interest is to investigate whether the CMF estimates differ greatly depending on whether the SPFs using motorcycle AADT or total AADT are applied in the EB before-after study approach. For this comparison, the results showed little difference. The average CMFs using the motorcycle and total AADT respectively were 1.01 and 1.02, and the average difference across all trials was 0.05.

The estimated CMFs that tended to be higher than 0.70 indicated that regression-to-the-mean was being overcorrected. Because the treatment sites were selected based on high crash rate, this left low crash rate sites for the reference sites, and the SPFs required predicted very few crashes since they are based on these remaining low crash sites.

In simulation 2, the project team used 200 sites as treatment sites and only 282 as reference sites. The CMF and number of years before and after stayed the same. The results, shown in table 70, were very similar to those in simulation 1, although the SEs of estimates were slightly smaller due to the larger sample size of treated sites.

Table 70. B2 results for simulation 2 using type 1 Florida data.

Trial	CMF MOTO	SE MOTO	CMF AADT	SE AADT	Difference	SE Difference
11	1.54	0.24	1.62	0.25	0.08	0.35
12	1.07	0.19	1.07	0.19	0.00	0.27
13	0.87	0.15	0.90	0.16	0.03	0.22
14	0.84	0.15	0.79	0.14	0.05	0.21
15	1.21	0.20	1.29	0.22	0.08	0.30
16	0.70	0.15	0.67	0.14	0.03	0.21
17	1.04	0.19	0.96	0.17	0.08	0.25
18	1.01	0.19	1.00	0.19	0.01	0.27
19	0.76	0.14	0.79	0.15	0.03	0.21
20	0.86	0.16	0.82	0.16	0.04	0.23
Average	0.99	0.18	0.99	0.18	0.04	0.25

In simulation 3, the number of sites and assumed CMF was identical to simulation 2, but now the treatment sites were selected by a random variable and not by taking the highest crash rate sites in the before period. This random selection would minimize any regression to the mean in the data. For these results, shown in table 71, there were near-identical estimates for the CMF between using the motorcycle or total AADT model. The average value of estimated CMFs was 0.69 for both the motorcycle and total AADT SPF approaches. However, significant variation in the CMF estimates for between trials remained.

Table 71. B2 results for simulation 3 using type 1 Florida.

Trial	CMF MOTO	SE MOTO	CMF AADT	SE AADT	Difference	SE Difference
21	0.93	0.16	0.93	0.16	0.00	0.23
22	0.55	0.11	0.55	0.11	0.00	0.16
23	0.58	0.12	0.59	0.12	0.01	0.17
24	0.58	0.11	0.58	0.11	0.00	0.16
25	0.61	0.11	0.59	0.11	0.02	0.16
26	0.53	0.11	0.53	0.11	0.00	0.16
27	0.62	0.12	0.64	0.12	0.02	0.17
28	0.80	0.14	0.78	0.14	0.02	0.20
29	0.81	0.13	0.80	0.14	0.01	0.19
30	0.93	0.16	0.93	0.16	0.00	0.23
Average	0.69	0.13	0.69	0.13	0.01	0.18

In simulation 4, the treatment sites were selected using a random variable only for sites that had one or more motorcycle crashes in the before period. The project team also reduced the number of treatment sites to 50. The results are shown in table 72. Again, the results showed that the estimates using the two SPFs were close, with the average difference being 0.05. The average values of the estimated CMFs were also close to 0.70.

Table 72. Results for simulation 4 using type 1 Florida data.

Trial	CMF MOTO	SE MOTO	CMF AADT	SE AADT	Difference	SE Difference
31	0.56	0.16	0.61	0.18	0.05	0.24
32	0.88	0.20	0.94	0.21	0.06	0.29
33	0.86	0.19	0.88	0.20	0.02	0.28
34	0.78	0.18	0.82	0.19	0.04	0.26
35	0.75	0.18	0.79	0.19	0.04	0.26
36	0.49	0.15	0.49	0.16	0.00	0.22
37	0.68	0.17	0.72	0.18	0.04	0.25
38	0.52	0.16	0.54	0.17	0.02	0.23
39	0.89	0.22	0.97	0.24	0.08	0.33
40	0.98	0.25	0.88	0.23	0.10	0.34
Average	0.74	0.19	0.76	0.20	0.05	0.27

Florida Type 3

Table 73 shows the results for simulation 1 using the Florida type 3 data. In simulation 1, the project team selected the treated sites using a random number for sites that had one or more before period crashes. There were 200 treatment sites and 2,997 reference sites. The CMF was 0.90 with 3 years before and 3 years after.

Similar to the results for the type 1 sites, the results from simulation 1 showed that the estimated CMF could vary a lot between trials. However, the differences for each trial between the estimates of the CMF using the SPF with motorcycle AADT versus total AADT were small. The average CMFs using the motorcycle and total AADT were 0.78 and 0.73, respectively, and the average difference across all trials was 0.05. The CMF estimates tended to be low, indicating that the correction for regression to the mean was not large enough.

Table 73. Results for simulation 1 using type 3 Florida data.

Trial	CMF MOTO	SE MOTO	CMF AADT	SE AADT	Difference	SE Difference
1	0.67	0.09	0.64	0.08	0.03	0.12
2	0.78	0.11	0.72	0.10	0.06	0.15
3	0.55	0.08	0.52	0.08	0.03	0.11
4	0.86	0.11	0.82	0.10	0.04	0.15
5	0.77	0.10	0.72	0.09	0.05	0.13
6	0.79	0.10	0.75	0.10	0.04	0.14
7	0.74	0.10	0.67	0.09	0.07	0.13
8	0.88	0.10	0.82	0.10	0.06	0.14
9	0.77	0.10	0.72	0.09	0.05	0.13
10	0.97	0.12	0.90	0.11	0.07	0.16
Average	0.78	0.10	0.73	0.09	0.05	0.14

In simulation 2, the number of treatment sites decreased from 200 to 100, the number of reference sites increased slightly from 2,997 to 3,097, and the true CMF reduced from 0.90 to 0.70.

The results from simulation 2 were consistent, as shown in table 74. However, the estimated CMF can vary greatly between trials, and the differences for each trial between the estimates of the CMF using the SPF with motorcycle AADT versus total AADT were small. The average CMFs using the motorcycle and total AADT respectively were 0.60 and 0.64, and the average difference across all trials was 0.04. As with simulation 1, the correction for regression to the mean appears not to be enough on average.

Table 74. Results for simulation 2 using type 3 Florida data.

Trial	CMF MOTO	SE MOTO	CMF AADT	SE AADT	Difference	SE Difference
11	0.73	0.13	0.78	0.14	0.05	0.19
12	0.30	0.08	0.31	0.08	0.01	0.11
13	0.50	0.11	0.50	0.11	0.00	0.16
14	0.65	0.12	0.63	0.12	0.02	0.17
15	0.72	0.14	0.76	0.15	0.04	0.21
16	0.57	0.12	0.64	0.14	0.07	0.18
17	0.54	0.11	0.58	0.13	0.04	0.17
18	0.64	0.13	0.69	0.15	0.05	0.20
19	0.81	0.15	0.89	0.16	0.08	0.22
20	0.55	0.12	0.59	0.13	0.04	0.18
Average	0.60	0.12	0.64	0.13	0.04	0.18

In simulation 3, the number of treatment sites increased to 1,000, the number of reference sites decreased to 2,197, and the true CMF remained 0.70.

As shown in table 75, the results from simulation 3 were also consistent in that the estimated CMF can vary between trials, and the differences for each trial between the estimates of the CMF using the SPF with motorcycle AADT versus total AADT were small. The average CMFs using the motorcycle and total AADT were 1.06 and 1.07, respectively, and the average difference across all trials was 0.03.

For simulation 3, the correction for regression to the mean appeared too large. With so many of the sites in the treatment group, the mean motorcycle crash rate of the reference group would be very low.

Table 75. Results for simulation 3 using type 3 Florida data.

Trial	CMF MOTO	SE MOTO	CMF AADT	SE AADT	Difference	SE Difference
21	0.99	0.08	0.99	0.08	0.00	0.11
22	1.21	0.09	1.23	0.09	0.02	0.13
23	0.96	0.07	0.99	0.08	0.03	0.11
24	0.93	0.07	0.94	0.07	0.01	0.10
25	1.38	0.10	1.45	0.10	0.07	0.14
26	1.00	0.08	0.98	0.08	0.02	0.11
27	0.98	0.08	1.03	0.09	0.05	0.12
28	1.15	0.09	1.12	0.09	0.03	0.13
29	0.99	0.08	1.00	0.08	0.01	0.11
30	0.97	0.08	0.96	0.08	0.01	0.11
Average	1.06	0.08	1.07	0.08	0.03	0.12

AVENUE B MODEL TYPE B2

The model type B2 analyses were conducted using the data for types 1 (rural freeway) and 3 (rural arterial) in Florida and for type 3 (rural non-freeway) in Pennsylvania.

Florida Type 1

Simulation 1

Figure 37 shows the model used to simulate the data for simulation 1 where dispersion equals 0.4975 and the CMF for AVGSHLDWID equals 0.89.

$$MOTOCRASHES = exp^{-5.3785} AVGMOTO^{0.8438} LENGTH^{0.8227} exp^{-0.118*AVGSHLDWID}$$

Figure 37. Equation. Florida type 1 simulation model 1.

The goal of the analysis was to re-estimate the model, including the parameter in the model for the average shoulder width variable, AVGSHLDWID. The project did this once with AVGMOTO and once with AVGAADT as the exposure measure.

Column 1 in table 76 indicates the trial number. Columns 2 and 3 show the parameter estimate and SE for the AVGSHLDWID variable using AVGMOTO in the model. Columns 4 and 5 show the same information for the model using AVGAADT in the model. Columns 6 and 7 provide the absolute value of the difference in parameter estimates and the SE of this difference. Columns 8 and 9 provide the inferred CMFs for the parameter values estimated. The last row shows the average value for all estimates across all trials. The estimated CMFs using the two exposure measures were very close for each trial and the average over all trials was very close to the true value of 0.89.

Table 76. B2 results simulation 1 for Florida type 1 data.

Trial	Estimate Using Motorcycle AADT	SE Using Motorcycle AADT	Estimate Using Total AADT	SE Using Total AADT	Difference in Estimates	Standard Deviation Of Difference	CMF Using Motorcycle AADT	CMF Using Total AADT
1	-0.0153	0.0567	-0.0146	0.0655	0.0007	0.0866	0.98	0.99
2	-0.1098	0.0571	-0.0775	0.0669	0.0323	0.0880	0.90	0.93
3	-0.1182	0.0570	-0.1342	0.0670	0.0160	0.0880	0.89	0.87
4	-0.1999	0.0624	-0.2226	0.0718	0.0227	0.0951	0.82	0.80
5	-0.1039	0.0585	-0.0976	0.0684	0.0063	0.0900	0.90	0.91
6	-0.0946	0.0544	-0.1419	0.0622	0.0473	0.0826	0.91	0.87
7	-0.0682	0.0586	-0.0511	0.0671	0.0171	0.0891	0.93	0.95
8	-0.1488	0.0572	-0.1584	0.0663	0.0096	0.0876	0.86	0.85
9	-0.1155	0.0553	-0.0678	0.0642	0.0477	0.0847	0.89	0.93
10	-0.1397	0.0511	-0.1606	0.0601	0.0209	0.0789	0.87	0.85
11	-0.1443	0.0635	-0.2093	0.0722	0.0650	0.0962	0.87	0.81
12	-0.2069	0.0645	-0.2454	0.0737	0.0385	0.0979	0.81	0.78
13	-0.2297	0.0589	-0.2360	0.0702	0.0063	0.0916	0.79	0.79
14	-0.1308	0.0541	-0.1274	0.0639	0.0034	0.0837	0.88	0.88
15	-0.1457	0.0561	-0.1459	0.0655	0.0002	0.0862	0.86	0.86
16	-0.1323	0.0574	-0.1216	0.0665	0.0107	0.0878	0.88	0.89
17	-0.0811	0.0554	-0.0614	0.0634	0.0197	0.0842	0.92	0.94
18	-0.1994	0.0577	-0.1742	0.0680	0.0252	0.0892	0.82	0.84
19	-0.1478	0.0548	-0.1817	0.0625	0.0339	0.0831	0.86	0.83
20	-0.0923	0.0589	-0.0910	0.0678	0.0013	0.0898	0.91	0.91
Average	-0.1312	0.0575	-0.1360	0.0667	0.0212	0.0880	0.88	0.87
STD	N/A	N/A	N/A	N/A	N/A	N/A	0.05	0.06
Minimum	N/A	N/A	N/A	N/A	N/A	N/A	0.79	0.78
Maximum	N/A	N/A	N/A	N/A	N/A	N/A	0.98	0.99

N/A = The statistic is not of interest and was not calculated.

Simulation 2

For simulation 2, the project team applied the same model used for simulation 1 (see figure 38) but with a much higher dispersion parameter equal to 5. The CMF for AVGSHLDWID equaled 0.89 per 1-ft increase in average shoulder width. The impact of the larger dispersion parameter was to create much more variability in crash counts between sites with similar road characteristics and traffic volumes.

$$MOTOCRASHES = exp^{-5.3785} AVGMOTO^{0.8438} LENGTH^{0.8227} exp^{-0.118*AVGSHLDWID}$$

Figure 38. Equation. Florida type 1 simulation model 2.

The results in table 77 again show that the estimated CMFs are close when using either the motorcycle or total AADT as an exposure measure. The average over all trials was close to the true value of 0.89: 0.92 for the motorcycle AADT model and 0.91 for the total AADT model. However, the standard deviation of the CMF estimates between trials is about double that for simulation 1.

Table 77. Model type B2 results simulation 2 for Florida type 1 data.

Trial	Estimate Using Motorcycle AADT	SE Using Motorcycle AADT	Estimate Using Total AADT	SE Using Total AADT	Difference in Estimates	Standard Deviation Of Difference	CMF Using Motorcycle AADT	CMF Using Total AADT
21	0.0017	0.0899	0.0692	0.1021	0.0675	0.1360	1.00	1.07
22	-0.0753	0.0871	-0.0740	0.0995	0.0013	0.1322	0.93	0.93
23	-0.1427	0.0894	-0.1690	0.1023	0.0263	0.1359	0.87	0.84
24	-0.0513	0.0919	-0.0833	0.1042	0.0320	0.1389	0.95	0.92
25	-0.1485	0.0850	-0.2201	0.0977	0.0716	0.1295	0.86	0.80
26	-0.2881	0.0975	-0.3399	0.1120	0.0518	0.1485	0.75	0.71
27	-0.0678	0.0852	-0.1569	0.0985	0.0891	0.1302	0.93	0.85
28	-0.1741	0.0901	-0.2295	0.1053	0.0554	0.1386	0.84	0.79
29	-0.0775	0.0901	-0.0503	0.1023	0.0272	0.1363	0.93	0.95
30	-0.0762	0.0821	-0.0764	0.0923	0.0002	0.1235	0.93	0.93
31	0.0582	0.0777	0.0634	0.0892	0.0052	0.1183	1.06	1.07
32	0.0217	0.0801	0.0276	0.0908	0.0059	0.1211	1.02	1.03
33	0.0406	0.0867	-0.0363	0.0976	0.0769	0.1305	1.04	0.96
34	-0.0211	0.0856	0.0024	0.0967	0.0235	0.1291	0.98	1.00
35	-0.1611	0.0912	-0.1655	0.1043	0.0044	0.1385	0.85	0.85
36	-0.2299	0.0951	-0.3021	0.1119	0.0722	0.1469	0.79	0.74
37	-0.2107	0.0916	-0.1018	0.1021	0.1089	0.1372	0.81	0.90
38	-0.2393	0.0834	-0.2532	0.0965	0.0139	0.1275	0.79	0.78
39	0.1332	0.0902	0.1433	0.1008	0.0101	0.1353	1.14	1.15
40	-0.1198	0.0808	-0.1585	0.0932	0.0387	0.1233	0.89	0.85
Average	-0.09	0.09	-0.11	0.10	0.04	0.13	0.92	0.91
STD	N/A	N/A	N/A	N/A	N/A	N/A	0.10	0.12
Minimum	N/A	N/A	N/A	N/A	N/A	N/A	0.75	0.71
Maximum	N/A	N/A	N/A	N/A	N/A	N/A	1.14	1.15

N/A = The statistic is not of interest and was not calculated.

Florida Type 3

Simulation 1

For the type 3 data in Florida, the project team pursued the simultaneous estimation of two CMFs. Figure 39 shows the model used to simulate the data for simulation 1. In this figure, dispersion equals 0.3390. The CMF for AVGSHLDWID equals 0.89 per 1-ft increase in average shoulder width. The CMF for MEDWIDTH equals 0.99 per 1-ft increase in median width.

$$MOTOCRASHES = exp^{-4.1264} AVGMOTO^{0.6155} LENGTH^{0.6683} exp^{-0.118*AVGSHLDWID-0.01*MEDWIDTH}$$

Figure 39. Equation. Florida type 3 simulation model 1.

The results in table 78 and table 79 show that the estimated CMFs were close when using either the motorcycle or total AADT as an exposure measure. The average over all trials was also very close to the true value CMF values.

Table 78. Model type B2 results simulation 1 for Florida type 3 data for AVGSHLDWID.

Trial	Estimate Using Motorcycle AADT	SE Using Motorcycle AADT	Estimate Using Total AADT	SE Using Total AADT	Difference in Estimates	Standard Deviation of Difference	CMF Using Motorcycle AADT	CMF Using Total AADT
1	-0.1039	0.0237	-0.1226	0.0245	0.0187	0.0341	0.90	0.88
2	-0.1580	0.0247	-0.1801	0.0256	0.0221	0.0356	0.85	0.84
3	-0.1532	0.0254	-0.1736	0.0267	0.0204	0.0369	0.86	0.84
4	-0.1054	0.0229	-0.1259	0.0237	0.0205	0.0330	0.90	0.88
5	-0.1086	0.0226	-0.1281	0.0233	0.0195	0.0325	0.90	0.88
6	-0.1122	0.0232	-0.1330	0.0240	0.0208	0.0334	0.89	0.88
7	-0.0117	0.0021	-0.1604	0.0255	0.1487	0.0256	0.99	0.85
8	-0.1490	0.0238	-0.1661	0.0245	0.0171	0.0342	0.86	0.85
9	-0.0937	0.0226	-0.1091	0.0231	0.0154	0.0323	0.91	0.90
10	-0.1248	0.0234	-0.1474	0.0242	0.0226	0.0269	0.88	0.86
11	-0.1088	0.0230	-0.1321	0.0240	0.0233	0.0332	0.90	0.88
12	-0.1175	0.0243	-0.1365	0.0250	0.0190	0.0349	0.89	0.87
13	-0.1180	0.0231	-0.1381	0.0238	0.0201	0.0332	0.89	0.87
14	-0.1680	0.0254	-0.1860	0.0260	0.0180	0.0363	0.85	0.83
15	-0.1233	0.0242	-0.1423	0.0246	0.0190	0.0345	0.88	0.87
16	-0.1047	0.0233	-0.1283	0.0242	0.0236	0.0336	0.90	0.88
17	-0.1164	0.0236	-0.1333	0.0245	0.0169	0.0340	0.89	0.88
18	-0.1392	0.0234	-0.0159	0.0243	0.1233	0.0337	0.87	0.98
19	-0.1568	0.0246	-0.1780	0.0252	0.0212	0.0352	0.85	0.84
20	-0.1251	0.0237	-0.1472	0.0246	0.0221	0.0342	0.88	0.86
Average	-0.1199	0.0227	-0.1392	0.0246	0.0316	0.0334	0.89	0.87
STD	N/A	N/A	N/A	N/A	N/A	N/A	0.03	0.03
Minimum	N/A	N/A	N/A	N/A	N/A	N/A	0.85	0.83
Maximum	N/A	N/A	N/A	N/A	N/A	N/A	0.99	0.98

N/A = The statistic is not of interest and was not calculated.

Table 79. Model type B2 results simulation 1 for Florida type 3 data for MEDWIDTH.

Trial	Estimate Using Motorcycle AADT	SE Using Motorcycle AADT	Estimate Using Total AADT	SE Using Total AADT	Difference in Estimates	Standard Deviation of Difference	CMF Using Motorcycle AADT	CMF Using Total AADT
.	-0.0062	0.0020	-0.0095	0.0024	0.0033	0.0031	0.99	0.99
2	-0.0107	0.0020	-0.0131	0.0024	0.0024	0.0031	0.99	0.99
3	-0.0114	0.0021	-0.0127	0.0025	0.0013	0.0033	0.99	0.99
4	-0.0122	0.0021	-0.0169	0.0025	0.0047	0.0033	0.99	0.98
5	-0.0125	0.0020	-0.0169	0.0024	0.0044	0.0031	0.99	0.98
6	-0.0113	0.0021	-0.0146	0.0024	0.0033	0.0032	0.99	0.99
7	-0.0117	0.0021	-0.0156	0.0025	0.0039	0.0033	0.99	0.98
8	-0.0101	0.0020	-0.0132	0.0023	0.0031	0.0030	0.99	0.99
9	-0.0111	0.0021	-0.0155	0.0024	0.0044	0.0032	0.99	0.98
10	-0.0087	0.0020	-0.0132	0.0024	0.0045	0.0031	0.99	0.99
11	-0.0075	0.0019	-0.0104	0.0023	0.0029	0.0030	0.99	0.99
12	-0.0071	0.0020	-0.0109	0.0024	0.0038	0.0031	0.99	0.99
13	-0.0105	0.0020	-0.0140	0.0023	0.0035	0.0030	0.99	0.99
14	-0.0117	0.0021	-0.0146	0.0024	0.0029	0.0032	0.99	0.99
15	-0.0090	0.0020	-0.0137	0.0024	0.0047	0.0031	0.99	0.99
16	-0.0103	0.0021	-0.0143	0.0025	0.0040	0.0033	0.99	0.99
17	-0.0101	0.0021	-0.0125	0.0024	0.0024	0.0032	0.99	0.99
18	-0.0147	0.0020	-0.0180	0.0024	0.0033	0.0031	0.99	0.98
19	-0.0121	0.0020	-0.0166	0.0024	0.0045	0.0031	0.99	0.98
20	-0.0110	0.0020	-0.0155	0.0024	0.0045	0.0031	0.99	0.98
Average	-0.0105	0.0020	-0.0141	0.0024	0.0036	0.0032	0.99	0.99
STD	N/A	N/A	N/A	N/A	N/A	STD	0.00	0.00
Minimum	N/A	N/A	N/A	N/A	N/A	Minimum	0.99	0.98
Maximum	N/A	N/A	N/A	N/A	N/A	Maximum	0.99	0.99

N/A = The statistic is not of interest and was not calculated.

Simulation 2

In simulation 2, the project team chose a random sample of 200 segments to investigate the results when the sample size is small. All other assumptions used in simulation 1 remained the same. The results in table 80 and table 81 show that as before, the CMF estimates for both geometric variables were very close between using motorcycle AADT or total AADT as the exposure measure. The average value over all trials was also close to the true value estimates. However, with the smaller sample size, the standard deviation of the CMF estimates was much higher.

Table 80. Model type B2 results simulation 2 for Florida type 3 data for AVGSHLDWID.

Trial	Estimate Using Motorcycle AADT	SE Using Motorcycle AADT	Estimate Using Total AADT	SE Using Total AADT	Difference in Estimates	Standard Deviation of Difference	CMF Using Motorcycle AADT	CMF Using Total AADT
1	-0.0864	0.2578	-0.0934	0.2509	0.0070	0.3597	0.92	0.91
2	-0.0921	0.2073	-0.1389	0.1859	0.0468	0.2784	0.91	0.87
3	-0.1128	0.3773	-0.1289	0.3831	0.0161	0.5377	0.89	0.88
4	-0.0865	0.1781	-0.1167	0.1825	0.0302	0.2550	0.92	0.89
5	-0.0328	0.1932	-0.0309	0.1874	0.0019	0.2692	0.97	0.97
6	-0.1992	0.2649	-0.2326	0.2808	0.0334	0.3860	0.82	0.79
7	-0.5430	0.3211	-0.5201	0.3140	0.0229	0.4491	0.58	0.59
8	-0.2165	0.2271	-0.2120	0.2226	0.0045	0.3180	0.81	0.81
9	0.0156	0.1320	0.0043	0.1488	0.0113	0.1989	1.02	1.00
10	0.0322	0.1865	0.0432	0.1783	0.0110	0.2580	1.03	1.04
Average	-0.1322	0.2345	-0.1426	0.2334	0.0185	0.3310	0.89	0.88
STD	N/A	N/A	N/A	N/A	N/A	N/A	0.13	0.13
Minimum	N/A	N/A	N/A	N/A	N/A	N/A	0.58	0.59
Maximum	N/A	N/A	N/A	N/A	N/A	N/A	1.03	1.04

N/A = The statistic is not of interest and was not calculated.

Table 81. Model type B2 results simulation 2 for Florida type 3 data for MEDWIDTH.

Trial	Estimate Using Motorcycle AADT	SE Using Motorcycle AADT	Estimate Using Total AADT	SE Using Total AADT	Difference in Estimates	Standard Deviation of Difference	CMF Using Motorcycle AADT	CMF Using Total AADT
1	-0.0065	0.0195	-0.0219	0.0248	0.0154	0.0315	0.99	0.98
2	-0.0066	0.0127	-0.0171	0.0153	0.0105	0.0199	0.99	0.98
3	0.0174	0.0167	0.0183	0.0184	0.0009	0.0248	1.02	1.02
4	0.0133	0.0112	0.0123	0.0153	0.0010	0.0190	1.01	1.01
5	0.0051	0.0193	0.0021	0.0205	0.0030	0.0282	1.01	1.00
6	-0.0121	0.0223	0.0028	0.0250	0.0149	0.0335	0.99	1.00
7	-0.0120	0.0201	-0.0114	0.0221	0.0006	0.0299	0.99	0.99
8	-0.0065	0.0161	-0.0082	0.0175	0.0017	0.0238	0.99	0.99
9	-0.0438	0.0308	-0.0482	0.0312	0.0044	0.0438	0.96	0.95
10	0.0002	0.0158	0.0024	0.0191	0.0022	0.0248	1.00	1.00
Average	-0.0052	0.0185	-0.0069	0.0209	0.0055	0.0279	0.99	0.99
STD	N/A	N/A	N/A	N/A	N/A	STD	0.02	0.02
Minimum	N/A	N/A	N/A	N/A	N/A	Minimum	0.96	0.95
Maximum	N/A	N/A	N/A	N/A	N/A	Maximum	1.02	1.02

N/A = The statistic is not of interest and was not calculated.

Pennsylvania Type 3

The project team ran one simulation using the type 3 sites in Pennsylvania. For this simulation, two geometric variables were included in the model: average shoulder width (AVGSHLDWID) and total surface width (SURFWIDTH). In this model, dispersion equals 0.8647. The CMF for AVGSHLDWID equals 0.97 per 1-ft increase in average shoulder width. The CMF for SURFWIDTH equals 1.03 per 1-ft increase in total surface width. Figure 40 shows the model used to simulate the data for simulation 1.

$$MOTOCRASHES = exp^{-5.0123} AVGMOTO^{0.4623} LENGTH^{0.7370} exp^{-0.0278*AVGSHLDWID + 0.0256*SURFWIDTH}$$

Figure 40. Equation. Pennsylvania type 3 simulation model 1.

For estimating the new models, the project team selected a random subset of 200 sites for each trial. This random selection allowed the project team to evaluate the impact of small sample sizes on the results.

The results in table 82 and table 83 show that, as before, the CMF estimates for both geometric variables are very close between using motorcycle AADT or total AADT as the exposure measure. The average value over all trials is also very close to the true value estimates.

Table 82. Model type B2 results simulation 1 for Pennsylvania type 3 data for AVGSHLDWID.

Trial	Estimate Using Motorcycle AADT	SE Using Motorcycle AADT	Estimate Using Total AADT	SE Using Total AADT	Difference in Estimates	Standard Deviation of Difference	CMF Using Motorcycle AADT	CMF Using Total AADT
1	0.0050	0.0572	0.0071	0.0597	0.0021	0.0827	1.01	1.01
2	0.0123	0.0536	0.0300	0.0555	0.0177	0.0772	1.01	1.03
3	-0.0634	0.0631	-0.0652	0.0655	0.0018	0.0909	0.94	0.94
4	-0.0944	0.0628	-0.0814	0.0651	0.0130	0.0905	0.91	0.92
5	-0.0294	0.0555	-0.0273	0.0577	0.0021	0.0801	0.97	0.97
6	-0.0954	0.0595	-0.0593	0.0620	0.0361	0.0859	0.91	0.94
7	-0.0981	0.0676	-0.1270	0.0708	0.0289	0.0979	0.91	0.88
8	-0.0457	0.0589	-0.0543	0.0618	0.0086	0.0854	0.96	0.95
9	-0.1707	0.0658	-0.1966	0.0690	0.0259	0.0953	0.84	0.82
10	-0.0223	0.0561	-0.0359	0.0584	0.0136	0.0810	0.98	0.96
11	-0.0400	0.0555	-0.0303	0.0578	0.0097	0.0801	0.96	0.97
12	-0.0504	0.0616	-0.0423	0.0642	0.0081	0.0890	0.95	0.96
13	0.0799	0.0545	0.0858	0.0564	0.0059	0.0784	1.08	1.09
14	-0.1335	0.0555	-0.1436	0.0585	0.0101	0.0806	0.88	0.87
15	-0.0061	0.0555	0.0012	0.0579	0.0073	0.0802	0.99	1.00
16	-0.0273	0.0560	-0.0276	0.0586	0.0003	0.0811	0.97	0.97
17	-0.0726	0.0645	-0.0939	0.0674	0.0213	0.0933	0.93	0.91
18	0.0023	0.0606	-0.0089	0.0629	0.0112	0.0873	1.00	0.99
19	-0.0777	0.0632	-0.0726	0.0657	0.0051	0.0912	0.93	0.93
20	-0.0379	0.0556	-0.0234	0.0578	0.0145	0.0802	0.96	0.98
Average	-0.0483	0.0591	-0.0483	0.0616	0.0122	0.0854	0.95	0.95
STD	N/A	N/A	N/A	N/A	N/A	N/A	0.05	0.06
Minimum	N/A	N/A	N/A	N/A	N/A	N/A	0.84	0.82
Maximum	N/A	N/A	N/A	N/A	N/A	N/A	1.08	1.09

N/A = The statistic is not of interest and was not calculated.

Table 83. Model type B2 results simulation 1 for Pennsylvania type 3 data for SURFWIDTH.

Trial	Estimate Using Motorcycle AADT	SE Using Motorcycle AADT	Estimate Using Total AADT	SE Using Total AADT	Difference in Estimates	Standard Deviation of Difference	CMF Using Motorcycle AADT	CMF Using Total AADT
1	0.0132	0.0275	0.0203	0.0275	0.0071	0.0389	1.01	1.02
2	0.0415	0.0221	0.0532	0.0216	0.0117	0.0309	1.04	1.05
3	0.0303	0.0289	0.0361	0.0290	0.0058	0.0409	1.03	1.03
4	0.0536	0.0246	0.0658	0.0241	0.0122	0.0344	1.06	1.07
5	0.0265	0.0243	0.0327	0.0241	0.0062	0.0342	1.03	1.03
6	0.0331	0.0244	0.0544	0.0228	0.0213	0.0334	1.03	1.06
7	0.0110	0.0323	0.0030	0.0339	0.0080	0.0468	1.01	1.00
8	-.0303	0.0336	-0.0229	0.0342	0.0074	0.0479	0.97	0.98
9	0.0258	0.0291	0.0185	0.0306	0.0073	0.0422	1.03	1.02
10	0.0201	0.0261	0.0207	0.0267	0.0006	0.0373	1.02	1.02
11	0.0233	0.0249	0.0333	0.0243	0.0100	0.0348	1.02	1.03
12	0.0209	0.0284	0.0315	0.0279	0.0106	0.0398	1.02	1.03
13	0.0485	0.0226	0.0568	0.0225	0.0083	0.0319	1.05	1.06
14	0.0587	0.0200	0.0608	0.0201	0.0021	0.0284	1.06	1.06
15	0.0296	0.0244	0.0392	0.0241	0.0096	0.0343	1.03	1.04
16	-.0245	0.0318	-0.0147	0.0320	0.0098	0.0451	0.98	0.99
17	0.0238	0.0298	0.0184	0.0309	0.0054	0.0429	1.02	1.02
18	0.0389	0.0264	0.0401	0.0267	0.0012	0.0375	1.04	1.04
19	0.0328	0.0293	0.0391	0.0294	0.0063	0.0415	1.03	1.04
20	0.0211	0.0257	0.0351	0.0250	0.0140	0.0359	1.02	1.04
Average	0.0249	0.0268	0.0311	0.0269	0.0082	0.0380	1.03	1.03
STD	N/A	N/A	N/A	N/A	N/A	N/A	0.02	0.02
Minimum	N/A	N/A	N/A	N/A	N/A	N/A	0.97	0.98
Maximum	N/A	N/A	N/A	N/A	N/A	N/A	1.06	1.07

N/A = The statistic is not of interest and was not calculated.

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

Data on traffic volumes are vital to the development of SPFs required for effective implementation of strategies to improve the safety of road networks. Mitigating motorcycle crashes can be especially challenging in this regard because few jurisdictions collect motorcycle traffic volume data systematically. To address this challenge, the project team conducted several analyses to explore (1) how much predictive power for an SPF is lost when motorcycle volumes are unknown and how this lack of information may affect a study of motorcycle countermeasures, and (2) alternative methods for deriving accurate predictions of motorcycle crashes or motorcycle volumes. The research investigated and demonstrated methods and the mathematical models that can be applied by jurisdictions that lack motorcycle volumes when undertaking the development and the evaluation of motorcycle-related safety countermeasures to estimate CMFs.

The project team investigated two groups, or avenues, of methods. The methods for avenue A focused on investigating (1) the difference in predictive performance for motorcycle SPFs calibrated with motorcycle AADT versus total AADT, (2) relating total crash SPFs to motorcycle crash SPFs so jurisdictions without motorcycle volumes could predict motorcycle crashes using total crash SPFs, and (3) methods to predict segment-level motorcycle AADT. The methods for avenue B focused on the differences in CMF estimates when using motorcycle AADT versus total AADT when applying before-after or cross-sectional regression CMF estimation methods.

For developing the avenue A models, data were collected from Florida and Pennsylvania. Both States had a large number of locations with an estimated motorcycle AADT and could provide linkable roadway inventory, traffic, and crash data. The project team also acquired data from Virginia to validate the models developed.

Table 84 summarizes the avenue A methods, with a final column on conclusions from the analysis. In addition to the analyses depicted, an assessment was conducted of how well EB estimates derived from the model type A1 models predict future motorcycle crashes for high-crash locations typically of interest in countermeasure applications that form the basis for future CMF development. The results of that assessment show that the models using total AADT and those using motorcycle AADT perform similarly, although the EB estimates from the total AADT models of the crashes at the high crash sites are marginally closer to the actual future crash counts than those based on the motorcycle AADT models.

Table 84. Summary of avenue A method elements and results.

Model Type and Intended Function	Basic Purpose	SPFs Developed	Approach	Conclusion
A1: Provide a direct measure of how the predictive power of a model is affected by either including or excluding motorcycle volumes.	To explore how much predictive power is lost when motorcycle volumes are unknown.	A1.1. Motorcycle crashes versus total AADT and other independent variables. A1.2. Motorcycle crashes versus motorcycle AADT and other independent variables.	1. Assess goodness-of-fit of two model sets and compare. 2. Assess how well each model set predicts motorcycle crashes at high crash locations. 3. After steps 1 and 2, assess predictive ability of SPF 1. 4. Consider SPF A1.1 for application to any jurisdiction if successful. 5. Use FHWA SPF calibration tool to assess application of SPF A1.1 to selected jurisdictions.	Overall, models with total AADT perform at least as well as those with motorcycle AADT for both arterials and freeways and even slightly better for freeways in Florida.
A2: Allow jurisdictions without motorcycle volumes to predict motorcycle crashes based on SPFs for total crashes.	Develop a relationship between predicted motorcycle crash frequency and predicted total crash frequency.	A2.1. Motorcycle crashes versus motorcycle AADT. A2.2. All crashes versus total AADT. A2.3. Predicted motorcycle crashes versus predicted total crashes and other variables.	1. Develop and assess a model that relates predictions from SPF A2.1 to predictions from SPF A2.2. 2. Consider SPF A2.3 for application to any jurisdiction if successful.	Models were successfully developed using the data for Florida. For Pennsylvania, no satisfactory models could be developed. In Florida, for urban and rural freeways, both A2 SPFs can outperform the corresponding models that predict motorcycle crashes from motorcycle AADT. For urban and rural arterials, the opposite is true.

Model Type and Intended Function	Basic Purpose	SPFs Developed	Approach	Conclusion
A3: Allow jurisdictions to directly estimate motorcycle volumes.	Develop models to estimate motorcycle traffic volumes based on roadway characteristics and other variables that may influence motorcycle trip generation.	A3 Motorcycle AADT versus variables related to roadway segment and county-level estimates of motorcycle registrations, licensing, and sociodemographic characteristics.	1. Assess/include variables that cause motorcycle AADT to vary. 2. Consider model A3 for estimating AADT in any jurisdiction where causal variables available if successful.	The models showed low R^2 values indicating that they are not explaining much of the variation in motorcycle AADT between road segments. For this reason, the A3 modeling was not considered a success.

The methods applied in avenue B make use of simulated data. Simulating data creates a database with many locations and with assumed relationships between roadway geometry or other countermeasures and motorcycle crashes. This tests the ability to accurately measure this true relationship when motorcycle volumes are and are not used in the process. The fixed relationships affecting motorcycle crashes were determined considering a likely range of values based on existing safety knowledge.

To investigate the impact of the lack of motorcycle AADTs on the estimation of CMFs, two CMF estimation approaches were investigated: model type B1, the EB before-after approach, and model type B2, cross-sectional generalized linear models. The project team chose the approach for avenue B because it will provide a direct measure of how the lack of motorcycle AADT affects CMF estimation by replicating the process of estimating CMFs.

For the EB before-after approach, a countermeasure was assumed with a known value of its CMF. The project team divided the simulated database into two time periods and adjusted by the value of the CMF in the after period the expected crash means for each location. The Poisson distribution generated the new after period counts. The project team then applied the EB approach to the data for these treated sites, using the remaining sites as a reference group. The project team completed this once using the motorcycle AADTs and once total AADT. The project team then made a comparison to see how the lack of motorcycle AADT affected the estimate of the CMF. The project team performed this entire process, beginning with the simulated data, multiple times and with multiple sample sizes and assumed CMF values.

For the cross-sectional regression model approach, an assumed CMF relationship based on logical considerations and related research was defined and added to the SPFs developed in model type A1. The project team used this modified SPF to simulate the data. The project team then used GLM to re-estimate the SPF, including the fictional variable, with and without motorcycle AADTs. The project team then made a comparison to see how the lack of motorcycle AADT affected the estimate of the CMF. The project team performed this entire process, beginning with the simulated data, multiple times with varying sample sizes and assumed CMF values.

The avenue B analyses used the roadway inventory, total AADT, and motorcycle AADT collected for the avenue A methods in Florida and Pennsylvania. For motorcycle crashes, the project team simulated the crash counts using the SPFs developed in the model type A1 models as a starting point.

The results for avenue B, which investigated the EB before-after approach, indicate that there was relatively little difference (0.05 or less) between the CMFs estimated using the motorcycle AADT SPF versus using the total AADT SPF. However, the estimated CMFs between simulation runs can vary considerably due to the relatively low frequency of motorcycle crashes.

The results for avenue B, which investigated the cross-sectional regression approach for estimating CMFs, also showed that the estimated CMFs were close when using either the motorcycle or total AADT as an exposure measure but that the variability in CMF estimates between simulations was large.

The findings of both the avenue A and avenue B modeling indicate that when motorcycle volumes are not known, using total AADT on its own is sufficient for developing SPFs and CMFs. The potential bias due to missing motorcycle-specific AADT is sufficiently negligible where it exists so as not to preclude SPF and CMF development.

The project team concluded that attempting to predict motorcycle volumes is not possible using typically available roadway and county-level data. Improvement could possibly be found in trip generation type modeling at a disaggregate scale, although given the success of the SPFs using total AADT, such an effort may not be worthwhile.

A more significant issue in developing motorcycle crash SPFs and CMFs is working with a crash type that is relatively rare. The analysis did not develop SPFs for all motorcycle crash types or site types. More evidently, in the estimation of CMFs, the CMF value varied significantly between simulation runs due to the low frequency of motorcycle crashes.

CHAPTER 7. LIMITATIONS AND FUTURE RESEARCH NEEDS

This section summarizes existing data limitations and research gaps identified through the assessment of available data sources, analytical methods, and the evaluation results.

DATA LIMITATIONS

With respect to developing SPFs and CMFs for motorcycle crashes, there are a number of data limitations related to traffic volumes (AADT), crash data, and roadway inventory data.

AADT Data

Technology

States collect traffic volume counts using a variety of techniques and technologies and convert these counts into estimated AADT volume estimates. With respect to motorcycle counts, NCHRP 08-36, Task 92, *Counting Motorcycles*, and NCHRP Report 760, *Improving Motorcycle Travel Data – Data Collection Protocols and Field Tests*, provide excellent detail on the strengths and weaknesses of the existing counter technologies.(29,30)

The second of these reports tested five detector technologies for their ability to accurately classify all vehicle types, specifically motorcycles. The five detector types tested were as follows:

- **Infrared (IR) Classifier:** This is a portable or fixed location system using IR beam interruption to identify and classify vehicles. The version tested in NCHRP 760 was The Infra-Red Traffic Logger (TIRTL).(30) The unit sends beams in four pathways (two perpendicular to the direction of travel and two at the diagonals). The receiver detects are records two timed events (beam interruption and re-establishment of the beam) as a vehicle passes through the four beam paths.

- **Inductive loops/piezo electric sensors:** The *Traffic Detector Handbook: Third Edition* describes the principal components of an inductive loop detector as loops of insulated wire placed in a slot sawed in the pavement and connected to an electronic controller unit.(31) As vehicles pass over the loop, they lower the inductance in the loop, and this change is recorded by the electronics in the controller unit. Each passage is time stamped. In units capable of classification counts, software is programmed to match the inductance changes over time with the pattern expected for each of the 23 FHWA-defined vehicle classification bins. Piezoelectric detectors are imbedded sensors that send an electronic pulse to the controller unit whenever an axle/tire travels over the sensor. The pulse varies by tire force, affecting the sensor. These sensors can be used in an array to provide vehicle classification data based on the weight and number of tires going over the sensor array.

- When combined, as is done in Virginia, the inductive loops and piezo sensors detect motorcycles with increased accuracy based on magnetic length (from the inductive loop) and rejection of energy detected from adjacent lanes (based on waveform analysis from the piezo).

- **Magnetometers:** Magnetometers work in a similar fashion to inductive loops in that they detect changes to a magnetic field as vehicles pass through the detection zone. Magnetometers installed in an array are used with accompanying software to classify vehicles by the timing and extent of changes in the magnetic field. Magnetometers are passive, meaning that a portion of the vehicle must pass over the detector. This makes this type of detector ideal (when installed in groups in a pattern that covers the lane correctly) for classification counts and for detecting vehicle spacing.

- **Multi-technology system:** This is a newer technology implemented specifically to count motorcycles. At the time of the research for NCHRP 760, the technology had not yet advanced to the point where it could reliably classify non-motorcycles.(30) As tested, it is a lower cost alternative than the other technologies.

- **Tracking video:** At the time of the research for NCHRP 760, video-based real-time counting and vehicle classification was undergoing significant changes.(30) The technology as tested used digital video processing algorithms to identify and classify vehicles based on shape profiles. Planned addition of IR video process would potentially improve the technology by allowing the unit to also detect the heat signatures of vehicles for use in classification.(30)

Table 85 compares the five technologies tested in NCHRP 760.(30) For each, there are two measures of accuracy: MC for motorcycle detection, and non-MC for detection of all other vehicles. Costs for two- and four-lane roadways are shown for initial installation. While there may be portable versions of many of the technologies, the only portable system tested was the TIRTL. The skill level required to set up and calibrate the detector and associated electronics/software is shown as well. Table 86 (table 2 from NCHRP 08-36) shows the technology used in 24 States.(29)

Table 85. Comparisons of five detector technologies from NCHRP 760.[(30)]

Technology	MC Accuracy (Percent)	Non-MC Accuracy (Percent)	Initial Cost (Two-Lane)	Initial Cost (Four-Lane)	Portability	Skill Level for Setup[a]
IR Classifier	95	98	$26,850	$26,850	Fixed/portable[b]	Expert
Inductive loops/ piezoelectric sensors (full lane-width)	45	95	$33,000[c]	$61,000	Fixed	Field technician
Magnetometers	80	95	$10,204	$15,964	Fixed[d]	Field technician
Multi-technology system	50	N/A	$6,000	$12,000	Fixed[d]	Field technician
Tracking video system	75	90	$15,000	$15,000	Fixed[d]	Field technician

[a]Setup skill level required: expert versus field technician with proper training.
[b]TIRTL is available as either portable or fixed, but only portable TIRTL was tested in this research.
[c]Estimated by Texas Department of Transportation: $61,000 total for four-lane site and $33,000 total for two-lane site.
[d]Some components could be portable, or the detector could be portable with modification.
N/A = Not available.

Table 86. Use of various detector technologies in 24 States as reported in NCHRP 08-36.[(29)]

Technology	Short Counts Tested	Short Counts Used	Continuous Counts Tested	Continuous Counts Used
Intrusive				
Road tubes	13	20	N/A	N/A
Piezoelectric cable	3	4	9	17
Conventional inductive loops	6	2	4	8
Piezoelectric film	1	0	4	3
Inductive loop signatures	1	0	2	1
Quadrupole loops	1	0	1	0
Magnetometers	1	0	2	0
Non-Intrusive				
Manual	0	1	N/A	N/A
Radar	7	3	4	5
Video	1	2	2	1
IR, including TIRTLs	5	0	4	3
Acoustic	1	0	2	0

N/A = Not applicable.

As can be seen from table 86, road tubes are the most common type of detector for short counts—NCHRP 760 did not test the road tubes because they were not considered to be capable of providing accurate motorcycle classification counts.[(30)] Piezoelectric and inductive loops together account for the vast majority of detector installations in the 24 States that responded to the survey. Of the non-intrusive methods, the TIRTL is second-most common (after radar), but the use of the non-intrusive technologies still lagged far behind the inductive loops and piezoelectric installations as of 2010 when the project was completed.

NCHRP 760 concluded that the TIRTL was the most accurate system for detecting both motorcycles and non-motorcycles.[30] It has some higher skill level requirements than the other detectors, and a relatively higher price compared to the next-most accurate system, magnetometers. The two technologies that performed least well in the test (multi-technology and tracking video) are undergoing rapid technology improvements and may become more useful in the future. If accuracy is improved, then their pricing and required skill level for installation are attractive.

AADT Data Limitations

FHWA's *Traffic Detector Handbook* provides details on the limitations of motorcycle count data available.[31] In general, the report finds the relatively small amount of metal in many motorcycles combined with the fact that many motorcyclists ride near lane lines in order to give themselves more time to avoid cars moving into their lanes means that inductive loop detectors and half-lane axle sensors often undercount motorcycles. When motorcycles ride in closely spaced groups, the closely spaced axles and cycles often confuse available traffic monitoring equipment that has not been designed to identify the resulting pattern of closely spaced axles and vehicles.[31]

The following summarizes the limitations of motorcycle AADT estimates in the United States:

- **Permanent counter installations:** Most permanent count system installations are not optimized for counting motorcycles. This is important because permanent counters are the source of most classification counts taken by States. There are two main concerns with current installations: calibration and detector configuration. Calibration refers to the sensitivity settings of the detector and software to arrive at correct bin assignments for each vehicle detected. The optimal calibration is different for each detector technology and will not be repeated here. However, it is important to recognize that detectors do go out of calibration and must be checked and adjusted or replaced. The software interprets input signals from the detectors and performs the assignment of vehicle counts into bins based on the detected size of the vehicle. Again, the technical details are different for each detector technology and will not be repeated here. The software is proprietary firmware with some user control over the associations between detected vehicle size and bin assignments. This is critical because when detecting motorcycles, the range of vehicle sizes can overlap with small cars, and the detector/software arrangement can mis-assign motorcycles to other vehicle classes when the motorcycles travel in groups. Depending on how the detectors are placed in the roadway, motorcycles traveling near the edge of a lane may be missed entirely.

- **Temporary count technology:** Temporary classification count methods and technology exist that can capture motorcycles accurately. This technology is not widely in use throughout the United States at present. The more accurate temporary installations are, unfortunately, more complex than older, more familiar technology. As a result, there is concern that it may not be installed properly in all instances or that it is more susceptible to failure (at least with respect to accurate motorcycle counting) than the older technology. As adoption of the new technologies is slow, there is not enough practical field data available to judge the benefit/cost of the newer technology. For the purposes of this summary, however, it is important to note is that there are relatively few reliable motorcycle classification counts outside of permanent count locations.

- **Detector calibration:** Taking temporary counts to be a negligible contribution to the motorcycle count data at present, the assessment of how accurate motorcycle AADT is in the United States requires detailed knowledge of just the permanent count locations. Installations and tracking of sensor calibration vary from State to State. It is difficult, if not impossible, to summarize how accurate motorcycle counts are at a national level. At the State level, it is possible to assess how well the existing count infrastructure matches the recently developed advice on detector types and configuration, though that evaluation has not yet been done for every State. Calibration of detectors also varies by State. Virginia, for example, keeps records that help to identify failing detectors in advance of a complete data loss. Other States may have this information, but Virginia supplies it as part of their count database so users can decide for themselves whether to use the data from any particular permanent count location.

  Some detector types (such as piezo) have a failure mode that affects small vehicle classifications earliest and most severely. When a detector starts to fail, motorcycle and small car counts are affected most. If a State is not proactively keeping the detectors calibrated and replacing those that are failing, analysts may have to selectively exclude some count sites for some time periods based on suspicions of unreliable data.

- **Motorcycles traveling in groups:** This issue is related to sensor configuration and calibration. When motorcycles travel in groups, they often ride in parallel or staggered formations at the outsides of a single lane. Detector systems that are not specifically configured (as discussed in NCHRP 08-81) to correctly recognize these groupings do a poor job classifying motorcycles into the proper bin.[30,32]

- **Count locations.** NCHRP 08-81 addresses another concern often expressed by motorcycling advocates—that the count locations are not selected in a way that optimizes their reliability for estimating motorcycle AADT.[30,32] In particular, the concern is that recreational riders go places where classification counts are rarely collected. The result is that State-level aggregated counts are lower than they should be because a large proportion of motorcycle trips take place off of the facilities with the best, most accurate counting technology; many of those trips are on facilities that have no classification counts collected. Estimated factors may be applied to arrive at an estimated motorcycle AADT for those locations; however, there is some concern that the factors themselves are not accurate for the facilities in question. Middleton et al. were able to show that crash locations are a reasonable surrogate for motorcycle traffic volume such that States could use that information when deciding where to place counters or to evaluate whether the current count locations are sufficient to capture reliable area- or State-level motorcycle counts.[33] In their study, the authors found that the placement of counters does vary considerably by State with the conclusion being that each State should examine the spatial correlation of permanent count locations and motorcycle crash locations when deciding if they need to make additions or changes to the count locations. The findings indicated that motorcycle crash locations are not too different from crashes involving other vehicle types and that, in general, permanent counters on higher functional class roadways are likely to be sufficient. At least on an area-wide basis, the locations with high crash counts correlated well with locations of high motorcycle volume counts and high total traffic volume counts.

- **Weekday versus weekend counts.** This is an issue that affects temporary count locations. Typically, these counts are scheduled as mid-week counts because they are aimed at obtaining typical travel volumes at a location. While motorcycle crash data would indicate that weekday travel is an important component of all motorcycle travel, it is unquestionably true that a significant (though unknown) proportion of all motorcycle trips take place on the weekends. Of purely recreational motorcycle trips, the weekend proportion is much higher. Estimating weekend motorcycle travel is thus complicated by the fact that weekend factors developed based on other available data (i.e., from permanent count locations) is inaccurate for recreational trips because it captures mainly the travel on major roadways. Motorcyclists do not completely avoid those major roadways, of course, but much of the recreational motorcycle travel is on weekends and off the major roadways.

For overall traffic safety analysis, these shortcomings indicate that motorcycle counts may be inaccurate, and the inaccuracies vary among the States in ways that are difficult to assess or make adjustments for an analysis. The data from any specific site (most likely a permanent count location in a State) will have unknown under-reporting problems depending on how closely the installation matches the ideal design.(30) With few exceptions, the health of the detector—its current ability to correctly detect motorcycles—is unknown. Even if a detector's installation was well designed originally, its current status might not be available to researchers who might use flawed data without knowing that the detector was beginning to fail during the period in which their data was collected. Ultimately, all of the problems would appear to lead to under-reporting of motorcycle counts, and thus an analyst might feel secure in viewing the data as a minimum. However, even that assurance may be misplaced. Detailed knowledge of the detector system's software and any user-defined settings is needed before analysts could be comfortable that they know what happens with mistaken detections of various types and how the software is set to record those counts in the various classification bins.

Crash Data Limitations

Crash data standards, completeness, and accuracy vary among States. The difficulties using crash data in safety analyses are well known but not particularly well documented in the literature. Practitioners are, however, well aware of the following data limitations:

- **Accident reports and data definitions**: Each State creates its own police accident report (PAR) and decides what data elements to include in its centralized State crash database. There is a national guideline—the Model Minimum Uniform Crash Criteria (MMUCC)—which provides a standardized set of data definitions for a minimum set of 110 data elements. However, this is a voluntary guideline, and most States' PAR and crash database are below 100 percent compliance with MMUCC. Analysts have to understand the specific data element definitions of each State's data that was used. Depending on which data elements are considered in the analysis, the ability to generalize the results from one State to another, or to the national picture, can be quite limited.

- **Location accuracy**: States generally do an excellent job of locating crashes on the roadway network for any State-maintained roads and for any roadways eligible for funding under the Highway Safety Improvement Program (HSIP). Recent FHWA requirements for each State to have an all-public-roads linear referencing system (LRS) and basic roadway inventory data means that soon all States will be able to place every crash in a single LRS. Today, however, there are still States that have not achieved that step of assigning LRS location codes to every crash. The result is that safety analysis over the entire network is complicated by the fact that only some crashes are readily associated with AADT and roadway descriptive data (for example, in a geographic information system (GIS) implementation of the State-roads LRS). Local roadway crashes may be located only using a non-linear location coding scheme that is incompatible with the State's GIS and LRS. As a result, the ability to access and use local roadway crashes is impeded. This may not be a problem for some types of motorcycle safety analyses (e.g., ones focusing on motorcycle crashes on interstates), but it can be a serious barrier to analyzing motorcycle crashes in intersections at the local level or rural crashes on low-volume, low-functional class roadways. These crashes also tend to be the ones that State crash location staff spend the most time trying to correct. Local law enforcement agencies typically use local designations for roadways in their jurisdiction and may not provide the information that would allow the State-level staff to identify the location on the official State GIS or LRS. As discussed earlier, these are also the locations that are least likely to have a classification count available.

- **Crash severity**: Most (but not all) States use a variation on the KABCO injury severity scale to code personal injuries in crashes. K, A, B, and C are coded as injuries to individuals involved in the crash and correspond with the overall crash severity—the highest severity single injury in the crash is assigned as the overall crash severity. The definitions of these terms can vary markedly between States. Fatalities are typically defined based on the FARS criterion of a crash-related death within 30 days of the event. A-level injuries are the most severe level, but the descriptions may include "serious," "incapacitating," or other terms. B-level injuries are typically defined as apparent injuries at a lower or moderate level, with "non-incapacitating" and "apparent" often used as descriptors. The C-level injury definition may include descriptors such as "slight" or "possible." PDO crashes are those in which there was no personal injury recorded but enough property damage to make the crash reportable under the State's threshold criterion. Those criteria also vary among States and in interpretation among law enforcement agencies within a State.

  Typically, States set a minimum dollar amount threshold plus any injury or fatality so that, in theory, all KABC crashes should be reported along with any crashes with property damage about the threshold dollar amount. In practice, the threshold dollar amount is treated as a rough approximation, and each agency's practices dictate how its officers interpret the physical damage to vehicles and other property to arrive at the decision of whether or not to report a PDO crash. Another factor affecting reported motorcycle crashes is that many single-vehicle motorcycle crashes are unreported.

Safety analysis addresses the variation in severity codes in two basic ways. One way is to use all crashes. In practice, this equates to using all reported crashes, which in turn means that there is likely to be some systematic under-reporting. If safety comparisons are planned among jurisdictions, using all levels of crash severity can cause problems if the law enforcement agencies interpret the crash reporting threshold differently across those jurisdictions. For that reason, and because data quality is usually best for reports of serious crashes, analysts sometimes concentrate on "serious" crashes by taking fatalities and A-level injuries together or sometimes combining K, A, and B injury crashes. This has the advantage of providing greater comparability across jurisdictions but at the price of missing the majority of reported crashes. Typically, PDO crashes account for about 60 to 70 percent of all crashes. Analyses focused on crash locations and their attributes, excluding PDO and C-level injury crashes, are likely to rob the analysis of generalizability.

No description of crash severity would be complete without also pointing out that the KABCO values are determined by law enforcement officers, not trained medical personnel. When crash data and medical records are linked (as has been done many times in several States), the results point to a large discrepancy between the officers' judgements of injury severity and the actual injury severity coded based on medical injury severity codes, medical treatments, or cost of treatment. As might be expected, fatalities are judged accurately most of the time (though there are cases of death away from the scene that are sometimes missed as well as successful patient resuscitations away from the scene). A, B, and C injury codes are often medically incorrect, so the calculated crash severity scores based on those injury codes are prone to error. These errors do not pose a large concern for most safety analyses because the errors are usually randomly distributed and the comparisons being made in the typical safety analysis would not be expected to change much. Unfortunately, typical motorcycle crash-related injuries are of a nature where the differences between officers' and medical judgements are largest. If an analysis requires comparison of motorcycle crashes to all other vehicles' crashes, the possibility exists that the KABC assignments to the motorcycle crashes are less accurate than the ones for occupants of other vehicles.

- **Crash contributing factors, harmful events, and characteristics**: Safety data analyses sometimes must obtain roadway and human factors information based solely on the PAR. For example, if the State lacks a robust roadway inventory system that can supply location descriptions, the only information about the crash site will come from the PAR.

  The officers record those circumstances that they judge to have contributed to the crash, the apparent sequence of events, most harmful events, etc. When there are inconsistencies in data definitions, the ability to reliably aggregate data and form valid comparisons is limited. In addition, different States record different aspects of the crash, including the basic location descriptions. Most States collect a minimum set of roadway attributes, but those attributes differ widely among the States. If the analysis relies on knowing something about the roadway attributes, or circumstances of the crash, the analyst must know details about the data definitions as set by the State and, preferably, have some measures of accuracy and completeness available in order to judge the sufficiency of the database to support the intended analysis.

Roadway Inventory Limitations

Safety analyses often make use of roadway inventory data in order to understand the roadway attributes' association with crash risk. There are a few notes worthy of consideration because they relate to AADT and crash data and the statewide roadway inventory data.

- **Roadway class/ownership**: Because of the limitations of AADT data and in some States the limited ability to locate local crashes on a statewide LRS, safety analyses are sometimes limited to the State-maintained/HSIP-eligible portions of the roadway network. As noted earlier, this selective use of higher functional classification roadways can miss an important subset of motorcycle crashes. Crashes on low-volume rural roads or at local roadway locations are more likely to be unlinked to any available roadway characteristics data. If the analysis focuses on specific roadway characteristics, that data may be missing for the local roadways.

- **Missing data**: As noted earlier, many State LRSs lack locations for local and low-volume roadways, although the systems are adding this information now. A related problem is that even when the LRS includes those locations, the roadway inventory file may have blanks for some of the key data. Just as AADT is less readily available for local and low-volume roadways in many States, so too are the roadway inventory data less complete. When the analysis requires knowledge of roadway attributes, the ability to compare jurisdictions may be impaired if too much of the detailed inventory data is missing.

ANALYSIS LIMITATIONS

As discussed in chapter 2, most of the current research concerning motorcycle crashes has focused on discrete outcomes (i.e., the probability of a given crash severity, presence of a roadway or traffic control feature given that a crash has occurred, or probability of injury severity given that feature or crash type). Very little research focuses on developing SPFs or CMFs specifically for motorcycle crashes.

The same analysis methods available for estimating SPFs and CMFs for other crash types are applicable for motorcycle crashes. While having estimates of motorcycle volumes is preferred, the results of the analyses undertaken for this project indicate that using total AADT volumes is a reasonable substitution when motorcycle AADT is not available.

Where motorcycle use is a small portion of traffic volume, such as the United States, some research has attempted to use motorcycle licensing and/or registration data as a surrogate, but these data are only available at the county level and so are not very useful for modeling site-level data.

Issues that apply to other rare crash types also hinder analyses of motorcycle crashes in order to develop SPFs and/or CMFs. Firstly, with rare crash types, low crash sample sizes make the development of reliable SPFs and CMFs difficult. Statistical models may not be estimable, or even if they are, estimated parameters may be very imprecise, and few variables related to crashes may be included in the model.

Another sample size issue is that roadway infrastructure treatments aimed at reducing motorcycle crashes are not common. When conducting a before-after study, the rarity of such treatment sites combined with low crash frequencies presents a formidable challenge.

The results of the research conducted for this project confirmed all of the analysis limitations stated above. Even with large datasets containing a substantial mileage of roads, SPFs could not be estimated for all site types or subtypes of motorcycle crashes. (Single-vehicle and multi-vehicle were attempted as well as total motorcycle crashes.) The simulation results estimating CMFs through before-after studies and cross-sectional regression modeling showed that the CMF estimates vary substantially between simulations due to low motorcycle crash frequencies.

RESEARCH GAPS AND NEEDS

In terms of research gaps with respect to motorcycle safety and CMFs, very little information is known on the effects of roadway geometric and traffic control features on motorcycle crash frequency and severity. The reason for this gap is likely twofold: motorcycle crashes are not usually the focus of safety related countermeasures, and, the rarity of motorcycle crashes combined with the scarcity of treatment locations would result in a small sample size for study. Future research will need to explore how to overcome the small sample size issue with appropriate methodologies. This discussion provides some thoughts in this regard.

With respect to SPFs for applying in network screening and other safety management tasks, few SPFs at the segment level or intersection level exist. The SPFs developed in this project may contribute to filling this void, but there remains work to be done in terms of site types for which no SPF was developed and ensuring that models exist that calibrate well in all jurisdictions.

A major need for the research community would be a database that includes countermeasures implemented that are expected to affect motorcycle crashes along with the location, date of treatment, and treatment description. This information would aid researchers in identifying treatments that are feasible for study.

In terms of analytical methods and other related gaps, the project team identified of the following research needs.

- **Exploration and validation of alternate sources of motorcycle volume data**: Given the lack of motorcycle volume data in most States, alternative sources of such data are one area of potential research. For example, crowd-sourced data from vendors who aggregate cell phone locator information are filling gaps for some traffic volume and origin/destination study needs, but to our knowledge, none of the commercial sources have a way to identify specific vehicle classes.

  There are other methods (including specific phone apps) that can fill this gap, but to the research team's knowledge, a motorcycle trip logger has not been developed (or at least not widely promoted) for this purpose.

- **Investigation of alternate route use by motorcyclists and weekday (commuting) versus weekend (recreational) riding patterns**: Motorcycle route choice differs on weekdays and weekends; however, Middleton et al. found that the weekday and weekend spatial correlation between crash locations and traffic volume were essentially the same.[32,33] This means that higher crash frequencies are found where higher total and/or motorcycle traffic volumes are found, regardless of whether the traffic patterns differ on the weekends or weekdays. Because Middleton's study was based on area-wide volume data collected at permanent count locations only, there is some reason to conclude that the area-wide volume differences in weekday versus weekend traffic track well with the changes in motorcycle riders' route selections. Again, this does not answer the question of which exact roads the riders chose, only that their crashes took place in the parts of the State that experienced the most traffic. Further research could help increase our knowledge of route choice and the differences between commuter and recreational trips as well as help determine whether those differences matter in modeling and CMF calculation.

- **Application of case-control study methods**: As discussed in the FHWA *CMF Guidebook*, case-control studies have been used in certain areas of highway safety but rarely focused on the safety effects of geometric design elements.[34] More recently, the case-control method was employed to estimate CMFs for geometric design elements, including lane and shoulder width.[34–36] Case-control studies select sites based on outcome status (e.g., crash or no crash) and then determine the prior risk factor status within each outcome group. Case-control studies assess whether exposure to a potential risk factor is disproportionately distributed between the cases and controls, thereby indicating the likelihood of an actual risk factor. Strictly speaking, case-control studies cannot be used to measure the probability of an event (e.g., crash or severe injury) in terms of expected frequency. However, when locations are most likely to have no crashes or only one crash, the case-control can get closer to the true CMF value. Use of the case-control method may hold promise for studying motorcycle crashes and could be explored.

- **Calibration of existing CMFs for motorcycle crashes**: NCHRP project 17-63 is developing methods for calibrating existing CMFs that are aggregated across crash types, (e.g., for total crashes so that they can be applied in any jurisdiction).[37] The principle is that the overall CMF should change for any site based on its proportion of crash types. For example, in figure 41, the CMF for total crashes is a function of the proportion of two crash types (type 1 and type 2) with CMF values of 0.4 and 1.0, respectively. The method for developing such an equation requires estimates of the aggregated CMF and the proportions of crash types for the data that developed the CMFs. The number of required aggregate CMFs is directly related to the desired number of crash type proportions to be included in the equation.

$$CMF_{total} = 0.4\ (TYPE1_{prop}) + 1.0\ (TYPE2_{prop})$$

Figure 41. Equation. Aggregate CMF calculation.

The development of the method focused on getting site-specific estimates of the total crash SPF. However, in doing so, estimates of the CMF for each crash type were derived.

Such a procedure could potentially be applied for developing CMFs for motorcycle treatments.

- **Investigation of methods for identifying and prioritizing cost-effective opportunities for reducing motorcycle crashes**: Because motorcycle crashes are relatively rare, state-of-the-art network screening methods documented in the *Highway Safety Manual* (HSM) for identifying sites with promise of safety improvements may not be very effective.[38] This difficulty is compounded by the fact that the SPFs needed to apply these methods are not as robust as they are for other types of crashes. The research challenge is to develop a more appropriate methodology for identifying and prioritizing cost-effective opportunities for reducing motorcycle crashes. Lessons can be learned in this regard from approaches used in countries such as New Zealand and the Netherlands, where crash blackspots are largely non-existent or have been eliminated. For example, methods can be focused on identifying routes or corridors for safety improvements rather than individual sites that are the focus of the HSM methods.[38]

- **Exploration of data and methods for the analysis of motor cycle crashes at intersections**: While this study focused on road segments, it is recognized that motorcycle crashes may be over-represented at intersections. Exposure data for this site type is either non-existent or sparse, and research is needed on methods for analyzing motorcycle crashes at intersections in the light of the challenges in obtaining exposure data for intersections. This analysis would also include the identification and prioritization of cost-effective opportunities for reducing motorcycle crashes specifically at intersections (see previous bullet). Research may also be needed on how motorcycle exposure data could be routinely obtained or estimated as part of a jurisdiction's intersection traffic counting program.

- **Application of probability models**: As noted in chapter 2, there has been substantial research on the use of these models to identify factors associated with motorcycle crash severity using a variety of approaches, mainly with ordered and unordered logit and probit specifications. These models have not traditionally been used in conventional safety management applications since they cannot be used directly for estimation of SPFs and CMFs. However, they do have the advantage of not requiring exposure data, and they could potentially be applied to prediction models for total crashes to estimate crash frequency by severity type, similar to the approach used for the freeway crash prediction methodology developed for the HSM.[38,39] With advances in modeling in this area, it may be worth revisiting this modeling approach with a view to pursuing HSM-type applications and exploring the further use of these models to identify opportunities for treatments targeted at specific features.[38] These features can then be subject to detailed engineering investigations at motorcycle crash blackspots. For example, one study cited earlier used probit modeling to find that curve radius is a significant factor influencing injury severity of single-motorcycle crashes with an increase of 1,000 ft in curve radius estimated to decrease the likelihood of fatalities and serious injuries by 0.2 and 0.15 percent, respectively, in a single-motorcycle crash along a curved roadway section. This finding could be used to develop treatments for a motorcycle crash blackspot on a curved roadway section.

ACKNOWLEDGEMENTS

The project team gratefully acknowledges the participation of the following organizations for their assistance in this study:

- FDOT.
- VDOT.
- Pennsylvania Department of Transportation.

REFERENCES

1. Flask, T., Lord, D., and Schneider, W.H. (2014). "A Segment Level Analysis of Multi-Vehicle Motorcycle Crashes in Ohio Using Bayesian Multi-Level Mixed Effects Models." *Safety Science*, *66*, 47–53.

2. French, M.T. and Gumus, G. (2014). "Macroeconomic Fluctuations and Motorcycle Fatalities in the U.S." *Social Science & Medicine*, *104*, 187–193.

3. Haque, M.M., Chin, H.C., and Huang, H. (2010). "Applying Bayesian Hierarchical Models to Examine Motorcycle Crashes at Signalized Intersections." *Accident Analysis and Prevention*, *42*(1), 203–212.

4. Manan, M., Jonsson, T., and Varhelyi, A. (2013). "Development of a Safety Performance Function for Motorcycle Accident Fatalities on Malaysian Primary Roads." *Safety Science*, *60*, 13–20.

5. Kim, K., Kim, S., and Yamashita, E. (2000). *An Analysis of Alcohol Impaired Motorcycle Crashes in Hawaii, 1986–1995*. Transportation Research Board 79th Annual Meeting, Washington, DC.

6. Connor, S. (2014). "Involvement of Unendorsed Motorcycle Operators in Fatal Crashes in Cuyahoga County, Ohio, 2005–2011." *Traffic Injury Prevention*, *15*(5), 508–512.

7. Akaateba, M.A., Amoh-Gyimah, R., and Yakuba, I. (2014). "A Cross-Sectional Observational Study of Helmet Use Among Motorcyclists in Wa, Ghana." *Accident Analysis and Prevention*, *64*, 18–22.

8. Gabauer, D. (2014). "Roadway Characteristics Associated with Motorcycle Crashes into Longitudinal Barriers and the Influence on Rider Injury." *Transportation Research Board 93d Annual Meeting Compendium of Papers*, Transportation Research Board, Washington, DC. (No. 14-0752).

9. Theofilatos, A. and Yannis, G. (2014). "Relationship Between Motorcyclists' Attitudes, Behavior, and Other Attributes with Declared Accident Involvement in Europe." *Traffic Injury Prevention*, *15*(2), 156–164.

10. Keall, M.D., Clark, B., and Rudin-Brown, C. (2013). "A Preliminary Estimation of Motorcyclist Fatal Injury Risk by BAC Level Relative to Car/Van Drivers." *Traffic Injury Prevention*, *14*(1), 7–12.

11. Haworth, N., Smith, R., Brument, I., and Pronk., N. (1997). *Case-Control Study of Motorcycle Crashes*. Report Number CR 174, Federal Office of Road Safety, Canberra Australia.

12. Kim, K. and Boski, J. (2001). "Finding Fault in Motorcycle Crashes in Hawaii: Environmental, Temporal, Spatial, and Human Factors." *Transportation Research Record 1779*, Transportation Research Board, Washington, DC, 182–188.

13. Mannering, F.L. and Grodsky, L.L. (1995) "Statistical Analysis of Motorcyclists' Perceived Accident Risk." *Accident Analysis and Prevention*, *27*(1), 21–31.

14. Shankar, V. and Mannering, F. (1996) "An Exploratory Multinomial Logit Analysis of Single-Vehicle Motorcycle Accident Severity." *Journal of Safety Research*, *27*(3), 183–194.

15. Savolainen, P. and Mannering, F. (2007) "Probabilistic Models of Motorcyclists' Injury Severities in Single- and Multi-Vehicle Crashes." *Accident Analysis and Prevention*, *39*, 955–963.

16. Jung, S., Xiao, Q., and Yoon, Y. (2013). "Evaluation of Motorcycle Safety Strategies Using the Severity of Injuries." *Accident Analysis & Prevention*, *59*, 357–364.

17. Jones, S., Gurupackiam, S., and Walsh, J. (2013). "Factors Influencing the Severity of Crashes Caused by Motorcyclists: Analysis of Data from Alabama." *Journal of Transportation Engineering*, *139*(9), 949–956.

18. Geedipally, S., Turner, P., and Patil, S. (2011). "Analysis of Motorcycle Crashes in Texas with Multinomial Logit Model." *Transportation Research Record 2265*, Transportation Research Board, Washington, DC, 62–69.

19. Barrette, T., Kirsch, T., Savolainen, P., Russo, B., and Gates, T. (2014). *A Disaggregate-Level Assessment of Changes to Michigan's Motorcycle Helmet Use Law on Motorcyclist Injury Outcomes*. Transportation Research Board 93d Annual Meeting, Washington, DC.

20. Ariannezhad, A., Razi-Ardakani, H., and Kermanshah, M. (2014) *Exploring Factors Contributing to Crash Severity of Motorcycles at Suburban Roads*. Transportation Research Board 93d Annual Meeting, Washington, DC.

21. Wang, Z., Lee, C., and Lin, P.S. (2014). *Modeling Injury Severity of Single-Motorcycle Crashes on Curved Roadway Segments*. Transportation Research Board 93d Annual Meeting, Washington, DC.

22. Blackman, R. and Haworth, N. (2013). "Comparison of Moped, Scooter and Motorcycle Crash Risk and Crash Severity." *Accident Analysis and Prevention*, *57*, 1–9.

23. Haque, M., Chin, H., and Debnath, A. (2012). "An Investigation on Multi-Vehicle Motorcycle Crashes Using Log-Linear Models." *Safety Science 50*(2), 352–362.

24. Chin, H. and Haque, M. (2012). "Effectiveness of Red Light Cameras on the Right-Angle Crash Involvement of Motorcycles." *Journal of Advanced Transportation*, *46*(1), 54–66.

25. de Rome, L, Ivers, R., Fitzharris, M., Du, W., Haworth, N., Heritier, S., and Richardson, D. (2011). "Motorcycle Protective Clothing: Protection from Injury or Just the Weather?" *Accident Analysis & Prevention*, *43*(6), 1893–1900.

26. Huang, W.S. and Lai, C.H. (2011). "Survival Risk Factors for Fatal Injured Car and Motorcycle Drivers in Single Alcohol-Related and Alcohol-Unrelated Vehicle Crashes." *Journal of Safety Research*, *42*(2), 93–99.

27. Chung, Y.S. (2013). "Factor Complexity of Crash Occurrence: An Empirical Demonstration Using Boosted Regression Trees." *Accident Analysis & Prevention*, *61*, 107–118.

28. Fridstrom, L., Ifver, J., Ingebrigtsen, S., Kulmala, R., and Thomsen, L.K. (1995). "Measuring the Contribution of Randomness, Exposure, Weather, and Daylight to the Variation in Road Accident Counts." *Accident Analysis and Prevention*, *27*, 1–20.

29. Cambridge Systematics, Inc. (2010). *Counting Motorcycles*. NCHRP Project 08-36, Task 92, Transportation Research Board, Washington, DC, obtained from: http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP08-36(92)_FR.pdf, last accessed May 19, 2016.

30. Middleton, D., Turner, P., Charara, H., Sunkari, S., Geedipally, S., and Scopatz, R. (2013). *Improving the Quality of Motorcycle Travel Data Collection*. NCHRP Report 760, Project 08-81, Transportation Research Board, Washington DC, obtained from: http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_rpt_760.pdf, last accessed May 19, 2016.

31. Klein, L.A., Mills, M.K., and Gibson, D.R.P. (2006) *Traffic Detector Handbook: Third Edition*. Report No. FHWA-HRT-06-108, Federal Highway Administration, Washington, DC, obtained from: http://www.fhwa.dot.gov/publications/research/operations/its/06108/, last accessed May 19, 2016.

32. Middleton, D., Turner, P., Charara, H., Sunkari, S., Geedipally, S., and Scopatz, R. (2014). *Improving the Quality of Motorcycle Travel Data Collection*. NCHRP Project 08-81, Transportation Research Board, Washington, DC.

33. Middleton, D., Geedipally, S., and Scopatz, R. (2014). *A Process to Determine Motorcycle Count Locations*. Paper 14-2988, 93d Annual Transportation Research Board Meeting, Washington, DC.

34. Gross, F., Persaud, B., and Lyon, C. (2010). *A Guide to Developing Quality Crash Modification Factors*. Report No. FHWA-SA-10-032, Federal Highway Administration, Washington, DC, obtained from: http://www.cmfclearinghouse.org/collateral/CMF_Guide.pdf, last accessed May 19, 2016.

35. Gross, F. (2006). *A Dissertation in Civil Engineering: Alternative Methods for Estimating Safety Effectiveness on Rural, Two-Lane Highways: Case-Control and Cohort Methods*. Dissertation, The Pennsylvania State University, State College, PA, obtained from: https://etda.libraries.psu.edu/catalog/7294, last accessed May 19, 2016.

36. Gross, F. and Jovanis, P.P. (2007). “Estimation of the Safety Effectiveness of Lane and Shoulder Width: The Case-Control Approach.” *ASCE Journal of Transportation Engineering*, *133*(6), pp. 362–369.

37. NCHRP. (In publication). *Guidance for the Development and Application of Crash Modification Factors*. NCHRP Project 17-63, Transportation Research Board, Washington, DC.

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

39. Texas Transportation Institute in association with CH2M-Hill. (2012). *Safety Prediction Methodology and Analysis Tool For Freeways and Interchanges*. National Cooperative Highway Research Program, Transportation Research Board, Washington, DC, obtained from: http://onlinepubs.trb.org/onlinepubs/nchrp/docs/NCHRP17-45_FR.pdf, last accessed March 28, 2016.

HRDS-20/10-16(WEB)E