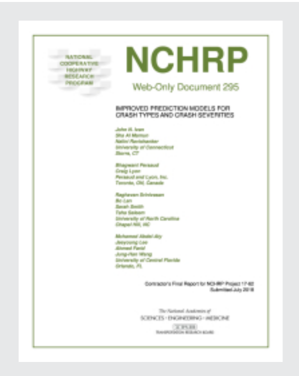


Improved Prediction Models for Crash Types and Crash Severities (2021)

DETAILS

138 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-48959-1 | DOI 10.17226/26164

CONTRIBUTORS

John N. Ivan, Sha Al Mamun, Nalini Ravishanker, University of Connecticut, Bhagwant Persaud, Craig Lyon, Persaud and Lyon, Inc., Raghavan Srinivasan, Bo Lan, Sarah Smith, Taha Saleem, University of North Carolina, and Mohamed Abdel-Aty, Jaeyoung Lee, Ahmed Farid, Jung-Han Wang, University of Central Florida; National Cooperative Highway Research Program; Transportation Research Board; National Academies of Sciences, Engineering, and Medicine

SUGGESTED CITATION

National Academies of Sciences, Engineering, and Medicine 2021. *Improved Prediction Models for Crash Types and Crash Severities*. Washington, DC: The National Academies Press. https://doi.org/10.17226/26164.

NCHRP

Web-Only Document 295:

IMPROVED PREDICTION MODELS FOR CRASH TYPES AND CRASH SEVERITIES

John N. Ivan
Sha Al Mamun
Nalini Ravishanker
University of Connecticut
Storrs, CT

Bhagwant Persaud
Craig Lyon
Persaud and Lyon, Inc.
Toronto, ON, Canada

Raghavan Srinivasan
Bo Lan
Sarah Smith
Taha Saleem
University of North Carolina
Chapel Hill, NC

Mohamed Abdel-Aty
Jaeyoung Lee
Ahmed Farid
Jung-Han Wang
University of Central Florida
Orlando, FL

Contractor's Final Report for NCHRP Project 17-62
Submitted July 2018

NATIONAL COOPERATIVE HIGHWAY RESEARCH PROGRAM

Systematic, well-designed, and implementable research is the most effective way to solve many problems facing state departments of transportation (DOTs) administrators and engineers. Often, highway problems are of local or regional interest and can best be studied by state DOTs individually or in cooperation with their state universities and others. However, the accelerating growth of highway transportation results in increasingly complex problems of wide interest to highway authorities. These problems are best studied through a coordinated program of cooperative research.

Recognizing this need, the leadership of the American Association of State Highway and Transportation Officials (AASHTO) in 1962 initiated an objective national highway research program using modern scientific techniques—the National Cooperative Highway Research Program (NCHRP). NCHRP is supported on a continuing basis by funds from participating member states of AASHTO and receives the full cooperation and support of the Federal Highway Administration (FHWA), United States Department of Transportation, under Agreement No. 693JJ31950003.

The National Academies of
SCIENCES · ENGINEERING · MEDICINE

TRB
TRANSPORTATION RESEARCH BOARD

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The **National Academy of Sciences** was established in 1863 by an Act of Congress, signed by President Lincoln, as a private, non-governmental institution to advise the nation on issues related to science and technology. Members are elected by their peers for outstanding contributions to research. Dr. Marcia McNutt is president.

The **National Academy of Engineering** was established in 1964 under the charter of the National Academy of Sciences to bring the practices of engineering to advising the nation. Members are elected by their peers for extraordinary contributions to engineering. Dr. John L. Anderson is president.

The **National Academy of Medicine** (formerly the Institute of Medicine) was established in 1970 under the charter of the National Academy of Sciences to advise the nation on medical and health issues. Members are elected by their peers for distinguished contributions to medicine and health. Dr. Victor J. Dzau is president.

The three Academies work together as the **National Academies of Sciences, Engineering, and Medicine** to provide independent, objective analysis and advice to the nation and conduct other activities to solve complex problems and inform public policy decisions. The National Academies also encourage education and research, recognize outstanding contributions to knowledge, and increase public understanding in matters of science, engineering, and medicine.

Learn more about the National Academies of Sciences, Engineering, and Medicine at **www.nationalacademies.org**.

The **Transportation Research Board** is one of seven major programs of the National Academies of Sciences, Engineering, and Medicine. The mission of the Transportation Research Board is to provide leadership in transportation improvements and innovation through trusted, timely, impartial, and evidence-based information exchange, research, and advice regarding all modes of transportation. The Board's varied activities annually engage about 8,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation.

Learn more about the Transportation Research Board at **www.TRB.org**.

COOPERATIVE RESEARCH PROGRAMS

CRP STAFF FOR NCHRP WEB-ONLY DOCUMENT 295

Christopher J. Hedges, *Director, Cooperative Research Programs*
Lori L. Sundstrom, *Deputy Director, Cooperative Research Programs*
David Jared, *Senior Program Officer*
Hana Vagnerova, *Senior Program Assistant*
Eileen P. Delaney, *Director of Publications*
Natalie Barnes, *Associate Director of Publications*
Kathleen Mion, *Senior Editorial Assistant*

NCHRP PROJECT 17-62 PANEL
Field of Traffic—Area of Safety

Richard Retting, *Sam Schwartz Consulting, Washington, DC* (Chair)
Peter A. Aiyuk, *Nevada DOT, Carson City, NV*
Larry J. Colclasure, *Maldonado-Burkett Intelligent Transportation Systems, Austin, TX*
Gianluca Dell'Acqua, *Universita degli Studi di Napoli Federico II, Naples*
Reza Jafari, *Road Safety and Transportation Solutions, Inc., Raleigh, NC*
In-Kyu Lim, *Virginia Department of Transportation, Richmond, VA*
Neil M. May, *City of Madison, Madison, WI*
John Milton, *Washington State Department of Transportation, Olympia, WA*
Brian Murphy, *North Carolina Department of Transportation, Garner, NC*
Craig Thor, *FHWA Liaison*
Kelly Hardy, *AASHTO Liaison*
Bernardo Kleiner, *TRB Liaison*

CONTENTS

Appendices A, B, and C are not published here, but are available on the TRB website (www.TRB.org) by searching for ”NCHRP Web-Only Document 295".

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Crash types

ANG	Right angle
ANIMAL	Animal related
BIKE	Bicycle–vehicle
FO	Fixed object
HO	Head-on
HO+SOD	Head-on plus sideswipe opposite direction
ID	Intersecting direction (all severity levels)
LEFT	Left turn
MO	Moving object
MV	Multiple vehicle (all severity levels)
MVD	Multiple-vehicle driveway related
MVN	Multiple-vehicle non-driveway related
MVN OTHER	(MVN minus RE, HO, SSD and SOD)
MVO	Multiple-vehicle other
NIGHT	Nighttime
OD	Opposite direction (all severity levels)
PED	Pedestrian–vehicle
RE	Rear end
RO	Overturn or roll over
ROR	Single-vehicle run-off-road
SD	Same direction (all severity levels)
SOD	Sideswipe opposite direction
SSD	Sideswipe same direction
SV	Single vehicle (all severity levels)
SV FIXEDOBJ	Single-vehicle fixed object
SV OTHER	Single-vehicle other
SV OTHEROBJ	Single-vehicle other object
TID	Turning intersecting direction
TOD	Turning opposite direction
TOT	Total
TSD	Turning same direction

Severity levels

K	Fatal injury
A	Incapacitating injury
B	Non-incapacitating injury
C	Possible injury
O	No injury or property damage only

Facility types

2U	Two-lane undivided
3SG	Three-leg signal-controlled
3ST	Three-leg stop-controlled
3T	Two-lane plus two-way left-turn lane
4D	Four-lane divided
4SG	Four-leg signal-controlled
4ST	Four-leg stop-controlled
4U	Four-lane undivided
5T	Four-lane plus two-way left-turn lane

Variables (with definitions)

AADT	Average annual daily traffic
$AADT_{maj}$	AADT on the major road (higher volume)
$AADT_{min}$	AADT on the minor road (lower volume)
$AADT_{tot}$	AADT (vehicles per day) for minor and major-road combined approaches
Automated Enforcement	Indicates if automated speed enforcement is present
DWYDENS	Number of driveways per mile
FODensity	Fixed object density per mile
Length	Segment length in miles
Lighting	Indicates if lighting is present or not
MajComm	Number of major commercial driveways
MajInd	Number of major industrial driveways
MajRes	Number of major residential driveways
MinComm	Number of minor commercial driveways
MinInd	Number of minor industrial driveways
MinRes	Number of minor residential driveways
MedWidth	Median width in feet
OffsetFO	Average distance from traveled way to fixed objects in feet
OtherDwy	Number of driveways of other type
Parking	Indicates presence of on-street parking

ParkingProp	Proportion of curb length with on-street parking
Parking Type	Indicates angled or parallel on-street parking
Speed Limit	Posted speed limit in miles per hour

Other Abbreviations

AASHTO	American Association of State Highway and Transportation Officials
AIC	Akaike's Information Criterion
BIC	Bayesian Information Criterion
CMF	Crash modification factor
DOT	Department of Transportation
FHWA	Federal Highway Administration
GLM	Generalized linear model
GOF	Goodness of fit
HSIS	Highway Safety Information System
HSM	Highway Safety Manual
MAD	Mean absolute deviation
MSPE	Mean squared prediction error
NB	Negative binomial
OH	Ohio
SPF	Safety performance function

SUMMARY

This report describes efforts to develop improved crash prediction methods for crash type and severity for the three facility types covered in the 2010 Highway Safety Manual (HSM)—specifically, two-lane rural highways, multilane rural highways, and urban/suburban arterials. For each, models were estimated for undivided and divided (multilane rural and urban/suburban only) segments and three- and four-leg stop-controlled intersections and four-leg signal-controlled intersections (also three-leg signal-controlled intersections for urban/suburban arterials). The models use data for segments and intersections with "base conditions" that are defined specifically for each facility type. Only the observations that satisfy the defined base conditions were used for estimating these models. For urban/suburban arterial segments, because no sites met all base conditions for roadside fixed objects and median width, these variables were included in the models only if considered appropriate for that crash type and if the variable was statistically significant in the model and with the expected direction of effect. For some crash types, the number of driveways was also directly included in the models where warranted. These base condition models provide predictions that can be adjusted for actual conditions at a place of prediction, such as lane and shoulder width, the presence of lighting, and other pertinent factors. Content describing these models and instructions for applying them has been prepared for inclusion in the second edition of the HSM. A revisit of the HSM's procedure for calibrating prediction models for transfer to other jurisdictions is also described and recommendations for updating that procedure offered. Average condition models were also estimated using all available valid data points available from the state data used for each facility type; these are provided in an appendix.

For most facility types, crash count models are estimated to predict four aggregated crash types: same-direction, intersecting-direction, opposite-direction, and single-vehicle crashes. For urban/suburban arterial segments, crash count models are estimated for more disaggregated types—for example, rear end, sideswipe-same-direction, combined head-on and opposite-direction sideswipe crashes and night crashes. Models for predicting total crashes were also estimated for all facility types. The count of total crashes may not be equal to the sum of the individual crash type counts because in a few cases there may have been missing variables that would have prevented a crash from being identified as a particular type of crash, but it was still included in total crashes. If a total crash prediction is required, the total crash model should be used rather than adding together all of the crash type models.

For all facility types except urban/suburban arterial segments, crash severity count models are also estimated for each aggregated crash type and for total crashes. For urban/suburban arterial segments, they are estimated for total crashes only. These models are estimated cumulatively—that is, for the following levels:

- Crashes resulting in fatal and incapacitating injuries (KA)
- Crashes resulting in fatal, incapacitating, and non-incapacitating injuries (KAB)
- Crashes resulting in fatal, incapacitating, non-incapacitating, and possible injuries (KABC)
- All severity levels (KABCO)

The sample of fatal injury crashes (K) was too small for all facility types to estimate meaningful models. Average proportions must be computed using local jurisdiction data to allocate KA crashes between K and A if a K crash predictive model is needed, provided the count of these crashes is large enough to estimate a proportion reliably. If a count for a specific crash type and level of severity is required, for example single-vehicle B crashes, the prediction for single-vehicle KA crashes can be subtracted from the prediction for single-vehicle KAB crashes to get a prediction for single-vehicle B crashes.

Also revisited was the procedure for calibrating HSM prediction models for application in jurisdictions other than the locations providing the data used to estimate the safety performance functions (SPFs). The current HSM method was compared with methods proposed by other researchers by calibrating the newly estimated models with data from other jurisdictions. This comparison included evaluating the association between calibration accuracy and calibration sample size, using both a constant calibration factor and an estimated calibration function that relates the factor to the model prediction. The findings suggest the procedure provided by the HSM is still reasonable, although the calibration function yields better accuracy than the constant factor, and sample sizes required for a reliable calibration are sometimes larger than the minimum recommended and can only be iteratively determined. The calibration function could not be estimated with very small sample sizes.

It is noted that estimation and application of crash prediction models is dependent upon having datasets of sufficient size and quality. It was not possible to estimate models for K only crashes for any crash types or in total for any facility type due to the small number of these crashes in any of the data sets. For some crash types, such as same direction crashes, KA crash models also could not be estimated. It is also noted that many of the roadway characteristic variables that are necessary for estimating and applying these models, for example numbers of driveways of different types and intersection skew angles, are not routinely archived by all transportation agencies. For estimation and validation of these models it was necessary to engage in data collection efforts to augment data provided by the transportation agencies that were used in the project. In order to use these prediction procedures, most agencies will likely need to augment their own data archives with additional roadway characteristics.

Appendices to the report provide the following:

1. Documentation of additional models that were estimated—specifically, for average crash models that were estimated using all available data, not just those that met the base conditions
2. Documentation of exploration of a probabilistic approach to predicting crash severity that is not being recommended for prediction
3. Content that has been prepared for inclusion in the second edition of the HSM

1 BACKGROUND

1.1 PROJECT BACKGROUND

The release of the Highway Safety Manual (HSM) by the American Association of State Highway and Transportation Officials (AASHTO) in 2010 was a landmark event in the practice of road safety analysis. Before it, the United States had no central repository for information about quantitative road safety analysis methodology. Consequently, road safety analysts would use methods they were already familiar with or that were easy to locate, which were not necessarily the most appropriate for the analysis context, let alone reflective of the most current knowledge. The HSM provides a single source for methodology and guidance for answering questions about road safety for road segments, intersections, and projects. Numerous state and local road authorities apply HSM methods through the AASHTO lead state initiative.

As revolutionary as it has been for the practice of road safety analysis, it is understood that the 2010 HSM is only a first edition, and room for improvement remains. The various predictive method chapters, for example, offer different approaches for predicting crashes by collision type and severity. Most of these apply aggregate proportions to predictions of total crashes, without accounting for the possibility that the proportion of crashes by type or severity level might be associated with a mixture of predictor variables observed at the location—in particular, traffic volume. Resolving this issue is the basis for this project.

Accurately predicting crashes by collision type and severity is important for the following reasons:

1. Many crash modification factors (CMFs) in the HSM apply only to certain collision types or crashes at certain severity levels. Their proper application requires accurate prediction of the number of crashes of the corresponding collision type and severity level.

2. The safety management methodology in Part B, Chapter 7 of the HSM includes economic evaluation of the expected crash outcomes of road improvement scenarios. These evaluations apply CMFs to improve estimates of crashes without the improvement obtained by applying standardized proportions of different crash types and severity levels to the predicted total crash count by type and severity level. Fully accounting for all of the factors associated with crash type and severity will result in better prediction of these counts and, thus, more accurate economic evaluations and more efficient allocation of scarce safety improvement resources.

3. Collision type and crash severity are usually associated with one another (Golob et al. 1987; Chang and Mannering 1999; Kockelman and Kweon 2002; Zhang *et al.* 2007). Predicting them individually potentially ignores strong associations, leading to less accurate predictions.

1.2 PROJECT OBJECTIVES

The objectives of the project (as defined in the scope of work) are to produce the following:

1. *Crash severity and crash type safety performance functions (SPFs) or distributions or both that can be used in the estimation of the types and severity of crashes likely on the facility types contained or intended for use in the HSM:*

   We present in this report SPFs for crash type and severity.

2. *Recommendations for how the research results can be incorporated into the HSM and associated tools, including the development of associated chapters or chapter content in AASHTO standard format for the HSM second edition, and recommended procedures for consistent use of crash severity and crash type SPFs or distributions or both:*

   Recommendations are provided in the conclusions; draft content for the HSM is provided in Appendix C.

3. *A description of the statistical and practical advantages and disadvantages of the methodology developed in the research and potential barriers to implementation:*

   The description is provided in this report.

The remainder of the report is organized as follows:

- Section 2 provides an overview of our modeling approach common to all facility types.
- Section 3 provides the results of the work on two-lane rural highways.
- Section 43.2.2 provides the results of the work on multilane rural highways.
- Section 5 provides the results of the work on urban and suburban arterials.
- Section 6 provides the results of the work on calibration and validation of all models.
- Section 7 provides conclusions and recommendations about how to incorporate the report findings into the HSM.

2 Analysis Approaches

2.1 Scope of Report

We report here two types of crash frequency models by crash type and crash severity.

Base condition models are estimated using only sites that meet the "base condition" and include only traffic volume as an explanatory variable; these models support the HSM Part C predictive methodology.

Average condition models are estimated using all sites and contain exposure-related variables, such as average annual daily traffic (AADT) and driveways; they apply for average conditions of non-exposure variables.

For most facility types, we report **base condition models** to keep these models compatible with the methodology of the current HSM. For a few facility types, we needed to relax some of the base condition definitions to achieve a large enough sample size to estimate significant models. For a few facility types, the total sample size was much smaller, so we had to use all cases to estimate significant models; we report **average condition models** for these facility types, as well as for the rest of the facility types in Appendix A.

This report does not contain probabilistic crash severity models or models that include both exposure and non-exposure covariates. As will be discussed later, our efforts to estimate these types of models were unsuccessful. This section of the report documents our crash type definitions, our estimation approach for crash count models, our exploration of probabilistic crash severity models, and our exploration of improvements for the model calibration procedure.

2.2 Crash Type Definitions

2.2.1 Crash Types

The selection of crash types for which models would be developed was based on several criteria:

1. The crash types included in the current HSM chapters for which proportions of total crashes are provided
2. The crash types identifiable from electronic crash records in the datasets used for the project
3. The crash types represented in the estimation and validation datasets
4. The crash types to which available CMFs in the HSM apply for each site type

While we tried to maintain consistency of the crash types estimated among all facility types, consideration of these criteria did result in some differences in the final array of crash type models from one facility type to another.

Note that models for pedestrian and bicycle crashes have not been estimated due to very small sample sizes in the available data. These crash types may still be analyzed using the existing HSM approach.

Note also that animal collisions are not included in any of the crash types (they are most likely to be identified as single-vehicle crashes). Our rationale for the omission is that animal crashes have more to do with environmental factors than road characteristics. Since the HSM predictive methods are focused more on providing guidance for selecting safety treatments or predicting expected crash counts related to road characteristics, it is not clear how models predicting animal collisions would fit into the model framework. We note the existence of a large body of research into animal–vehicle collisions and suggest that body of work be consulted for consideration of this collision type in safety management procedures.

We have defined the crash types shown in Figure 2-1 to estimate models:

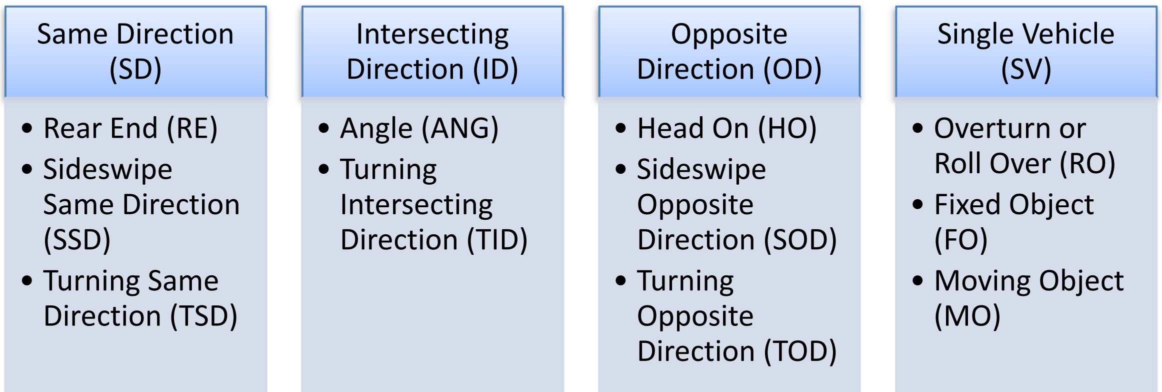


Figure 2-1: General Taxonomy of Crash Types

The taxonomy shown in Figure 2-1 provides for several levels of disaggregation of the crash types according to the number of vehicles involved, their direction of travel, and the manner of the collision. The justification for creating these categories is as follows:

- Each crash type within each category involves vehicles colliding in the same way—that is, front to front, front to rear, front to side, and so on. This results in similar crash severity profiles, as confirmed by Zhang et al. (2007).
- Each crash type within each category is associated with a similar distribution of contributing factors, as assigned by investigating officers (Zhang et al. 2007). This suggests common covariates and exposure functions for these associated collision types.
- Single-vehicle and opposite-direction crashes have very different relationships with exposure (Ivan 2004), so while their collision patterns and contributing factors are similar, they could have very different model forms.
- Experience with crash type prediction suggests that splitting the crash count into too many categories cripples the estimation process, as the crash count for each type gets smaller and smaller. The aggregation categories defined here permit finding a balance that maximizes differences in crash severity and likely causal factors between groups and minimizes them within groups.

The data did not support successful estimation of models for all of these crash types for each facility type, such that coefficients on the AADT variables were not significant or received negative coefficients, there were insufficient numbers of observed crashes or the models did not converge. Also, for the urban/suburban segment models, multiple-vehicle crashes were classified as "driveway related" (MVD) and "multiple-vehicle non-driveway other" (MVN). In these cases, MVD included the following subtypes:

turning same direction (TSD), all intersecting direction (ID) types, and turning opposite direction (TOD). MVN included rear end (RE), head-on (HO), sideswipe same direction (SSD), sideswipe opposite direction (SOD), and MVN other (that is, crashes coded as parked vehicle or angle, though not at driveways or intersections). In addition to the above taxonomy, we estimated nighttime crashes (Night) for some facility types (Urban/suburban segments). Table 2-1 lists the base condition crash type models that were estimated for each facility type.

Table 2-1: Base Condition Crash Type Models Estimated for Each Facility Type

Facility Type	MVD	MVN	MVN OTHER	SD	RE	SSD	ID	OD	HO	HO + SOD	SV	NIGHT
Two-lane rural												
2U				X				X			X	
3ST				X			X	X			X	
4ST				X			X	X			X	
4SG				X			X	X			X	
Multilane rural												
4U				X			X	X			X	
4D				X			X	X			X	
3ST				X			X	X			X	
4ST				X			X	X			X	
4SG				X			X	X			X	
Urban/Suburban arterials												
2U	X	X	X		X	X				X		X
3T	X	X	X		X	X				X	X	X
4U	X	X	X		X	X				X	X	X
4D	X	X	X		X	X				X	X	X
5T	X	X	X		X	X				X	X	X
3ST				X			X	X			X	
4ST				X			X	X			X	
3SG				X			X	X			X	
4SG				X			X	X			X	

Notes: *Facility type codes*—2U = two-lane undivided segments; 3T = two-lane segments with two-way left-turn lane; 4U = four-lane undivided segments; 4D = four-lane divided segments; 5T = four-lane segments with two-way left-turn lane; 3ST = 3 leg stop-controlled intersections; 4ST = four-leg stop-controlled intersections; 3SG = three-leg signal-controlled intersections; 4SG = four-leg signal-controlled intersections.
Crash type codes—MVD = multiple-vehicle driveway related; MVN = multiple-vehicle non-driveway related; MVN OTHER = multiple-vehicle other; SD = same direction (all severity levels); RE = rear end; SSD = sideswipe same direction; ID = intersecting direction; OD = opposite direction (all severity levels); HO = head-on; HO+SOD = sideswipe + opposite direction; SV = single vehicle (all severity levels); NIGHT = nighttime.

2.2.2 Delineation of Intersection Versus Segment Crashes

In the HSM methodology, roadway segment models are used to predict all crashes that occur on portions of roadway segments that are more than 250 feet from an intersection and non-intersection-related crashes that occur on portions of roadway segments that are within 250 feet of an intersection. Intersection models are used to predict all intersection and intersection-related crashes that occur within 250 feet of the intersection. The models for two-lane rural roads and for urban and suburban and suburban arterials apparently were developed to facilitate this application directly.

For multilane rural roads in states where the crash records do not indicate "intersection" or "intersection-related," all crashes occurring within 250 feet of the middle of an intersection are assigned to that intersection. The calibration procedure is expected to allow models developed for such cases to be applied to cases specified in the HSM methodology, and vice versa.

These models were developed to be as consistent with the HSM methodology as possible. In the Ohio database used for urban and suburban arterials and the California database used for multilane rural roads, however, crashes cannot reliably be identified as intersection or intersection-related. Thus, the intersection models being developed for those two databases and facility types will pertain to *all* crashes occurring within 250 feet of the center of an intersection, and the segment models will apply to crashes occurring outside this boundary. As noted previously, the calibration procedure will allow these models to apply to cases where intersection and intersection-related crashes can be identified in accordance with the HSM methodology.

2.3 Model Estimation Approach

2.3.1 Crash Count Models

Because crash frequency is a count phenomenon, negative binomial (NB) regression models, or other count distribution estimation methods, are commonly used to build crash prediction models. Even though the NB model has some limitations (for example, it cannot overcome potential underdispersion problems, and the dispersion parameter may be biased for small sample sizes), this model is still the one most commonly used in univariate crash frequency data analysis. The NB model also provides the dispersion parameter that is required for the empirical Bayes weighting of model predictions and observed crashes in the HSM. In this research, the NB model has been applied for all count models developed.

The NB model, also called the Poisson-Gamma model, is well known to be able handle the issue of overdispersion in count data, where the variance exceeds the mean in violation of the definition of the Poisson distribution. In the NB model, the mean parameter for each site, *i*, is

$$\lambda_i = f(\beta X_i) \times \exp(\varepsilon_i) \quad (2\text{-}1)$$

where ε_i is a gamma-distributed disturbance term, X_i is a vector of explanatory variables, and β is a vector of estimable parameters (coefficients on X_i). The most common relationship between the explanatory variables and λ_i is

$$f(\beta X_i) = \exp(\beta X_i) \text{ or } \ln[f(\beta X_i)] = \beta X_i. \quad (2\text{-}2)$$

With this form, the relationship is also called a log-linear model. One reason the log-linear model is popular for counts is that it ensures the dependent variable (that is, the expected number of crashes during a certain time period) is always positive or zero. Another reason is that taking the log of both sides of the equation results in a linear combination of the predictor variables (that is, the *X*'s) on the right-hand side. This model form belongs to a category called generalized linear models (GLMs). In a GLM, the regression coefficients and their standard errors are typically estimated by maximizing the likelihood or log likelihood of the parameters for the data observed.

The variance of the NB model can be estimated as

$$\mathrm{VAR}[y_i] = \mathrm{E}[y_i] + \alpha(\mathrm{E}[y_i])^2, \quad (2\text{-}3)$$

where *y* is the crash frequency data and *α* is the dispersion parameter.

2.3.2 Alternatives for Model Form

SPFs for roadway segments are formulated as

$$N = exp[b_0 + b_1 \times \ln(AADT) + \ln(L)] \quad (2\text{-}4)$$

where

N = expected average crash frequency per year for a roadway segment;
$AADT$ = annual average daily traffic (vehicles per day) on a roadway segment;
L = length of roadway segment (miles); and
b_0, b_1 = regression coefficients.

The value of the overdispersion parameter associated with *N* is determined as a function of segment length for two-lane and multilane rural facility segments as follows:

$$k = {}^{1}\!/_{exp[c + \ln(L)]} \quad (2\text{-}5)$$

The following function was used for the overdispersion parameter for urban/suburban facility segments (except as noted for individual models):

$$k = e^{\alpha_2} L^{\beta_2} \quad (2\text{-}6)$$

For intersections, two alternative functional forms were considered:

$$N = exp\left[b_0 + b_1 \times \ln(AADT_{maj}) + b_2 \times \ln(AADT_{min})\right] \quad (2\text{-}7)$$

and

$$N = exp[b_0 + b_3 \times \ln(AADT_{total})], \quad (2\text{-}8)$$

where

N = base total expected average crash frequency per year for an intersection;
$AADT_{maj}$ = AADT (vehicles per day) for major-road approaches;
$AADT_{min}$ = AADT (vehicles per day) for minor-road approaches;
$AADT_{total}$ = AADT (vehicles per day) for minor- and major-road approaches combined; and
b_0, b_1, b_2, b_3 = regression coefficients.

In this research, only $AADT_{maj}$, $AADT_{min}$ or $AADT_{total}$ were used for exposure for the SPFs, to be consistent with the HSM. Nevertheless, it is possible that different combinations of exposure variables can better explain the number of crashes (Wang et al. 2017). For some facility types, other model forms were used; this is explained in detail in the relevant sections below.

2.3.3 Model Estimation and Fit Statistics

SPFs for all facility types and crash categories were estimated using standard statistical packages, such as SAS®. As indicated above, the negative binomial distribution was used to start. When the negative binomial overdispersion parameter estimated by maximum likelihood (k) is found to be 0, which happened for several intersection models, this indicates a Poisson distribution is more appropriate (IDRE-UCLA, SAS User Guide). We re-estimated the models with a Poisson distribution in those cases and report both models.

In addition to the parameter estimates and standard errors and the overdispersion parameter, the tables also provide the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Both consist of a goodness-of-fit term (log likelihood), along with a penalty to control for overfitting, and this penalty is a function of the number of parameters estimated. With both the AIC and BIC, lower is better. For a discussion of AIC and BIC, readers are referred to Dziak et al. (2012); suffice to say here that BIC provides a larger penalty for the number of parameters. Dziak et al. (2012) indicate that, while the BIC is more likely to lead to a more parsimonious model with some risk of underfitting, the AIC could lead to a model with good future prediction with some risk of overfitting, and the use of AIC versus BIC may depend on the application.

The mean absolute deviation (MAD) gives a measure of the average magnitude of variability of prediction. Smaller values are preferred to larger in comparing two or more competing SPFs. The MAD is the sum of the absolute value of predicted crashes minus observed crashes, divided by the number of sites. The values of predicted and observed crashes are from the calibration data:

$$MAD = \frac{\sum_i |\hat{y}_i - y_i|}{n}, \qquad (2\text{-}9)$$

where

y_i = observed counts;

$\hat{y}_i$ = predicted values from the SPF; and

n = validation data sample size.

The mean squared prediction error (MSPE) is the sum of squared differences between observed and predicted crash frequencies, divided by sample size. MSPE is typically used to assess error associated with a validation or external data set:

$$MSPE = \frac{\sum_i (\hat{y}_i - y_i)^2}{n}, \qquad (2\text{-}10)$$

where

y_i = observed counts;

$\hat{y}_i$ = predicted values from the SPF; and

n = validation data sample size.

Washington et al. (2005) gives guidelines for interpreting fit statistics and evaluating the suitability of crash prediction models.

2.3.4 Crash Severity Modeling

In general, crashes are classified into five severity levels: fatal injury (K); incapacitating injury (A); non-incapacitating injury (B); possible injury (C); and no injury or property damage only (O). Cumulative values of these levels are commonly defined, building from the highest level, *e.g.*, KA indicates K and A level crashes, KAB indicates K, A and B crashes, etc. For analyzing crash severities, the research team considered several methodologies. First, we considered ordered logit and probit models, using each crash as an observation. These models would have been used to split crash counts into categories of severity. In the preliminary results, some roadway geometric characteristics were found to be statistically significant. They showed that higher maximum speed limits and paved shoulders decrease the severity of a crash, whereas wider lanes increase it, which is clearly counterintuitive. Consequently, we suspected omitted variable bias occurred in the models causing these erroneous results, as they did not include individual or crash characteristics (such as driver, passenger, vehicle, and so on), which are usually found most valuable for predicting the severity of individual crashes.

Consequently, we considered an alternative approach to investigating crashes by severity on an aggregate basis. This better suited the available data as well as the implementation context for the HSM, in which prediction by road segment or intersection is required, and demographic information about travelers is not available. Specifically, we considered a fractional split modeling approach, in which the proportion of crashes by severity level is predicted for each segment or intersection. The methodology and modeling results are excerpted from Yasmin et al. (2016) and summarized in Appendix B. The rest of this section summarizes the fractional split approach and our findings and recommendations regarding crash severity prediction.

Traditionally, the transportation safety literature has evolved along two major streams: crash frequency analysis and crash severity analysis. In crash frequency analysis, the focus is on identifying attributes that result in traffic crashes and effective countermeasures to improve the roadway design, and operational attributes are proposed. Crash severity analysis, on the other hand, is focused on examining crash events, identifying factors that affect the outcome, and providing solutions to reduce the consequences—injuries and fatalities—in the unfortunate event of traffic crashes. Recently, research in transportation safety has begun to bridge the gap between crash frequency and crash severity models. Specifically, researchers are examining crash frequency levels by severity while recognizing that, for the same observation record, crash frequencies by different levels of severity are likely to be dependent. Hence, as opposed to adopting the earlier univariate crash frequency models, researchers have developed multivariate models.

In multivariate approaches that are aimed at studying frequency and severity, the impact of exogenous variables is quantified through the propensity component of count models. The main interaction across different severity-level variables is sought through unobserved effects—that is, no interaction of observed effects occurs across the multiple count models. While this might not be a limitation per se, it might be beneficial to evaluate the impact of exogenous variables in a framework that directly relates a single exogenous variable to all severity count variables simultaneously. It is a framework where the observed propensities of crashes by severity level are modeled directly, while also recognizing the inherent ordering of crash severity outcomes.

The fractional split approach is not without limitations. In field data, there are often no crashes for some specific crash severities in a given case—for example, fatal injury crashes. When this happens, such a segment cannot be used for modeling. To avoid cases with zero crashes for any of the severity levels, the research team aggregated roadway segments into extended super-segments (or arterials). To do this, the severity proportions had to be assumed to be consistent over all segments and intersections included in each super-segment, which was not very practical. In addition, once we aggregated the segments, information specific to them was lost. For these reasons, the research team decided not to adopt the fractional split model for predicting crash severity. Instead, we recommend predicting crash severity using count models, as we do for crash type.

2.4 Estimation and Validation Data

Estimating crash prediction models for the HSM requires datasets with adequate size, quality and scope of variables. Very few highway agencies have such data readily available. In order to limit the extent of the project budget expended on data collection, existing data sources were acquired to the extent possible for each facility type. It was also considered to be desirable to use data from the same states as were used to estimate models for the First Edition of the HSM for consistency. Two sources of readily available data were considered:

- The Highway Safety Information System (HSIS). HSIS is a multistate database that contains crash, roadway inventory, and traffic volume data for a select group of states. When HSIS was initially established, participating states were selected based on the quality and quantity of data available, and their ability to merge data from various files. For estimating the prediction models, HSIS data from Washington (two-lane rural segments), Minnesota (two-lane, multilane rural intersections and urban and suburban segments) and California (multilane four-lane divided segments) were used.
- Ohio Department of Transportation (ODOT). Ohio is part of HSIS. However, in addition to the Ohio data that is part of HSIS, Ohio embarked upon a comprehensive project to collect data for implementation of the HSM and graciously provided the data they have assembled.

In order to validate the estimated models it was necessary to have data from at least one more jurisdiction. The above datasets were sufficient for two-lane rural highways, but additional data sources had to be identified and in most cases data elements collected in order to form validation datasets. Table 2-2 lists the source of the data for estimation and validation for segments and intersections for each facility type. The subsequent chapters discuss the datasets in more detail, but a few overall notes about the selection of data are in order at this stage:

- For 4-leg signalized (4SG) intersections on two-lane and multilane rural highways, the Ohio dataset is used for model estimation because it has more cases than the Minnesota dataset. In the First Edition of the HSM, Minnesota data were used to estimate those models. Consequently, the base predictions for these models will be quite different from those made by the First Edition models.
- For four-lane undivided segments on multilane rural highways, only one state (Texas) could provide a useful dataset. Consequently, three years of the data were used for estimation and the fourth year used for validation.

- For four-lane divided segments on multilane rural highways, data from two states are used for validation as all none of the three state databases were as large as would have been preferred, and having two states to validate against helped to better test the resulting models.

Table 2-2: Data Used for Estimation and Validation

Facility Type	Segments Estimation	Segments Validation	Intersections Estimation	Intersections Validation
Two-lane rural highways	Washington	Ohio	3ST: Minnesota 4ST: Minnesota 4SG: Ohio	3ST: Ohio 4ST: Ohio 4SG: Minnesota
Multilane rural highways	4U: Texas (2009-11) 4D: California	4U: Texas (2012) 4D: Illinois & Washington	3ST: Minnesota 4ST: Minnesota 4SG: Ohio	3ST: Ohio 4ST: Ohio 4SG: Minnesota
Urban/suburban arterials	Ohio	Minnesota	Ohio	North Carolina

Notes: *Facility type codes*—2U = two-lane undivided segments; 3T = two-lane segments with two-way left-turn lane; 4U = four-lane undivided segments; 4D = four-lane divided segments; 5T = four-lane segments with two-way left-turn lane; 3ST = 3 leg stop-controlled intersections; 4ST = four-leg stop-controlled intersections; 3SG = three-leg signal-controlled intersections; 4SG = four-leg signal-controlled intersections.

3 MODELS FOR TWO-LANE RURAL HIGHWAYS

3.1 ROADWAY SEGMENTS

3.1.1 Estimation and Validation Data

To predict crash frequency and severity on two-lane rural highways, the research team estimated and validated base condition SPFs for undivided roadway segments. To develop SPFs for undivided segments (2U), we used segment crash and road characteristics data from Washington State (2008–12). To validate 2U SPFs, we used the same kind of data from Ohio (2009–11).

As shown in Table 3-1, some of the variables defining base conditions in the current HSM are not available in the Washington data: driveway density, vertical curvature, lighting, and use of automated speed enforcement. Knowledge of the roads of this facility type in Washington suggests we can safely assume lighting and automated speed enforcement are absent from nearly all of the segments in the database. A total of 361 2U segments meet the HSM base conditions (other than the four missing variables). We used these to estimate base condition models. Table 3-2 and Table 3-3, respectively, present descriptive statistics for the base condition SPFs and their validation datasets. For validating 2U SPFs, only 21 segments meet the HSM base conditions with shoulder width of six feet; therefore, we used shoulder width ranging from four to seven feet in our dataset to represent the base condition sites. We found a total of 321 segments with this relaxed shoulder width and used them for validating the 2U SPFs.

Table 3-1: HSM Base Conditions and Data Availability, Two-Lane Undivided (2U) Segments

Base Condition	HSM Base Condition	Available in Washington Data?	Available in Ohio Data?
Lane width	12 feet	YES	YES
Shoulder width	6 feet	YES	YES
Shoulder type	Paved	YES	YES
Roadside hazard rating	3	YES	YES
Driveway density	5/mile	NO	YES
Horizontal curvature	None	YES	YES
Vertical curvature	None	NO	NO
Centerline rumble strips	None	YES	YES
Passing lanes	None	YES	YES
Two-way left-turn lanes	None	YES	YES
Lighting	None	NO	YES
Automated speed enforcement	None	NO	YES
Grade level	0%	YES	NO

Table 3-2: Descriptive Statistics for Base Condition SPF Estimation, Two-Lane Undivided (2U) Segments

Variable WA (N = 361, 164.19 miles)			Mean	S.D.	Min	Max
Segment length (mi)			0.454	0.528	0.1	5.42
AADT (veh/day)			4,573	4,121	210	21,622
Lane width (ft)			12	0	12	12
Shoulder width (ft)			6	0	6	6
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	996	2.759	3.703	0	31
	KABC	330	0.914	1.583	0	12
	KAB	187	0.518	1.041	0	7
	KA	57	0.158	0.441	0	3
Same direction	KABCO	204	0.565	1.375	0	12
	KABC	79	0.219	0.641	0	4
	KAB	30	0.083	0.340	0	3
	KA	2	0.006	0.074	0	1
Intersecting direction	KABCO	0	0	0	0	0
	KABC	0	0	0	0	0
	KAB	0	0	0	0	0
	KA	0	0	0	0	0
Opposite direction	KABCO	176	0.488	1.216	0	16
	KABC	80	0.222	0.633	0	7
	KAB	55	0.152	0.496	0	5
	KA	31	0.086	0.309	0	2
Single vehicle	KABCO	616	1.706	2.343	0	17
	KABC	171	0.474	0.960	0	7
	KAB	102	0.283	0.694	0	5
	KA	24	0.066	0.281	0	2

Table 3-3: Descriptive Statistics for Base Condition Validation Data, Two-Lane Undivided (2U) Segments

Variable Ohio (N = 321, 131.1 miles)			Mean	S.D.	Min	Max
Segment length (mi)			0.408	0.397	0.1	2.65
AADT (veh/day)			5,162	2,852	490	15,340
Lane width (ft)			12	0	12	12
Shoulder width (ft)			4.623	0.973	4	7
Crash type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	850	2.648	4.235	0	33
	KABC	195	0.607	1.176	0	10
	KAB	146	0.455	0.925	0	6
	KA	34	0.106	0.346	0	3
Same direction	KABCO	115	0.358	0.901	0	8
	KABC	55	0.171	0.486	0	4
	KAB	33	0.103	0.324	0	2
	KA	5	0.016	0.124	0	1
Intersecting direction	KABCO	0	0	0	0	0
	KABC	0	0	0	0	0
	KAB	0	0	0	0	0
	KA	0	0	0	0	0
Opposite direction	KABCO	58	0.181	0.479	0	3
	KABC	32	0.100	0.339	0	2
	KAB	27	0.084	0.300	0	2
	KA	8	0.025	0.175	0	2
Single vehicle	KABCO	652	2.031	3.368	0	22
	KABC	100	0.312	0.691	0	5
	KAB	79	0.246	0.574	0	4
	KA	18	0.056	0.230	0	1

3.1.2 Estimated Models

We estimated SPFs for rural two-lane highway segments as described in Section 2. Again, for convenience, the model form is given by

$$N = exp[b_0 + b_1 \times \ln(AADT) + \ln(L)]\ , \tag{3-1}$$

and the overdispersion parameter is determined by

$$k = {}^{1}\!/_{exp[c + \ln(L)]}. \tag{3-2}$$

Table 3-4 presents the estimated base condition SPFs for rural two-lane segments; estimated coefficient values and standard errors (in parentheses) are shown, along with the estimated dispersion parameter and fit statistics, as defined in Section 2.3.3. All of the estimated coefficients look reasonable, and all but a few are statistically significant. The fit statistics are also within reasonable ranges.

Table 3-4: Base Condition SPFs, Two-Lane Undivided (2U) Segments

Crash Type Washington (N = 361, 164.19 mi.)	Severity	b_0	b_1	c	–2LL	AIC	MAD
Total	KABCO	-7.463 (0.520)	0.927 (0.062)	1.999 (0.166)	1364.6	1370.6	1.722
	KABC	-9.006 (0.798)	0.977 (0.095)	1.479 (0.255)	825.9	831.9	0.831
	KAB	-8.499 (1.003)	0.852 (0.120)	1.100 (0.327)	618.2	624.2	0.574
	KA	-9.853 (1.472)	0.872 (0.172)	2.527# (2.703)	301.7	307.7	0.239
Same direction	KABCO	-15.456 (1.168)	1.658 (0.135)	1.214 (0.292)	580	586	0.583
	KABC	-17.721 (1.684)	1.807 (0.190)	1.326 (0.550)	334.2	340.2	0.283
	KAB	-16.183 (2.313)	1.526 (0.262)	1.355# (1.339)	192.7	198.7	0.133
	KA* (2 crashes)	-17.266 (7.845)	1.341# (0.887)	13.434 (.)	27.6	33.6	0.011
Opposite direction	KABCO	-10.525 (1.230)	1.085 (0.147)	0.636 (0.254)	628.1	634.1	0.594
	KABC	-11.461 (1.573)	1.100 (0.185)	0.582# (0.430)	379.7	385.7	0.318
	KAB	-10.972 (1.842)	0.999 (0.218)	0.228# (0.517)	292.7	298.7	0.234
	KA	-11.190 (2.021)	0.947 (0.235)	30.408 (0.014)	191.3	197.3	0.137
Single vehicle	KABCO	-5.798 (0.572)	0.674 (0.069)	2.005 (0.223)	1120.7	1126.7	1.217
	KABC	-6.582 (0.975)	0.613 (0.117)	1.117 (0.347)	573.4	579.4	0.520
	KAB	-6.919 (1.227)	0.592 (0.148)	0.809 (0.460)	422.2	428.2	0.372
	KA	-10.949 (2.381)	0.899 (0.280)	0.446# (1.254)	166	172	0.118

* Moore-Penrose inverse is used in covariance matrix.
\# Not significant at 90th percentile confidence interval.

3.1.3 Validation of Models

The prediction models for road segments were validated using the Ohio data. Table 3-5 displays the results for each model, including the total observed crashes, the total crashes predicted using these estimated SPFs and the HSM methodology, and two measures of effectiveness, the MAD and the MSPE (as defined in Section 2.3.3). We then calibrated these predictions using the HSM calibration methodology and provided the MAD and MSPE of these values. Finally, we estimated the calibration function defined in Srinivasan et al. (2016) for the dataset for each crash category; the Srinivasan method is described in detail in Section 6.1.1.

The resulting predictions and fit statistics are provided, and the predicted results are reasonable. In general, the calibration function performs slightly better than the calibration factor. Because the number of observed same-direction KA crashes is small, the calibration factor is the highest of all crash categories (3.848, compared to the second highest, 1.749, for SV KABCO), and the calibration function fails to converge. In general, the models calibrate reasonably well for the Ohio data.

Table 3-5: Calibration and Validation of Washington SPFs Using Ohio Data, Two-Lane Undivided (2U) Segments

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM)				Calibration Function (Srinivasan et al. 2016)					
					Calibration Factor	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	850	630.218	1.732	9.906	1.349	850	1.819	9.021	1.392 (0.070)	0.972 (0.064)	0.511 (0.089)	850.076	1.814	8.992
KABC	195	208.075	0.603	0.896	0.937	195	0.588	0.889	0.934[#] (0.088)	0.917 (0.090)	0.411 (0.177)	194.205	0.590	0.895
KAB	146	116.631	0.457	0.574	1.252	146	0.476	0.549	01.212 (0.118)	0.958 (0.100)	0.304 (0.206)	144.783	0.480	0.554
KA	34	35.819	0.169	0.106	0.949	34	0.166	0.106	0.950[#] (0.360)	1.000 (0.182)	0.001 (0.455)	34.038	0.166	0.106
SD KABCO	115	130.238	0.467	0.637	0.883	115	0.446	0.615	0.844[#] (0.144)	0.859 (0.109)	0.967 (0.371)	114.661	0.452	0.603
SD KABC	55	50.739	0.234	0.212	1.084	55	0.241	0.215	0.731[#] (0.245)	0.699 (0.121)	0.473 (0.450)	54.951	0.253	0.203
SD KAB	33	19.586	0.130	0.093	1.685	33	0.151	0.091	1.100[#] (0.411)	0.811 (0.156)	0.000[#] (0.003)	33.049	0.158	0.090
SD KA	5	1.299	0.020	0.016	3.848	5	0.031	0.016	Failed to Converge					
OD KABCO	58	116.777	0.367	0.295	0.497	58	0.256	0.194	0.441 (0.170)	0.806 (0.137)	0.255 (0.425)	57.834	0.180	0.062
OD KABC	32	52.208	0.204	0.115	0.613	32	0.161	0.103	0.493 (0.323)	0.842 (0.180)	0.215 (0.739)	32.008	0.100	0.020
OD KAB	27	35.284	0.156	0.083	0.765	27	0.139	0.081	0.568 (0.415)	0.833 (0.196)	0.000[#] (0.013)	27.046	0.084	0.014
OD KA	8	18.046	0.075	0.032	0.443	8	0.047	0.030	0.494 (0.951)	1.045 (0.373)	0.014[#] (1.017)	8.010	0.025	0.002
SV KABCO	652	372.703	1.505	8.067	1.749	652	1.536	6.425	1.714 (0.064)	1.050 (0.075)	0.558 (0.101)	657.934	1.897	8.420
SV KABC	100	76.806	0.353	0.369	1.302	100	0.373	0.355	1.409 (0.167)	1.077 (0.122)	0.201 (0.215)	100.256	0.315	0.211
SV KAB	79	59.956	0.297	0.258	1.318	79	0.315	0.247	1.436 (0.192)	1.069 (0.127)	0.001[#] (0.220)	78.998	0.260	0.145
SV KA	18	15.132	0.090	0.048	1.190	18	0.097	0.048	1.441 (0.678)	1.075 (0.252)	0.000[#] (0.000)	17.995	0.075	0.033

[#] Not significant at 90th percentile confidence interval.

3.2 INTERSECTIONS

3.2.1 Estimation and Validation Data

SPFs for two-lane rural highway intersections were estimated and validated using data collected from Minnesota and Ohio. The base conditions for three-leg stop-controlled (3ST), four-leg stop-controlled (4ST), and four-leg signal-controlled (4SG) intersections, as shown in Table 3-6, are specified in the current HSM. All of the variables defining base conditions for all three intersection types are available in Minnesota and Ohio. The Minnesota data include 141 base condition 3ST intersection sites and 198 base condition 4ST intersection sites. The Ohio data have total of 2,081 base condition 3ST intersections and 662 base condition 4ST intersections. Minnesota data were used for model estimation and Ohio data for model validation for both the 3ST and 4ST intersections.

Table 3-6: HSM Base Conditions and Data Availability, Two-Lane Intersections

Base Condition (3ST, 4ST, and 4SG)	Criteria	Available in Minnesota Data	Available in Ohio Data
Intersection skew angle	0 degrees (Not applicable for 4SG)	YES	YES
Intersection left-turn lanes	None	YES	YES
Intersection right-turn lanes	None	YES	YES
Lighting	None	YES	YES

Table 3-7 shows the sample sizes for 4SG intersections for the base and modified base conditions. None of the 4SG intersections from the Minnesota data satisfied the base conditions, and only 48 intersections from the Ohio data satisfied the HSM base conditions. We therefore used a modified base condition for lighting (that is, presence of lighting = “Present”) to get a large enough sample for both model estimation and validation. We used data from Ohio on a total of 202 4SG intersections with modified base conditions to estimate the SPFs, and 25 4SG intersections from Minnesota were used for validation.

Table 3-7: Base Condition Criteria and Data Availability, Two-Lane Four-Leg Signal-Controlled (4SG) Intersections

Data	Condition	Lighting	Intersection Left-Turn Lanes	Intersection Right-Turn Lanes	Sample Size
Minnesota	Base condition	Not present	0	0	0
	Modified	Present	0	0	25
Ohio	Base condition	Not present	0	0	48
	Modified	Present	0	0	202

Table 3-8 and Table summarize descriptive statistics for, respectively, the data used to develop models and the data used to validate them for base conditions at 3ST intersections, including the total number of crashes at all intersections. Table 3-10 and Table 3-11 present descriptive statistics for the data used to develop and validate models for base conditions at 4ST intersections. Table 3-12 and Table 3-13 show descriptive statistics for 4SG intersections.

Table 3-8: Descriptive Statistics for Base Condition SPFs, Two-Lane Three-Leg Stop-Controlled (3ST) Intersections

Variable Minnesota (N = 141)			Mean	S.D.	Min	Max
Major AADT (veh/day)			3,033	3,393	308	20,092
Minor AADT (veh/day)			360	451	4	3,064
Total entering vehicles (veh/day)			3,392	3,513	316	20,824
Presence of lighting			None	None	None	None
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			0	0	0	0
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	323	2.291	3.476	0	22
	KABC	114	0.809	1.711	0	15
	KAB	47	0.333	0.808	0	5
	KA	10	0.071	0.308	0	2
Same direction	KABCO	83	0.589	1.942	0	19
	KABC	35	0.248	1.190	0	13
	KAB	12	0.085	0.485	0	5
	KA	3	0.021	0.188	0	2
Intersecting direction	KABCO	39	0.277	0.634	0	4
	KABC	18	0.092	0.357	0	2
	KAB	6	0.064	0.298	0	2
	KA	2	0.021	0.188	0	2
Opposite direction	KABCO	39	0.277	0.728	0	5
	KABC	13	0.128	0.445	0	3
	KAB	9	0.043	0.203	0	1
	KA	3	0.014	0.119	0	1
Single vehicle	KABCO	162	1.149	1.521	0	8
	KABC	48	0.340	0.664	0	3
	KAB	20	0.142	0.407	0	2
	KA	2	0.014	0.119	0	1

Table 3-9: Descriptive Statistics for Base Condition Validation Data, Two-Lane Three-Leg Stop-Controlled (3ST) Intersections

Variable Ohio (N = 2,081)			Mean	S.D.	Min	Max
Major AADT (veh/day)			2,214	1,889	90	15,340
Minor AADT (veh/day)			817	721	19	3,050
Total entering vehicles (veh/day)			3,031	2,280	135	15,845
Presence of lighting			None	None	None	None
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			0	0	0	0
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	736	0.354	0.879	0	12
	KABC	288	0.138	0.446	0	6
	KAB	211	0.101	0.360	0	5
	KA	62	0.030	0.170	0	1
Same direction	KABCO	135	0.065	0.356	0	7
	KABC	54	0.026	0.194	0	4
	KAB	31	0.015	0.136	0	2
	KA	9	0.004	0.066	0	1
Intersecting direction	KABCO	76	0.032	0.187	0	2
	KABC	33	0.019	0.139	0	2
	KAB	24	0.015	0.127	0	2
	KA	3	0.007	0.085	0	1
Opposite direction	KABCO	67	0.037	0.223	0	3
	KABC	39	0.016	0.139	0	2
	KAB	32	0.012	0.111	0	2
	KA	15	0.001	0.038	0	1
Single vehicle	KABCO	403	0.194	0.554	0	5
	KABC	140	0.067	0.290	0	3
	KAB	105	0.050	0.248	0	3
	KA	32	0.015	0.123	0	1

Table 3-10: Descriptive Statistics for Base Condition SPFs, Two-Lane Four-Leg Stop-Controlled (4ST) Intersections

Variable Minnesota (N = 198)			Mean	S.D.	Min	Max
Major AADT (veh/day)			1,842.46	1,419.77	147.00	8,461.40
Minor AADT (veh/day)			395.83	667.09	4.00	4,740.00
Total entering vehicles (veh/day)			2,238.30	1,741.04	196.60	9,912.80
Presence of lighting			None	None	None	None
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			0	0	0	0
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	345	1.742	2.273	0	15
	KABC	123	0.621	1.172	0	10
	KAB	70	0.354	0.804	0	6
	KA	17	0.086	0.346	0	2
Same direction	KABCO	70	0.354	0.626	0	3
	KABC	19	0.096	0.295	0	1
	KAB	10	0.051	0.220	0	1
	KA	3	0.015	0.122	0	1
Intersecting direction	KABCO	107	0.207	0.475	0	2
	KABC	57	0.051	0.220	0	1
	KAB	36	0.015	0.122	0	1
	KA	11	0.000	0.000	0	0
Opposite direction	KABCO	41	0.540	1.285	0	12
	KABC	10	0.288	0.897	0	8
	KAB	3	0.182	0.594	0	5
	KA	0	0.056	0.270	0	2
Single vehicle	KABCO	127	0.641	0.883	0	4
	KABC	37	0.187	0.451	0	2
	KAB	21	0.106	0.309	0	1
	KA	3	0.015	0.122	0	1

Table 3-11: Descriptive Statistics for Base Condition Validation Data, Two-Lane Four-Leg Stop-Controlled (4ST) Intersections

Variable Ohio (N = 662)			Mean	S.D.	Min	Max
Major AADT (veh/day)			2,238	1,565	160	7,740
Minor AADT (veh/day)			987	970	33	4,496
Total entering vehicles (veh/day)			3,225	2,266	270	9,780
Presence of lighting			None	None	None	None
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			0	0	0	0
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	201	0.304	0.946	0	12
	KABC	103	0.156	0.562	0	7
	KAB	84	0.127	0.491	0	7
	KA	24	0.036	0.210	0	2
Same direction	KABCO	30	0.045	0.229	0	2
	KABC	12	0.018	0.144	0	2
	KAB	6	0.009	0.095	0	1
	KA	0	0.000	0.000	0	0
Intersecting direction	KABCO	91	0.014	0.116	0	1
	KABC	63	0.005	0.067	0	1
	KAB	54	0.005	0.067	0	1
	KA	17	0.003	0.055	0	1
Opposite direction	KABCO	9	0.137	0.676	0	10
	KABC	3	0.095	0.484	0	7
	KAB	3	0.082	0.439	0	7
	KA	2	0.026	0.176	0	2
Single vehicle	KABCO	54	0.082	0.320	0	3
	KABC	15	0.023	0.159	0	2
	KAB	14	0.021	0.154	0	2
	KA	4	0.006	0.078	0	1

Table 3-12: Descriptive Statistics for Modified Base Condition SPFs, Two-Lane Four-Leg Signal-Controlled (4SG) Intersections

Variable Ohio (N = 202)			Mean	S.D.	Min	Max
Major AADT (veh/day)			5,344.55	2,740.54	910	14,790
Minor AADT (veh/day)			2,476.67	2,069.78	95	11,641
Total entering vehicles (veh/day)			7,821.22	4,022.69	1,201	24,690
Presence of lighting			Present	Present	Present	Present
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			N/A	N/A	N/A	N/A
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	454	2.248	3.390	0	25
	KABC	108	0.535	1.070	0	8
	KAB	63	0.312	0.724	0	6
	KA	16	0.079	0.305	0	2
Same direction	KABCO	249	1.233	2.299	0	18
	KABC	49	0.243	0.635	0	5
	KAB	22	0.109	0.357	0	2
	KA	4	0.020	0.140	0	1
Intersecting direction	KABCO	137	0.074	0.263	0	1
	KABC	43	0.025	0.156	0	1
	KAB	29	0.020	0.140	0	1
	KA	6	0.010	0.099	0	1
Opposite direction	KABCO	15	0.678	1.293	0	8
	KABC	5	0.213	0.589	0	5
	KAB	4	0.144	0.451	0	4
	KA	2	0.030	0.170	0	1
Single vehicle	KABCO	53	0.262	0.635	0	4
	KABC	11	0.054	0.227	0	1
	KAB	8	0.040	0.196	0	1
	KA	4	0.020	0.140	0	1

Table 3-13: Descriptive Statistics for Modified Base Condition Validation Data, Two-Lane Four-Leg Signal-Controlled (4SG) Intersections

Variable Minnesota (N = 25)			Mean	S.D.	Min	Max
Major AADT (veh/day)			7,780	3,178	2,450	18,525
Minor AADT (veh/day)			3,472	2,692	353	10,500
Total entering vehicles (veh/day)			11,252	5,238	2,803	29,025
Presence of lighting			Present	Present	Present	Present
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			N/A	N/A	N/A	N/A
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	136	5.440	3.536	0	16
	KABC	34	1.360	1.705	0	8
	KAB	11	0.440	0.961	0	4
	KA	1	0.040	0.200	0	1
Same direction	KABCO	68	2.720	2.390	0	9
	KABC	12	0.480	0.963	0	3
	KAB	1	0.040	0.200	0	1
	KA	0	0.000	0.000	0	0
Intersecting direction	KABCO	21	0.840	1.179	0	4
	KABC	9	0.360	0.638	0	2
	KAB	4	0.160	0.374	0	1
	KA	1	0.040	0.200	0	1
Opposite direction	KABCO	33	1.320	1.145	0	3
	KABC	9	0.360	0.569	0	2
	KAB	3	0.120	0.332	0	1
	KA	0	0.000	0.000	0	0
Single vehicle	KABCO	14	0.560	0.821	0	3
	KABC	4	0.160	0.473	0	2
	KAB	3	0.120	0.440	0	2
	KA	0	0.000	0.000	0	0

3.2.2 Estimated Models

We first estimated base condition SPFs for all two-lane rural intersections using NB modeling, as defined in Section 2. For some of these crash type–crash severity combinations, a dispersion factor of 0 was found; for those types, we show the Poisson modeling results as well. For some models, the parameter on AADT_min was not significant, so we estimated and show models for those crash types with total AADT instead. Following are the model forms used (as defined in Section 2):

$$N = exp\left[b_0 + b_1 \times \ln(AADT_{maj}) + b_2 \times \ln(AADT_{min})\right] \quad (3\text{-}3)$$

or

$$N = exp[b_0 + b_3 \times \ln(AADT_{total})] \quad (3\text{-}4)$$

Table 3-14 shows the NB modeling results for base conditions at 3ST intersections, while Table 3-15 shows the Poisson modeling results for the model types with a 0 dispersion factor for these intersections. Some of the base condition models have AADT parameter values that are not significant; this is largely due to the small number of crashes and small sample sizes. Otherwise, the parameter values are all within expected ranges.

Base condition SPFs for 4ST intersections are presented in Table 3-16 and Table 3-17 for base condition models estimated using NB distribution and Poisson distribution, respectively. These results are similar to those for the 3ST intersections and also reasonable.

Modified base condition SPFs for 4SG intersections are presented in Table 3-18 and Table 3-19 for base condition models estimated using NB distribution and Poisson distribution, respectively.

Table 3-14: Base Condition SPFs, Two-Lane Three-Leg Stop-Controlled (3ST) Intersections

Crash Type Minnesota (N = 141)	Severity	b_0	b_1	b_2	b_3	k	–2LL	AIC	MAD
Total	KABCO	-7.924 (0.857)	0.656 (0.099)	0.295 (0.067)	-	0.622 (0.153)	253.6	515.2	1.633
	KABC	-9.628 (1.301)	0.725 (0.144)	0.312 (0.102)	-	0.974 (0.340)	153.9	315.9	0.845
	KAB	-10.241 (1.891)	0.581 (0.214)	0.468 (0.153)	-	1.383 (0.708)	93.2	194.5	0.461
	KA	-11.873 (4.477)	-	-	0.908 (0.548)	5.123 (4.870)	34.1	74.2	0.137
Same direction	KABCO	-15.506 (2.178)	1.291 (0.221)	0.452 (0.141)	-	1.777 (0.600)	108.9	225.9	0.751
	KABC	-18.598 (3.506)	1.569 (0.337)	0.420 (0.231)	-	2.775 (1.370)	60.1	128.2	0.387
	KAB	-16.952 (4.337)	-	-	1.501 (0.506)	5.281 (4.458)	30.8	67.6	0.146
	KA (3 crashes)	-13.794 (5.693)	-	-	0.984[#] (0.665)	0.000[*] (0.063)	14.1	34.2	0.042
Intersecting direction	KABCO	-14.120 (2.365)	0.818 (0.244)	0.753 (0.191)	-	0.995 (0.653)	73.7	155.5	0.353
	KABC	-15.174 (3.164)	0.977 (0.320)	0.583 (0.261)	-	1.583 (1.494)	45.5	99.0	0.208
	KAB (6 crashes)	-13.383 (4.052)	-	-	1.017 (0.472)	0.000[*] (0.005)	22.5	51.0	0.078
	KA (2 crashes)	-10.629 (6.539)	-	-	0.556[#] (0.795)	0.000[*] (0.000)	10.2	26.5	0.028
Opposite direction	KABCO	-11.716 (1.581)	0.746 (0.181)	0.455 (0.133)	-	0.000[*] (0.000)	78.8	165.6	0.340
	KABC	-15.272 (3.605)	1.025 (0.406)	0.476 (0.258)	-	1.415 (1.720)	37.0	82.0	0.146
	KAB (9 crashes)	-15.571 (4.585)	0.371[#] (0.523)	1.301 (0.511)		1.167 (1.626)	26.0	60.1	0.105
	KA (3 crashes)	-12.867 (5.580)	-	-	0.875[#] (0.658)	0.000[*] (0.050)	14.3	34.6	0.042
Single vehicle	KABCO	-5.916 (0.907)	0.409 (0.112)	0.173 (0.076)	-	0.535 (0.203)	198.6	405.3	1.023
	KABC	-5.398 (1.464)	-	-	0.302 (0.183)	0.787 (0.569)	105.7	217.4	0.492
	KAB (20crashes)	-5.854 (2.264)	-	-	0.249[#] (0.285)	1.309 (1.486)	59.9	125.9	0.249
	KA (2 crashes)	-7.515[#] (6.355)	-	-	0.168[#] (0.802)	0.000[*] (0.000)	10.4	26.9	0.028

[*] Poisson distribution used; scale = square root of Deviance/DOF.

[#] Not significant at 90th percentile confidence interval.

Table 3-15: Base Condition SPF Using Poisson Distribution, Two-Lane Three-Leg Stop-Controlled (3ST) Intersections

Crash Type Minnesota (N = 141)	Severity	b_0	b_1	b_2	b_3	*Scale*	–2LL	AIC	MAD
Same direction	KA (3 crashes)	-13.794 (2.346)	-	-	0.984 (0.274)	0.412 (0)	14.1	34.2	0.042
Intersecting direction	KAB (6 crashes)	-13.383 (1.976)	-	-	1.017 (0.230)	0.487 (0)	22.5	49.0	0.078
	KA (2 crashes)	-10.629 (2.255)	-	-	0.556 (0.274)	0.344 (0)	10.2	24.5	0.028
Opposite direction	KABCO	-11.716 (1.307)	0.746 (0.150)	0.455 (0.110)	-	0.826 (0)	78.8	163.6	0.340
	KA (3 crashes)	-12.867 (2.323)	-	-	0.8752 (0.274)	0.416 (0)	14.3	32.6	0.042
Single vehicle	KA (2 crashes)	-7.513 (2.221)	-	-	0.1681[#] (0.280)	0.349 (0)	10.4	24.9	0.028

[#] Not significant at 90th percentile confidence interval.

Table 3-16: Base Condition SPFs, Two-Lane Four-Leg Stop-Controlled (4ST) Intersections

Crash Type Minnesota (N = 198)	Severity	b_0	b_1	b_2	b_3	k	–2LL	AIC	MAD
Total	KABCO	-6.620 (0.805)	0.451 (0.112)	0.339 (0.06)	-	0.435 (0.119)	325.9	659.9	1.308
	KABC	-8.747 (1.306)	-	-	0.825 (0.168)	0.929 (0.312)	200.8	407.7	0.730
	KAB	-8.511 (1.676)	-	-	0.723 (0.217)	1.564 (0.630)	146.5	299.1	0.527
	KA	-10.539 (3.235)	-	-	0.799 (0.416)	4.683 (3.524)	55.6	117.2	0.156
Same direction	KABCO	-7.914 (1.294)	0.364 (0.184)	0.399 (0.101)	-	0.000* (0.001)	138.3	284.6	0.439
	KABC	-7.538 (2.469)	-	-	0.429# (0.320)	0.000* (0.000)	62.6	131.2	0.172
	KAB (10 crashes)	-4.284# (3.337)	-	-	-0.087# (0.448)	0.000* (0.002)	39.8	85.6	0.096
	KA (0 crash)	-	-	-	-	-	-	-	-
Intersecting direction	KABCO	-10.362 (1.320)	0.475 (0.181)	0.722 (0.103)	-	0.415 (0.219)	158.4	324.8	0.568
	KABC	-12.896 (2.284)	-	-	1.248 (0.292)	2.906 (1.094)	115.9	237.8	0.438
	KAB	-12.779 (2.425)	-	-	1.175 (0.306)	2.178 (1.183)	89.1	184.2	0.297
	KA	-15.115 (4.079)	-	-	1.318 (0.508)	3.094 (3.627)	38.3	82.6	0.101
Opposite direction	KABCO	-10.514 (1.776)	0.769 (0.242)	0.224 (0.125)	-	0.000* (0.000)	99.8	207.6	0.303
	KABC	-11.702 (3.572)	-	-	0.881 (0.450)	0.000* (0.002)	37.8	81.7	0.094
	KAB (3 crashes)	-9.979# (6.251)	-	-	0.506# (0.806)	0.000* (0.000)	15.3	36.7	0.030
	KA (2 crashes)	-32.191# (4290)	-	-	-0.117# (57709)	849.22* (0.000)	0	-	-
Single vehicle	KABCO	-5.533 (1.044)	-	-	0.415 (0.136)	0.256 (0.203)	210.2	426.5	0.688
	KABC	-6.412 (1.838)	-	-	0.369# (0.239)	0.439 (0.714)	101.0	208.0	0.310
	KAB (21 crashes)	-6.874 (2.337)	-	-	0.355# (0.304)	0.000* (0.001)	67.4	140.8	0.188
	KA (3 crashes)	-2.405# (6.079)	-	-	-0.508# (0.839)	0.000* (0.000)	15.3	36.7	0.030

*Poisson distribution used; scale = square root of Deviance/DOF.

Not significant at 90th percentile confidence interval.

Table 3-17: Base Condition SPFs Using Poisson Distribution, Two-Lane Four-Leg Stop-Controlled (4ST) Intersections

Crash Type Minnesota (N = 198)	Severity	b_0	b_1	b_2	b_3	*Scale*	–2LL	AIC	MAD
Same direction	KABCO	-7.914 (1.159)	0.364 (0.165)	0.399 (0.090)	-	0.895 (0.000)	138.3	282.6	0.439
	KABC	-7.538 (1.647)	-	-	0.429# (0.213)	0.667 (0.000)	62.6	129.2	0.172
	KAB (10 crashes)	-4.284 (1.841)	-	-	-0.087# (0.247)	0.551 (0.002)	39.8	83.6	0.096
	KA (3 crashes)	-8.564 (2.203)	-	-	0.322# (0.287)	0.357 (0.000)	15.4	34.9	0.030
Opposite direction	KABCO	-10.514 (1.428)	0.769 (0.194)	0.224 (0.101)	-	0.803 (0.000)	99.8	205.6	0.303
	KABC	-11.702 (1.905)	-	-	0.881 (0.240)	0.535 (0.000)	37.8	79.7	0.094
	KAB (3 crashes)	-9.979 (2.221)	-	-	0.506 (0.286)	0.355 (0.000)	15.3	34.7	0.030
Single vehicle	KAB (21 crashes)	-6.874 (1.609)	-	-	0.355 (0.209)	0.688 (0.000)	67.4	138.8	0.188
	KA (3 crashes)	-2.406# (2.160)	-	-	-0.508 (0.298)	0.355 (0.000)	15.3	34.7	0.030

Not significant at 90th percentile confidence interval.

Table 3-18: Modified Base Condition SPFs, Two-Lane Four-Leg Signal-Controlled (4SG) Intersections

Crash Type Ohio (N = 202)	Severity	b_0	b_1	b_2	b_3	k	−2LL	AIC	MAD
Total	KABCO	-8.163 (1.692)	-	-	0.877 (0.189)	1.829 (0.299)	381.1	768.3	2.248
	KABC	-12.337 (2.393)	1.028 (0.280)	0.231 (0.132)	-	1.403 (0.470)	185.9	379.9	0.690
	KAB	-11.059 (2.828)	-	-	0.981 (0.313)	1.376 (0.630)	139.1	284.2	0.459
	KA (16 crashes)	-8.788 (4.898)	-	-	0.578# (0.545)	2.349 (2.592)	56.3	118.6	0.147
Same direction	KABCO	-14.523 (2.184)	-	-	1.509 (0.242)	1.613 (0.340)	277.2	560.5	1.324
	KABC	-15.878 (3.219)	1.242 (0.372)	0.341 (0.176)	-	0.976 (0.632)	111.9	231.8	359
	KAB	-14.740 (4.419)	-	-	1.269 (0.483)	0.831 (1.221)	68.0	142.0	0.188
	KA (4 crashes)	0.422# (7.202)	-	-	-0.623# (0.833)	0.000* (0.003)	19.4	44.8	0.039
Intersecting direction	KABCO	-5.767 (2.209)	-	-	0.480 (0.248)	2.358 (0.571)	220.4	446.9	0.896
	KABC	-11.026 (3.282)	0.675 (0.388)	0.341 (0.199)	-	1.731 (0.926)	108.0	224.1	0.343
	KAB	-10.318 (3.738)	-	-	0.813 (0.414)	1.679 (1.238)	84.7	175.5	0.248
	KA (6 crashes)	-14.890 (7.891)	-	-	1.143# (0.863)	0.000* (0.000)	26.1	58.2	0.057
Opposite direction	KABCO (15 crashes)	-9.404 (2.793)	-	-	0.861 (0.528)	0.000* (0.000)	52.5	111.1	0.135
	KABC (5 crashes)	-13.000# (8.371)	-	-	0.916# (0.921)	0.000* (0.000)	22.9	51.9	0.048
	KAB (4 crashes)	-9.041# (8.695)	-	-	0.452# (0.969)	0.000* (0.012)	19.5	45.1	0.039
	KA (2 crashes)	-7.825# (11.851)	-	-	0.238# (1.329)	0.000* (0.000)	11.2	28.4	0.020
Single vehicle	KABCO (53 crashes)	-5.325 (2.701)	-	-	0.325# (0.303)	2.029 (0.909)	128.6	263.2	0.423
	KABC (11 crashes)	-11.854 (5.610)	-	-	0.876# (0.618)	0.000* (0.001)	41.9	89.8	0.102
	KAB (8 crashes)	-14.053 (6.778)	-	-	1.083# (0.743)	0.000* (0.002)	32.6	71.3	0.075
	KA (4 crashes)	-17.692 (9.993)	-	-	1.405# (1.087)	0.000* (0.008)	18.7	43.5	0.039

* Poisson distribution used; scale = square root of Deviance/DOF.
# Not significant at 90th percentile confidence interval.

Table 3-19: Modified Base Condition SPFs Using Poisson Distribution, Two-Lane Four-Leg Signal-Controlled (4SG) Intersections

Crash Type Ohio (N = 202)	Severity	b_0	b_1	b_2	b_3	Scale	–2LL	AIC	MAD
Same direction	KA (4 crashes)	0.422[#] (2.828)	-	-	-0.623[#] (0.327)	0.392 (0.000)	19.4	42.8	0.039
Intersecting direction	KA (6 crashes)	-14.890 (3.540)	-	-	1.143 (0.387)	0.448 (0.000)	26.1	56.2	0.057
Opposite direction	KABCO (15 crashes)	-11.404 (2.938)	-	-	0.861 (0.323)	0.613 (0.000)	52.5	109.1	0.135
	KABC (5 crashes)	-13.000 (3.547)	-	-	0.916 (0.390)	0.423 (0.000)	22.9	49.9	0.048
	KAB (4 crashes)	-9.041 (3.431)	-	-	0.452[#] (0.382)	0.394 (0.000)	19.5	43.1	0.039
	KA (2 crashes)	-7.825[#] (3.597)	-	-	0.238[#] (0.403)	0.303 (0.000)	11.2	26.4	0.020
Single vehicle	KABC (11 crashes)	-11.854 (3.120)	-	-	0.876 (0.343)	0.556 (0.001)	41.9	87.8	0.102
	KAB (8 crashes)	-14.053 (3.366)	-	-	1.083 (0.369)	0.496 (0.002)	32.6	69.3	0.075
	KA (4 crashes)	-17.692 (3.838)	-	-	1.405 (0.417)	0.384 (0.000)	18.7	41.5	0.039

[#] Not significant at 90th percentile confidence interval.

3.2.3 Validation of Models

The prediction models for all three types of intersections were validated in the same way as the segment models. The 3ST and 4ST models were estimated using Minnesota data and validated using Ohio data, while the 4SG models were estimated using Ohio data and validated using Minnesota data. The results are presented in Table 3-20 through Table 3-22, in the same format as Table 3-5.

As for the segment models, the predicted results for intersections are reasonable. Again, the calibration function generally performs slightly better than the calibration factor. Also, the calibration function fails to converge for many of the crash categories that have very few or no crashes, most commonly categories for KA crashes of various types. For the 3ST and 4ST models, the calibration factors are almost all less than 1.0, and the MAD and MSPE values are within reasonable ranges. Because the estimation and validation data sources are reversed for the 4SG models, the calibration factors are all greater than 1.0, some substantially so. It is noted that a total of only 136 crashes is in this dataset, and a small number of intersections (202 estimation and 25 validation).

Table 3-20: Calibration and Validation of Minnesota SPFs Using Ohio Data, Three-Leg Stop-Controlled (3ST) Intersections

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (CF) (HSM)				Calibration Function (Srinivasan et al. 2016)					
					CF*	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	736	2342.83	1.006	1.643	0.314	736	0.514	0.725	0.330 (0.052)	0.832 (0.074)	2.168 (0.238)	733.442	0.519	0.725
KABC	288	825.61	0.415	0.301	0.349	288	0.234	0.191	0.306 (0.103)	0.819 (0.095)	1.842 (0.413)	286.799	0.236	0.191
KAB	211	416.92	0.253	0.145	0.506	211	0.178	0.126	0.349 (0.173)	0.735 (0.102)	1.595 (0.498)	210.762	0.180	0.126
KA	62	61.76	0.057	0.029	1.004	62	0.057	0.029	0.559[#] (0.680)	0.824 (0.198)	0.000[#] (0.001)	62.053	0.057	0.029
SD KABCO	135	586.58	0.291	0.239	0.230	135	0.116	0.118	0.219 (0.145)	0.949 (0.105)	4.197 (1.095)	132.966	0.116	0.118
SD KABC	54	211.45	0.116	0.056	0.255	54	0.049	0.036	0.213 (0.301)	0.891 (0.139)	5.221 (2.430)	53.532	0.049	0.036
SD KAB	31	54.66	0.040	0.018	0.567	31	0.029	0.018	1.449[#] (0.679)	1.310 (0.224)	0.000[#] (0.505)	31.013	0.029	0.018
SD KA	9	16.94	0.012	0.004	0.531	9	0.009	0.004	Failed to Converge					
ID KABCO	76	420.62	0.218	0.121	0.181	76	0.069	0.050	0.101 (0.244)	0.539 (0.122)	7.292 (2.673)	76.667	0.070	0.049
ID KABC	33	158.80	0.088	0.028	0.208	33	0.031	0.019	0.110 (0.444)	0.707 (0.176)	0.000[#] (0.500)	33.138	0.031	0.019
ID KAB	24	33.57	0.027	0.012	0.715	24	0.023	0.012	Failed to Converge					
ID KA	3	12.17	0.007	0.001	0.247	3	0.003	0.001	Failed to Converge					
OD KABCO	67	322.61	0.172	0.062	0.208	67	0.061	0.034	0.154 (0.295)	0.812 (0.160)	1.584 (1.363)	66.867	0.061	0.034
OD KABC	39	100.26	0.064	0.022	0.389	39	0.036	0.019	0.142 (0.491)	0.626 (0.161)	0.0035 (1.008)	39.138	0.037	0.019
OD KAB	32	142.49	0.080	0.025	0.225	32	0.030	0.016	Failed to Converge					
OD KA	15	17.42	0.015	0.007	0.861	15	0.014	0.007	Failed to Converge					
SV KABCO	403	1122.74	0.562	0.450	0.359	403	0.325	0.302	0.336 (0.105)	0.871 (0.143)	2.552 (0.398)	404.057	0.326	0.302
SV KABC	140	299.21	0.193	0.089	0.468	140	0.126	0.083	1.257[#] (0.786)	1.521 (0.411)	3.319 (1.073)	140.070	0.126	0.083
SV KAB	105	124.80	0.105	0.061	0.841	105	0.096	0.061	5.611[#] (1.562)	1.682 (0.561)	3.791 (1.475)	106.050	0.096	0.061
SV KA	32	12.55	0.021	0.015	2.550	32	0.030	0.015	Failed to Converge					

[*] CF = Calibration Factor; [#] Not significant at 90th percentile confidence interval.

Table 3-21: Calibration and Validation of Minnesota SPFs Using Ohio Data, Four-Leg Stop-Controlled (4ST) Intersections

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM)				Calibration Function (Srinivasan et al. 2016)					
					CF*	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	201	806.693	1.169	2.162	0.249	201	0.503	0.905	0.277 (0.118)	0.739 (0.234)	5.357 (0.915)	201.628	0.508	0.894
KABC	103	239.390	0.437	0.390	0.430	103	0.277	0.319	0.316 (0.324)	0.647 (0.267)	6.222 (1.530)	104.274	0.279	0.315
KAB	84	131.173	0.285	0.252	0.640	84	0.228	0.242	0.356 (0.520)	0.613 (0.300)	5.231 (1.525)	84.700	0.230	0.240
KA	24	32.206	0.081	0.044	0.745	24	0.070	0.044	0.236# (1.293)	0.607# (0.420)	6.493 (4.727)	24.015	0.070	0.044
SD KABCO	30	205.831	0.330	0.258	0.146	30	0.087	0.055	0.084 (0.317)	0.388 (0.177)	3.487 (2.767)	30.286	0.087	0.052
SD KABC	12	31.880	0.064	0.022	0.376	12	0.036	0.021	10.215# (3.544)	2.118 (1.204)	8.030# (9.483)	12.204	0.036	0.021
SD KAB	6	54.202	0.089	0.014	0.111	6	0.018	0.009	Failed to converge					
SD KA	0	Model is not significant							Failed to converge					
ID KABCO	91	351.659	0.590	0.868	0.259	91	0.253	0.471	0.221 (0.307)	0.522 (0.251)	16.108 (3.810)	91.802	0.259	0.458
ID KABC	63	127.534	0.265	0.272	0.494	63	0.179	0.241	0.155 (0.537)	0.249# (0.254)	14.789 (4.247)	63.502	0.180	0.234
ID KAB	54	77.719	0.186	0.204	0.695	54	0.154	0.198	0.129 (0.691)	0.191# (0.271)	10.052 (4.443)	53.848	0.154	0.193
ID KA	17	24.986	0.061	0.032	0.680	17	0.050	0.031	0.093 (1.094)	0.362# (0.306)	9.086 (7.736)	17.119	0.050	0.031
OD KABCO	9	90.192	0.145	0.036	0.100	9	0.027	0.013	0.055 (1.015)	0.678# (0.503)	0.000# (0.000)	8.967	0.027	0.013
OD KABC	3	19.784	0.034	0.005	0.152	3	0.009	0.005	0.067# (3.357)	0.761# (0.970)	0.000# (0.000)	2.965	0.009	0.005
OD KAB	3	5.167	0.012	0.005	0.581	3	0.009	0.005	Failed to converge					
OD KA	2	Model is not significant							Failed to converge					
SV KABCO	54	211.523	0.356	0.166	0.255	54	0.152	0.102	0.170 (0.621)	0.634# (0.526)	3.252 (1.642)	54.027	0.152	0.102
SV KABC	15	60.695	0.110	0.030	0.247	15	0.044	0.025	0.259# (2.488)	1.021# (1.047)	5.015# (6.305)	14.962	0.044	0.025
SV KAB	14	34.180	0.071	0.025	0.410	14	0.041	0.024	1.125# (3.368)	1.345# (1.149)	5.742# (7.100)	14.042	0.041	0.024
SV KA	4	Model is not significant							Failed to converge					

* CF = Calibration Factor; # Not significant at 90th percentile confidence interval.

Table 3-22: Calibration and Validation of OH SPFs using MN Data (4SG)

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM)				Calibration Function (Srinivasan et al. 2016)					
					CF	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	136	26.288	4.468	28.48	5.174	136	2.384	9.039	5.746 (0.105)	0.376 (0.117)	0.087 (0.084)	135.561	2.189	7.967
KABC	34	5.480	1.247	3.994	6.204	34	0.945	2.464	4.162 (0.778)	0.730[#] (0.497)	0.355 (0.297)	33.799	0.980	2.509
KAB	11	3.491	0.489	0.891	3.151	11	0.594	0.690	1.287[#] (0.917)	0.520[#] (0.407)	2.004 (2.006)	10.550	0.631	0.781
KA	1	1.164	0.082	0.038	0.859	1	0.076	0.038	Failed to converge					
SD KABCO	68	9.944	2.401	9.933	6.838	68	1.897	5.751	4.223 (0.195)	0.310 (0.103)	0.156 (0.169)	67.746	1.578	3.780
SD KABC	12	1.407	0.498	1.038	8.532	12	0.634	0.678	Failed to converge					
SD KAB	1	1.019	0.069	0.031	0.981	1	0.068	0.031	Failed to converge					
SD KA	0	Model not significant							Failed to converge					
ID KABCO	33	10.678	1.160	2.093	3.090	33	1.108	1.608	1.253[#] (0.361)	0.052[#] (0.318)	0.000 (0.004)	29.783	0.988	1.278
ID KABC	9	1.073	0.377	0.415	8.384	9	0.505	0.352	Failed to converge					
ID KAB	3	1.780	0.171	0.108	1.685	3	0.207	0.107	Failed to converge					
ID KA	0	Model not significant							Failed to converge					
OD KABCO	21	0.899	0.830	1.493	3.161	21	0.829	1.018	2.082 (0.502)	0.644 (0.336)	0.426 (0.604)	20.746	0.860	1.072
OD KABC	9	Model not significant							Failed to converge					
OD KAB	4	Model not significant							Failed to converge					
OD KA	1	Model not significant							Failed to converge					
SV KABCO	14	4.723	0.574	0.754	2.964	14	0.601	0.580	6.612[#] (1.529)	1.505[#] (0.960)	0.013 (0.512)	14.005	0.575	0.564
SV KABC	4	0.650	0.178	0.228	6.151	4	0.271	0.200	Failed to converge					
SV KAB	3	0.415	0.131	0.192	7.228	3	0.202	0.162	Failed to converge					
SV KA	0	Model not significant							Failed to converge					

[*] CF = Calibration Factor; [#] Not significant at 90th percentile confidence interval.

4 MODELS FOR MULTILANE RURAL HIGHWAYS

4.1 ROADWAY SEGMENTS

4.1.1 Estimation and Validation Data

The data we used for estimation of segment SPFs were collected from Texas (2009–11) and California (2010–11), for rural undivided multilane highway segments and rural divided multilane highway segments, respectively. Since not all the base conditions in the HSM were available, we slightly modified the base conditions and conducted tests to check whether the modified factors were statistically significant in the SPFs. Data used for validation were from Texas (2012) for undivided segments and Illinois (2009–10) and also Washington State (2009–11) for divided segments. Although data were available from North Carolina and Ohio for undivided segments, the samples were limited for SPF development and validation. For divided segments, we used the California data for SPF development because the range of AADTs in California is larger than in Illinois and Washington.

4.1.1.1 Undivided Segments

Not all the base conditions in the current HSM are available in both training (Table 4-1) and validation data. Also, some base conditions are rare in the dataset, such as segments with six-foot shoulder widths.

Table 4-1: HSM Base Conditions and Data Availability, Four-Lane Undivided (4U) Segments

Base Condition	HSM Criteria	Texas
Lane width	12 feet	YES
Shoulder width	6 feet	YES
Shoulder type	Paved	YES
Side slopes	1V:7H or flatter	NO
Lighting	None	NO
Automated speed enforcement	None	YES

In the training data, only 48 divided segments have twelve-foot lanes, six-foot paved shoulders, and no automated speed enforcement. Since estimating all crash types with 48 segments was not possible, we slightly modified the base conditions, changing the shoulder width in the base condition definition from “six feet” to “six feet or wider.” This increased the sample size in the training dataset from 48 to 401, as shown in Table 4-2. Table 4-3 and Table 4 4 present descriptive statistics for, respectively, base condition SPFs and validation data for undivided segments.

Table 4-2: Sample Size by Conditions, Four-Lane Undivided (4U) Segments

Condition	Lane width	Shoulder width	Shoulder type	Auto-speed enforcement	Texas
Base condition	12.ft.	6.ft.	paved	not present	48
Modified condition	12.ft.	6.ft. or wider	paved	not present	401

Table 4-3: Descriptive Statistics for Base Condition SPFs, Four-Lane Undivided (4U) Segments

Texas (N = 401, 176.925 mi)	No. of Crashes	Mean	S.D.	Min	Max
Segment length (ft)	-	0.441	0.626	0.1	5.226
AADT (veh/day)	-	7,193	5,108	250	21,667
Lane width (ft)	-	12	0	12	12
Shoulder width (ft)	-	8.688	1.812	6	17
KABCO	738	1.840	4.952	0	61
KABC	288	0.718	1.920	0	20
KAB	158	0.394	1.111	0	12
KA	56	0.140	0.510	0	5
SD KABCO	287	0.716	2.652	0	35
SD KABC	102	0.254	0.875	0	9
SD KAB	43	0.107	0.419	0	3
SD KA	9	0.022	0.179	0	2
ID KABCO	147	0.367	1.184	0	11
ID KABC	58	0.145	0.569	0	6
ID KAB	32	0.080	0.344	0	3
ID KA	9	0.022	0.179	0	2
OD KABCO	83	0.207	0.889	0	12
OD KABC	40	0.100	0.447	0	5
OD KAB	29	0.072	0.342	0	4
OD KA	22	0.055	0.286	0	3
SV KABCO	192	0.479	1.315	0	15
SV KABC	78	0.195	0.646	0	9
SV KAB	50	0.125	0.529	0	8
SV KA	16	0.040	0.220	0	2

Table 4-4: Descriptive Statistics for Base Condition Validation Data, Four-Lane Undivided (4U) Segments

Texas (N = 402, 170.531 mi)	No. of Crashes	Mean	S.D.	Min	Max
Segment length (ft)	-	0.424	0.596	0.100	5.226
AADT (veh/day)	-	7,022	4,844	90	23,000
Lane width (ft)	-	12	0	12	12
Shoulder width (ft)	-	8.662	1.814	6	17
KABCO	195	0.485	1.399	0	18
KABC	66	0.164	0.581	0	7
KAB	38	0.095	0.362	0	4
KA	14	0.035	0.184	0	1
SD KABCO	78	0.194	0.739	0	10
SD KABC	24	0.06	0.285	0	3
SD KAB	11	0.027	0.178	0	2
SD KA	3	0.007	0.086	0	1
ID KABCO	30	0.075	0.345	0	4
ID KABC	3	0.007	0.086	0	1
ID KAB	3	0.007	0.086	0	1
ID KA	0	0	0	0	0
OD KABCO	15	0.037	0.266	0	4
OD KABC	11	0.027	0.191	0	2
OD KAB	7	0.017	0.131	0	1
OD KA	3	0.007	0.086	0	1
SV KABCO	72	0.179	0.614	0	7
SV KABC	28	0.07	0.339	0	4
SV KAB	17	0.042	0.236	0	3
SV KA	8	0.02	0.14	0	1

4.1.1.2 Divided Segments

Similarly, in the base condition SPFs for divided roadway segments, some base conditions in the current HSM are not available for California, which is the state used for SPF development (see Table 4-5).

Table 4-5: HSM Base Conditions and Data Availability, Four-Lane Divided (4D) Segments

Base Condition	HSM Criteria	California
Lane width	12 ft	X
Right shoulder width	8 ft	X
Median width	30 ft	X
Lighting	none	
Automated speed enforcement	none	X

Moreover, some base conditions occur infrequently in the dataset, such as segments with median widths of 30 feet (0.88 percent). To increase the sample size, we changed the median width in the base condition from “30 feet” to “30 feet or wider.” This increased the sample size in the dataset from 0 to 138, as shown in Table 4 6. Table 4-7 shows descriptive statistics for base condition SPFs for divided segments. Also, descriptive statistics produced from the validation data for the divided segments are presented in Table 4-8 and Table 4-9. We chose California's data for SPF estimation because California has the largest range of AADTs, as shown in Table 4-7; Illinois and Washington data were chosen for validation.

Table 4-6: Sample Size by Condition, Four-Lane Divided (4D) Segments

Condition	Lane Width	Shoulder Width	Median Width	Auto-Speed Enforcement	California
Base condition	12 feet	8 feet	30 feet	Not present	0
Modified condition	12 feet	8 feet	30 feet or wider	Not present	138

Table 4-7: Descriptive Statistics for Base Condition SPFs, Four-Lane Divided (4D) Segments

California (N = 138, 73.366 mi)	No. of Crashes	Mean	S.D.	Min	Max
Segment length (mi)	-	0.532	0.585	0.104	3.65
AADT (veh/day)	-	16,212	10,083	2,325	66,504
Lane width (ft)	-	12	0	12	12
Shoulder width (ft)	-	9.123	0.734	8	11
Median width (ft)	-	68.384	18.904	30	99
KABCO	263	1.906	4.159	0	30
KABC	93	0.674	1.476	0	11
KAB	47	0.341	0.759	0	5
KA	9	0.065	0.248	0	1
SD KABCO	115	0.833	2.536	0	22
SD KABC	30	0.217	0.835	0	7
SD KAB	10	0.072	0.334	0	2
SD KA	0	0	0	0	0
ID KABCO	1	0.007	0.085	0	1
ID KABC	0	0	0	0	0
ID KAB	0	0	0	0	0
ID KA	0	0	0	0	0
OD KABCO	7	0.051	0.251	0	2
OD KABC	5	0.036	0.223	0	2
OD KAB	3	0.022	0.146	0	1
OD KA	3	0.022	0.146	0	1
SV KABCO	136	0.986	2.007	0	19
SV KABC	57	0.413	0.843	0	6
SV KAB	33	0.239	0.561	0	3
SV KA	6	0.043	0.205	0	1

Table 4-8: Descriptive Statistics for Base Condition Validation Data, Four-Lane Divided (4D) Segments - Illinois

Illinois (N = 592, 145.500 mi)	No. of Crashes	Mean	S.D.	Min	Max
Segment length (mi)	-	0.246	0.162	0.100	1.410
AADT (veh/day)	-	8,198	4,376	1,050	27,750
Lane width (ft)	-	12	0	12	12
Shoulder width (ft)	-	10.856	1.239	8	14
Median width (ft)	-	48.785	11.105	30	88
KABCO	170	0.287	0.675	0	6
KABC	70	0.118	0.376	0	3
KAB	59	0.1	0.342	0	3
KA	21	0.035	0.185	0	1
SD KABCO	47	0.079	0.317	0	3
SD KABC	20	0.034	0.199	0	2
SD KAB	18	0.03	0.191	0	2
SD KA	8	0.014	0.116	0	1
ID KABCO	0	0	0	0	0
ID KABC	0	0	0	0	0
ID KAB	0	0	0	0	0
ID KA	0	0	0	0	0
OD KABCO	5	0.008	0.092	0	1
OD KABC	4	0.007	0.082	0	1
OD KAB	4	0.007	0.082	0	1
OD KA	2	0.003	0.058	0	1
SV KABCO	116	0.196	0.513	0	4
SV KABC	44	0.074	0.299	0	2
SV KAB	37	0.063	0.262	0	2
SV KA	11	0.019	0.135	0	1

Table 4-9: Descriptive Statistics for Base Condition Validation Data, Four-Lane Divided (4D) Segments - Washington

Washington (N = 214, 91.727 mi)	No. of Crashes	Mean	S.D.	Min	Max
Segment length (mi)	-	0.429	0.439	0.103	2.373
AADT (veh/day)	-	15,660	8221	3947	42,310
Lane width (ft)	-	12	0	12	12
Shoulder width (ft)	-	9.853	0.525	8	10.5
Median width (ft)	-	59.051	22.833	35	180
KABCO	721	3.369	4.159	0	24
KABC	179	0.836	1.288	0	7
KAB	116	0.542	0.991	0	6
KA	21	0.098	0.342	0	2
SD KABCO	131	0.612	1.212	0	7
SD KABC	36	0.168	0.454	0	2
SD KAB	20	0.093	0.337	0	2
SD KA	3	0.014	0.118	0	1
ID KABCO	20	0.093	0.4	0	3
ID KABC	11	0.051	0.278	0	2
ID KAB	9	0.042	0.262	0	2
ID KA	2	0.009	0.096	0	1
OD KABCO	8	0.037	0.234	0	2
OD KABC	4	0.019	0.136	0	1
OD KAB	2	0.009	0.096	0	1
OD KA	1	0.005	0.068	0	1
SV KABCO	554	2.589	3.329	0	22
SV KABC	126	0.589	0.983	0	6
SV KAB	84	0.393	0.76	0	5
SV KA	15	0.07	0.306	0	2

4.1.2 Estimated Models

Base condition SPFs for undivided segments are displayed in Table 4-10. The coefficients of the traffic volumes range from 0.518 to 1.711, indicating that the relationship between crash occurrences and traffic volume is not necessarily linear. Note that for same-direction, intersecting-direction, and single-vehicle KA crashes, the volume coefficients of the corresponding SPF are not statistically significant at a confidence interval of 90 percent. Such SPFs should be used with care. Also, for the SPF of same-direction KAB crashes, the standard error of the overdispersion parameter function, *c*, is not estimated. Thus, this SPF should also be used with care. Furthermore, the MADs are reasonably small.

Base condition SPFs for divided segments are displayed in Table 4-11. The lowest and highest traffic volume coefficients are –0.176 and 1.730, respectively. They indicate the nonlinear relationship between crashes and traffic volume. The coefficients of the intersecting-direction KABCO crash SPF are all statistically insignificant at all levels of significance because of their extremely large standard errors. Thus, it is not recommended to apply the SPF for practical purposes without care. The SPFs of the remaining crash categories of intersecting-direction crashes and same-direction KA crashes cannot be estimated because there are no observed crashes to model. Finally, the SPF MADs are small.

Table 4-10: Base Condition SPFs, Four-Lane Undivided (4U) Segments

Crash Type Texas (N = 401, 176.925 mi)	Severity	b_0	b_1	c	–2LL	AIC	MAD
Total	KABCO	-9.129 (1.001)	1.055 (0.112)	0.476 (0.130)	1,199.4	1,205.4	1.677
	KABC	-9.6520 (1.2192)	1.0088 (0.1350)	0.611 (0.221)	731.6	737.6	0.711
	KAB	-9.704 (1.447)	0.950 (0.160)	0.783 (0.390)	509.7	515.7	0.453
	KA	-9.799 (2.335)	0.847 (0.259)	-0.2157# (0.5427)	256.1	262.1	0.199
Same direction	KABCO	-13.541 (1.616)	1.431 (0.178)	0.0327# (0.183)	709.0	715.0	0.813
	KABC	-16.6504 (2.2606)	1.654 (0.245)	0.365# (0.365)	374.7	380.7	0.313
	KAB	-15.173 (2.711)	1.404 (0.292)	11.832# (.)	229.2	235.2	0.163
	KA	-12.032 (6.410)	0.895# (0.713)	-1.983 (1.101)	65.6	71.6	0.042
Intersecting direction	KABCO	-10.209 (2.145)	1.000 (0.241)	-0.825 (0.211)	546.1	552.1	0.559
	KABC	-10.944 (2.913)	0.978 (0.325)	-1.199 (0.331)	301.9	307.9	0.243
	KAB	-11.340 (3.227)	0.955 (0.356)	-0.764# (0.567)	197.2	203.2	0.135
	KA	-10.025 (5.702)	0.671# (0.637)	-2.249 (0.921)	76.6	82.6	0.043
Opposite direction	KABCO	-15.344 (2.912)	1.495 (0.321)	-0.923 (0.304)	307.6	313.6	0.290
	KABC	-16.518 (3.174)	1.540 (0.343)	0.365# (0.824)	190.0	196.0	0.147
	KAB	-18.421 (3.572)	1.711 (0.382)	13.203# (224.650)	142.7	148.7	0.108
	KA	-16.573 (3.998)	1.482 (0.431)	0.885# (2.254)	124.0	130.0	0.089
Single vehicle	KABCO	-7.127 (1.196)	0.688 (0.133)	1.018 (0.379)	598.3	604.3	0.532
	KABC	-6.738 (1.558)	0.545 (0.173)	13.202# (121.940)	325.5	331.5	0.253
	KAB	-6.941 (2.044)	0.518 (0.228)	0.476# (0.879)	243.1	249.1	0.184
	KA	-6.931 (3.378)	0.390# (0.379)	-0.255# (1.421)	113.2	119.2	0.071

* Moore-Penrose inverse matrix used.
# Not significant at 90th percentile confidence interval.

Table 4-11: Base Condition SPFs, Four-Lane Divided (4D) Segments

Crash Type California (N = 138, 73.366 mi)	Severity	b_0	b_1	c	−2LL	AIC	MAD
Total	KABCO	-9.644 (1.519)	1.050 (0.156)	0.669 (0.296)	441.2	447.2	1.384
	KABC	-10.817 (1.999)	1.064 (0.203)	1.023 (0.851)	249.2	255.2	0.641
	KAB	-10.69 (2.456)	0.983 (0.248)	2.090 (3.255)	173.6	179.6	0.402
	KA	-7.690 (5.401)	0.508 (0.554)	11.238 (289.120)	57.2	63.2	0.108
Same direction	KABCO	-14.701 (2.920)	1.479 (0.299)	-0.473 (0.210)	282.9	288.9	0.957
	KABC	-18.512 (6.115)	1.730 (0.625)	-1.620 (0.521)	121.8	127.8	0.351
	KAB	-14.914 (9.572)	1.261 (0.983)	-2.190 (0.883)	61.3	67.3	0.130
	KA	0 Observed Crashes: Failed to Converge					
Intersecting direction	KABCO	-192.600# (1360.500)	17.207# (122.530)	9.094# (316.050)	2.200	8.2	0.001
	KABC	0 Observed Crashes: Failed to Converge					
	KAB	0 Observed Crashes: Failed to Converge					
	KA	0 Observed Crashes: Failed to Converge					
Opposite direction	KABCO	-17.478 (5.829)	1.470 (0.575)	9.638# (438.060)	36.7	42.7	0.074
	KABC	-17.132 (7.121)	1.403 (0.707)	1.553# (42.172)	29.8	35.8	0.054
	KAB	-20.211 (8.927)	1.656 (0.874)	9.871# (396.870)	20.0	26.0	0.031
	KA	-20.211 (8.927)	1.656 (0.874)	9.871# (396.870)	20.0	26.0	0.031
Single vehicle	KABCO	-7.990 (1.580)	0.816 (0.161)	1.262 (0.715)	329.4	335.4	0.863
	KABC	-9.473 (2.093)	0.879 (0.212)	10.025# (586.580)	191.6	197.6	0.424
	KAB	-10.952 (2.925)	0.973 (0.296)	1.422# (2.264)	144.2	150.2	0.317
	KA	-1.524 (6.838)	-0.176 (0.719)	9.978# (913.790)	46.1	52.1	0.081

* Moore-Penrose inverse matrix used.

Not significant at 90th percentile confidence interval.

4.1.3 Validation of Models

Calibration and validation of the SPFs of rural multilane undivided and divided segments are presented in Table 4-12 through Table 4-14. The Texas 2012 data are used for calibration and validation of undivided segment SPFs and the Illinois and Washington data for divided segment SPFs. The calibration factors obtained using the HSM method are less than 1, except for SV SPFs. Use of the calibration function (Srinivasan et al. 2016) improves model fit better than the HSM calibration technique in some cases, as indicated by the MADs and MSPEs. Due to the lack of samples, the same-direction KA, intersecting-direction KAB, intersecting-direction KA, and opposite-direction KA crash SPFs cannot be calibrated using the calibration function.

Table 4-12: Calibration of the Texas (2009–11) Safety Performance Functions Using the Texas (2012) Data, Multilane Undivided Segments Base Conditions

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM)				Calibration Function (Srinivasan et al. 2016)					
					CF+	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	195	233.28	0.583	1.371	0.836	195	0.542	1.320	0.825 (0.089)	0.838 (0.084)	1.074 (0.289)	188.96	0.554	1.362
KABC	66	90.49	0.249	0.201	0.729	66	0.215	0.206	0.670 (0.120)	0.936 (0.118)	0.437# (0.392)	63.76	0.217	0.217
KAB	38	50.04	0.154	0.086	0.759	38	0.140	0.088	0.741 (0.183)	0.988 (0.137)	8.95×10^{-7}# (0.013)	37.71	0.140	0.089
KA	14	17.90	0.069	0.032	0.782	14	0.061	0.031	0.447 (0.116)	0.781 (0.224)	0# (0.007)	14.00	0.063	0.031
SD KABCO	78	89.80	0.290	0.452	0.869	78	0.274	0.443	0.751 (0.161)	0.839 (0.118)	1.853 (0.721)	75.81	0.278	0.442
SD KABC	24	31.89	0.105	0.061	0.753	24	0.092	0.061	0.776 (0.265)	1.021 (0.172)	0.047# (0.525)	23.95	0.092	0.061
SD KAB	11	13.71	0.054	0.028	0.803	11	0.048	0.028	1.065# (0.720)	1.117 (0.259)	0.058# (1.125)	11.00	0.048	0.028
SD KA	3	2.96	0.015	0.007	1.015	3	0.015	0.007	Failed to Converge					
ID KABCO	30	48.02	0.170	0.135	0.625	30	0.134	0.118	0.396 (0.172)	0.710 (0.183)	3.074 (1.880)	30.08	0.136	0.113
ID KABC	3	18.78	0.053	0.014	0.160	3	0.015	0.007	0.056 (0.079)	0.596 (0.468)	0 (0)	3.00	0.015	0.007
ID KAB	3	10.25	0.032	0.009	0.293	3	0.015	0.007	Failed to Converge					
ID KA	0	2.90	0.007	1.834×10^{-4}	0	0	0	0	Failed to Converge					
OD KABCO	15	26.73	0.086	0.067	0.561	15	0.065	0.064	0.173# (0.117)	0.495 (0.208)	10.690# (7.997)	14.14	0.068	0.068
OD KABC	11	12.63	0.052	0.032	0.871	11	0.049	0.032	0.238# (0.202)	0.570 (0.233)	5.241# (6.143)	10.40	0.050	0.035
OD KAB	7	9.18	0.036	0.016	0.763	7	0.032	0.016	0.215# (0.213)	0.596 (0.266)	0# (0.004)	7.00	0.033	0.016
OD KA	3	6.98	0.025	0.009	0.430	3	0.015	0.008	Failed to Converge					
SV KABCO	72	61.33	0.247	0.273	1.174	72	0.259	0.266	1.153 (0.301)	0.997 (0.136)	1.299 (0.646)	70.98	0.258	0.266
SV KABC	28	25.01	0.110	0.094	1.119	28	0.115	0.093	1.180 (0.593)	1.035 (0.194)	1.526# (1.355)	27.40	0.113	0.093
SV KAB	17	15.98	0.073	0.047	1.064	17	0.075	0.047	1.231# (0.789)	1.062 (0.227)	0.386# (1.085)	16.76	0.074	0.046
SV KA	8	5.12	0.031	0.018	1.561	8	0.036	0.018	2.672 (1.550)	1.146 (0.314)	0 (0.014)	8.00	0.036	0.018

+ CF = Calibration Factor; # Not significant at 90th percentile confidence interval. * Moore-Penrose inverse matrix used.

Table 4-13: Calibration of the California (2009–10) Safety Performance Functions Using the Illinois (2009–11) Data, Multilane Divided Segments Base Conditions

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM)				Calibration Function (Srinivasan et al. 2016)					
					CF+	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	170	233.745	0.463	0.406	0.727	170	0.413	0.390	0.747 (0.102)	1.046 (0.131)	0.978 (0.343)	169.46	0.411	0.390
KABC	70	82.320	0.213	0.130	0.850	70	0.199	0.130	0.755 (0.252)	0.934 (0.174)	0.624# (0.592)	69.615	0.199	0.130
KAB	59	44.393	0.153	0.109	1.329	59	0.171	0.107	1.638 (0.763)	1.092 (0.192)	0.370# (0.566)	58.761	0.169	0.106
KA	21	12.280	0.054	0.034	1.710	21	0.067	0.033	2.295# (2.462)	1.085 (0.291)	4.379 $\times10^{-6\#}$ (0.058)	20.589	0.066	0.033
SD KABCO	47	75.301	0.170	0.091	0.624	47	0.136	0.089	0.904 (0.312)	1.233 (0.193)	0.569# (0.662)	46.748	0.132	0.088
SD KABC	20	16.978	0.058	0.038	1.178	20	0.063	0.037	0.999# (0.862)	0.951 (0.260)	1.535# (2.066)	19.739	0.063	0.037
SD KAB	18	8.248	0.043	0.035	2.182	18	0.057	0.034	6.391# (8.344)	1.282 (0.334)	1.504# (2.099)	17.682	0.056	0.034
SD KA	8	Failed to Converge			Failed to Converge				Failed to Converge					
ID KABCO	0	Unable to calibrate			Unable to calibrate				Failed to Converge					
ID KABC	0	Failed to Converge			Failed to Converge				Failed to Converge					
ID KAB	0	Failed to Converge			Failed to Converge				Failed to Converge					
ID KA	0	Failed to Converge			Failed to Converge				Failed to Converge					
OD KABCO	5	4.344	0.015	0.008	1.151	5	0.016	0.008	26.667* (4.934)	1.729* (0.063)	$3.84\times10^{-13*}$ (0)	5.000*	0.016*	0.008*
OD KABC	4	3.298	0.012	0.007	1.213	4	0.013	0.007	1.011* (0.243)	0.999* (0.061)	$5.061\times10^{-8*}$ (.)	3.345*	0.012*	0.007*
OD KAB	4	1.569	0.009	0.007	2.550	4	0.013	0.006	2071.8* (7100.31)	2.322* (0.730)	$1.350\times10^{-13*}$ (0.000)	4.000*	0.012*	0.006*
OD KA	2	1.569	0.006	0.003	1.275	2	0.007	0.003	10.039* (0.108)	1.420* (0.151)	$4.426\times10^{-7*}$ (.)	1.658*	0.006*	0.003*
SV KABCO	116	145.206	0.337	0.252	0.799	116	0.308	0.246	0.795 (0.191)	0.991 (0.167)	1.087 (0.470)	116.61	0.309	0.246
SV KABC	44	58.380	0.156	0.089	0.754	44	0.136	0.087	0.463 (0.270)	0.772 (0.248)	2.306# (1.526)	44.003	0.137	0.087
SV KAB	37	31.385	0.107	0.067	1.179	37	0.115	0.067	0.628# (0.448)	0.771 (0.245)	1.091 (1.308)	37.002	0.116	0.067
SV KA	11	13.494	0.040	0.018	0.815	11	0.036	0.018	0.516* (0.242)	0.874* (0.148)	$9.000\times10^{-7*}$ (.)	11.00*	0.036*	0.018*

+ CF = Calibration Factor; # Not significant at 90th percentile confidence interval. * Moore-Penrose inverse matrix used.

Table 4-14: Calibration of the California (2009–10) Safety Performance Functions Using the Washington (2009–11) Data, Multilane Divided Segments Base Conditions

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM, 2010)				Calibration Function (Srinivasan et al. 2016)					
					CF[+]	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	721	437.481	2.001	10.098	1.648	721	1.978	8.385	1.960 (0.142)	0.849 (0.065)	0.376 (0.082)	723.568	1.942	7.708
KABC	179	155.458	0.728	1.165	1.151	179	0.754	1.198	1.161 (0.107)	0.761 (0.092)	0.340 (0.171)	178.321	0.762	1.133
KAB	116	79.609	0.504	0.753	1.457	116	0.553	0.726	1.323 (0.190)	0.842 (0.114)	0.401 (0.239)	115.680	0.554	0.715
KA	21	16.234	0.150	0.108	1.294	21	0.165	0.108	1.481 (0.949)	1.059 (0.270)	0.655 (1.065)	21.163	0.164	0.108
SD KABCO	131	184.025	0.728	1.288	0.712	131	0.629	1.061	0.741 (0.085)	0.890 (0.113)	0.680 (0.267)	131.193	0.633	1.044
SD KABC	36	48.355	0.255	0.172	0.744	36	0.226	0.161	0.700 (0.257)	0.935 (0.207)	5.300×10^{-7}[#] (0.029)	35.943	0.230	0.161
SD KAB	20	17.619	0.139	0.093	1.135	20	0.145	0.091	1.639 (1.307)	1.193 (0.378)	2.15×10^{-7} (0.007)	20.000	0.138	0.089
SD KA	3	Failed to Converge			Failed to Converge				Failed to Converge					
ID KABCO	20	3.883×10^{-4}	0.093	0.168	51506.57	20	0.187	0.736	0.144[#] (0.156)	0.015[#] (0.037)	8.937 (5.005)	20.033	0.175	0.159
ID KABC	11	Failed to Converge			Failed to Converge				Failed to Converge					
ID KAB	9	Failed to Converge			Failed to Converge				Failed to Converge					
ID KA	2	Failed to Converge			Failed to Converge				Failed to Converge					
OD KABCO	8	10.563	0.081	0.056	0.757	8	0.071	0.055	0.300[#] (0.425)	0.647[#] (0.435)	12.599[#] (11.914)	8.035	0.071	0.054
OD KABC	4	7.693	0.051	0.019	0.520	4	0.036	0.018	0.394[#] (2.027)	0.902[#] (1.715)	1.530×10^{-6}[#] (0.219)	4.000	0.036	0.018
OD KAB	2	4.272	0.028	0.009	0.468	2	0.018	0.009	7.468[*] (0.545)	1.937[*] (0.275)	-[*] (.)	2.001[*]	0.018[*]	0.009[*]
OD KA	1	4.272	0.024	0.005	0.234	1	0.009	0.005	1.069[*] (3.836)	1.488[*] (1.184)	4×10^{-5}[*] (.)	1.000[*]	0.009[*]	0.005[*]
SV KABCO	554	234.130	1.838	8.985	2.366	554	1.608	5.539	2.502 (0.160)	0.902 (0.071)	0.351 (0.089)	558.213	1.562	5.324
SV KABC	126	98.006	0.566	0.796	1.286	126	0.601	0.821	1.146 (0.145)	0.739 (0.108)	0.339 (0.203)	126.156	0.611	0.764
SV KAB	84	55.951	0.397	0.494	1.501	84	0.440	0.492	1.227 (0.229)	0.790 (0.124)	0.264[#] (0.240)	84.167	0.448	0.472
SV KA	15	11.346	0.111	0.089	1.322	15	0.124	0.088	1.662[#] (1.658)	1.085 (0.355)	2.784[#] (2.635)	15.225	0.124	0.088

[+] CF = Calibration Factor; [#] Not significant at 90th percentile confidence interval. [*] Moore-Penrose inverse matrix used.

4.2 INTERSECTIONS

4.2.1 Estimation and Validation Data

The data used for estimation of intersection SPFs were collected from Minnesota (2009–11) and Ohio (2009–11) for rural multilane stop-controlled intersections and rural multilane four-leg signal-controlled intersections, respectively. Since not all the base conditions in the HSM were available, we slightly modified the base conditions and conducted tests to check whether the modified factors were statistically significant in the SPFs. For validation, the Ohio data are used for stop-controlled intersections and the Minnesota data for four-leg signal-controlled intersections. For both types of stop-controlled intersections, the sample sizes in the Minnesota data are larger than those in the Ohio data, which prompted us to use the Minnesota data for SPF development, and thus, the Ohio data for validation. For four-leg signal-controlled intersections, the larger number of samples in Ohio than in Minnesota motivated us to use the Ohio data for developing SPFs and the Minnesota data for validation.

The base conditions for three-leg stop-controlled intersections (3ST) and four-leg stop-controlled intersections (4ST) are specified in the current HSM. As shown in Table 4-15, most of the base conditions for intersections in the current HSM are available in Minnesota; an exception is skew angle. Specifically, according to the data description, the type of intersection is interpreted as either skewed or not skewed. Skewed intersections are excluded from the data.

Table 4-15: HSM Base Conditions and Data Availability, Multilane Intersections

Base Condition (3ST and 4ST)	Criteria	Minnesota
Intersection skew angle	0°–5°	NO
Intersection left-turn lanes	None	YES
Intersection right-turn lanes	None	YES
Lighting	None	YES

Table 4-16 summarizes descriptive statistics for base condition SPFs for 3ST intersections; the sample size is 149. The descriptive statistics for 3ST intersections from the Ohio validation data are shown in Table 4-17.

Table 4-18 presents the descriptive statistics for base condition SPFs for 4ST intersections; the sample size is 139. The descriptive statistics for 4ST intersections for the Ohio validation data are shown in Table 4-17.

As no base conditions were defined for 4SG intersections, we needed to define them; they are listed in Table 4-20. The defined base condition for lighting is not consistent with that of the stop-controlled intersections because most four-leg signal-controlled intersections are lit. Also, no information is known regarding the presence of red light–running cameras.

Table 4-21 shows descriptive statistics for base condition SPFs for 4SG intersections; the sample size is 53. The sample size for the validation data, from Minnesota, is only 24. Descriptive statistics for the validation data are shown in Table 4-22.

Table 4-16: Descriptive Statistics for Base Condition SPFs, Multilane Three-Leg Stop-Controlled (3ST) Intersections

Variable Minnesota (N = 149)			Mean	S.D.	Min	Max
Major AADT (veh/day)			11,651	7,759	1,325	36,000
Minor AADT (veh/day)			760	984	3	5,800
Total entering vehicles (veh/day)			12,031	7,730	2,025	36,028
Presence of lighting			0	0	0	0
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	338	2.268	2.426	0	12
	KABC	139	0.933	1.277	0	8
	KAB	62	0.416	0.754	0	4
	KA	10	0.067	0.277	0	2
Same direction	KABCO	85	0.57	0.988	0	6
	KABC	34	0.228	0.534	0	2
	KAB	13	0.087	0.327	0	2
	KA	2	0.013	0.115	0	1
Intersecting direction	KABCO	92	0.617	1.211	0	8
	KABC	50	0.336	0.827	0	6
	KAB	29	0.195	0.541	0	4
	KA	6	0.04	0.229	0	2
Opposite direction	KABCO	10	0.067	0.251	0	1
	KABC	4	0.027	0.162	0	1
	KAB	2	0.013	0.115	0	1
	KA	1	0.007	0.082	0	1
Single vehicle	KABCO	152	1.020	1.500	0	11
	KABC	51	0.342	0.624	0	3
	KAB	18	0.121	0.347	0	2
	KA	1	0.007	0.082	0	1

Table 4-17: Descriptive Statistics for Base Condition Validation Data, Multilane Three-Leg Stop-Controlled (3ST) Intersections

Variable Ohio (N = 117)			Mean	S.D.	Min	Max
Major AADT (veh/day)			8,859	5,485	830	25,000
Minor AADT (veh/day)			1,033	1,231	88	10,450
Total entering vehicles (veh/day)			9,375	5,482	915	25,320
Presence of lighting			0	0	0	0
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew Angle (°)			1.880	1.748	0	5
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	157	1.342	2.146	0	15
	KABC	63	0.538	1.071	0	5
	KAB	46	0.393	0.861	0	5
	KA	15	0.128	0.446	0	3
Same direction	KABCO	61	0.521	1.142	0	7
	KABC	21	0.179	0.448	0	2
	KAB	15	0.128	0.384	0	2
	KA	4	0.034	0.182	0	1
Intersecting direction	KABCO	47	0.402	0.956	0	5
	KABC	23	0.197	0.605	0	3
	KAB	17	0.145	0.478	0	3
	KA	8	0.068	0.253	0	1
Opposite direction	KABCO	12	0.103	0.402	0	3
	KABC	5	0.043	0.203	0	1
	KAB	3	0.026	0.159	0	1
	KA	0	0	0	0	0
Single vehicle	KABCO	37	0.316	0.611	0	3
	KABC	14	0.12	0.375	0	2
	KAB	11	0.094	0.347	0	2
	KA	3	0.026	0.206	0	2

Table 4-18: Descriptive Statistics for Base Condition SPFs, Multilane Four-Leg Stop-Controlled (4ST) Intersections

Variable Minnesota (N = 139)			Mean	S.D.	Min	Max
Major AADT (veh/day)			10,803	6,606	2,422	34,500
Minor AADT (veh/day)			589	629	25	4,654
Total entering vehicles (veh/day)			11,392	6,667	2,499.5	34,583
Presence of lighting			0	0	0	0
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	390	2.806	3.476	0	23
	KABC	165	1.187	1.662	0	12
	KAB	80	0.576	1.028	0	7
	KA	14	0.101	0.439	0	4
Same direction	KABCO	98	0.705	1.207	0	8
	KABC	27	0.194	0.48	0	2
	KAB	12	0.086	0.306	0	2
	KA	1	0.007	0.085	0	1
Intersecting direction	KABCO	151	1.086	2.118	0	17
	KABC	85	0.612	1.299	0	10
	KAB	41	0.295	0.847	0	7
	KA	10	0.072	0.393	0	4
Opposite direction	KABCO	15	0.108	0.334	0	2
	KABC	9	0.065	0.247	0	1
	KAB	5	0.036	0.187	0	1
	KA	1	0.007	0.085	0	1
Single vehicle	KABCO	126	0.906	1.197	0	6
	KABC	44	0.317	0.59	0	2
	KAB	22	0.158	0.404	0	2
	KA	2	0.014	0.120	0	1

Table 4-19: Descriptive Statistics for Base Condition Validation Data, Multilane Four-Leg Stop-Controlled (4ST) Intersections

Variable Ohio (N = 83)			Mean	S.D.	Min	Max
Major AADT (veh/day)			9,853.771	4,799.732	2,690	20,623
Minor AADT (veh/day)			1,320.446	2,542.512	95	20,623
Total entering vehicles (veh/day)			11,174.217	5,949.802	3,300	41,246
Presence of lighting			0	0	0	0
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			1.831	1.576	0	5
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	199	2.398	3.072	0	16
	KABC	83	1.000	1.593	0	8
	KAB	69	0.831	1.387	0	7
	KA	23	0.277	0.786	0	5
Same direction	KABCO	42	0.506	0.875	0	4
	KABC	17	0.205	0.435	0	2
	KAB	12	0.145	0.354	0	1
	KA	2	0.024	0.154	0	1
Intersecting direction	KABCO	96	1.157	2.167	0	12
	KABC	47	0.566	1.241	0	7
	KAB	40	0.482	1.063	0	5
	KA	19	0.229	0.687	0	4
Opposite direction	KABCO	18	0.217	0.47	0	2
	KABC	8	0.096	0.335	0	2
	KAB	8	0.096	0.335	0	2
	KA	1	0.012	0.11	0	1
Single vehicle	KABCO	43	0.518	0.888	0	6
	KABC	11	0.133	0.341	0	1
	KAB	9	0.108	0.313	0	1
	KA	1	0.012	0.11	0	1

Table 4-20: Base Condition Criteria and Data Availability, Multilane Four-Leg Signal-Controlled (4SG) Intersections

Base Condition	Criteria	Ohio
Intersection skew angle	0°–5°	X
Intersection left-turn lanes	None	X
Intersection right-turn lanes	None	X
Red light violation cameras	None	
Lighting	Present	X

Table 4-21: Descriptive Statistics for Base Condition SPFs, Multilane Four-Leg Signal-Controlled (4SG) Intersections

Variable - Ohio (N = 53)			Mean	S.D.	Min	Max
Major AADT (veh/day)			4,686	2,704	880	12,420
Minor AADT (veh/day)			1,902	1,748	157	7,992
Total entering vehicles (veh/day)			6,587	3,158	1,522	14,472
Presence of lighting			1	1	1	1
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			1.264	1.211	0	5
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	249	4.698	4.126	0	16
	KABC	33	0.623	1.078	0	5
	KAB	16	0.302	0.638	0	3
	KA	4	0.075	0.267	0	1
Same direction	KABCO	105	1.981	2.374	0	9
	KABC	14	0.264	0.625	0	3
	KAB	4	0.075	0.331	0	2
	KA	1	0.019	0.137	0	1
Intersecting direction	KABCO	77	1.453	1.927	0	11
	KABC	13	0.245	0.617	0	3
	KAB	8	0.151	0.496	0	3
	KA	1	0.019	0.137	0	1
Opposite direction	KABCO	38	0.717	0.885	0	4
	KABC	4	0.075	0.267	0	1
	KAB	3	0.057	0.233	0	1
	KA	1	0.019	0.137	0	1
Single vehicle	KABCO	29	0.547	0.867	0	4
	KABC	2	0.038	0.192	0	1
	KAB	1	0.019	0.137	0	1
	KA	1	0.019	0.137	0	1

Table 4-22: Descriptive Statistics for Base Condition Validation Data, Multilane Four-Leg Signal-Controlled (4SG) Intersections

Variable Minnesota (N = 24)			Mean	S.D.	Min	Max
Major AADT (veh/day)			11,371	4543	5101	22,468
Minor AADT (veh/day)			3304	1764	417	6649
Total entering vehicles (veh/day)			14,676	4739	7228	24,727
Presence of lighting			Not Available	Not Available	Not Available	Not Available
Presence of left-turn lanes			0	0	0	0
Presence of right-turn lanes			0	0	0	0
Skew angle (°)			Not Available	Not Available	Not Available	Not Available
Crash Type	**Severity**	**No. of Crashes**	**Mean**	**S.D.**	**Min**	**Max**
Total	KABCO	202	8.417	7.994	1	32
	KABC	57	2.375	2.568	0	12
	KAB	16	0.667	0.761	0	2
	KA	2	0.083	0.408	0	2
Same direction	KABCO	96	4	4.314	0	16
	KABC	24	1	1.445	0	5
	KAB	5	0.208	0.415	0	1
	KA	1	0.042	0.204	0	1
Intersecting direction	KABCO	65	2.708	2.662	0	11
	KABC	25	1.042	1.197	0	5
	KAB	10	0.417	0.584	0	2
	KA	1	0.042	0.204	0	1
Opposite direction	KABCO	16	0.667	1.049	0	4
	KABC	5	0.208	0.509	0	2
	KAB	0	0	0	0	0
	KA	0	0	0	0	0
Single vehicle	KABCO	25	1.042	1.654	0	7
	KABC	3	0.125	0.338	0	1
	KAB	1	0.042	0.204	0	1
	KA	0	0	0	0	0

4.2.2 Estimated Models

Base condition SPFs for 3ST intersections are exhibited in Table 4-23. We tried to estimate regression coefficients for both the major- and minor-road traffic volumes. In cases where the minor-road traffic volume was statistically insignificant, we estimated a coefficient for the total entering volume instead. In either case, the coefficients do not indicate the relationship between crashes at 3ST intersections, and entering volumes are linear. In addition, only 10 reported KA crashes are sampled for analysis, rendering all KA crash model results unreliable. Such SPFs should be used with caution. This also applies for the single-vehicle KAB crash SPF, all opposite-direction crash SPFs, and the same-direction KAB crash SPF. The MADs indicate the average deviations between crash counts, predicted by SPFs, and the observed ones are relatively low.

Base condition SPFs for 4ST intersections are presented in Table 4-24. Apart from KABCO, KABC, KAB, same-direction KABCO, and intersecting-direction KABCO crash SPFs, the total entering volume is used as the exposure variable in the SPF development process. This is because statistically insignificant minor-road traffic volumes result when estimating SPFs using both major- and minor-road volumes. Similar to the 3SG intersection KA crash patterns, only 14 reported KA crashes are available in the data, and any KA crash SPF should be used with caution. We also suggest the same-direction KAB crash SPF, the single-vehicle KAB crash SPF, and all opposite-direction SPFs be used only with extra care due to the small samples modeled. The MAD measures indicate low average residuals.

Base condition SPFs for various crash types for 4SG intersections are shown in Table 4-25. In all SPFs, we used the total entering volume, since minor-road volumes are insignificant when using both the major- and minor-road volumes as independent variables. The total entering-volume estimated coefficients range from –0.682 to 1.921. They indicate nonlinear relationships between crashes and entering volume. Yet, the volume is almost linearly correlated with intersecting-direction KABCO crashes, as indicated by the volume coefficient. In addition, we used Moore-Penrose inverse matrices for all KA SPFs, the opposite-direction KABC SPF, the opposite-direction KAB SPF, the single-vehicle KABC SPF, and the single-vehicle KAB SPF, due to inadequate samples. We suggest those SPFs be used with caution. Finally, the average residuals are reasonably low, as indicated by the MADs.

4.2.3 Validation of Models

We conducted calibration and validation of rural multilane intersection SPFs using the Ohio data for 3ST and 4ST intersections and the Minnesota data for the 4SG intersections. The results are presented in tables 4-26 through 4-27. The calibration factors, obtained using the HSM calibration method, are not near 1. The calibration function performs slightly better than the HSM calibration method in a few cases. It should be noted that we used the Moore-Penrose inverse matrix for several SPFs for the severe crash categories due to limited samples.

Table 4-23: Base Condition SPFs, Multilane Three-Leg Stop-Controlled (3ST) Intersections

Crash Type Minnesota (N = 149)	Severity	b_0	b_1	b_2	b_3	k	–2LL	AIC	MAD
Total	KABCO	-9.118 (1.259)	0.776 (0.123)	0.270 (0.053)	-	0.323 (0.100)	548.0	556	1.516
	KABC	-9.392 (1.632)	0.659 (0.160)	0.346 (0.073)	-	0.261# (0.159)	365.6	373.6	0.793
	KAB	-9.208 (2.264)	0.546 (0.221)	0.357 (0.106)	-	0.367# (0.346)	240.6	248.6	0.526
	KA (10 crashes)	2.910# (5.145)	-	-	-0.741# (0.573)	1.913# (3.120)	72.7	78.7	0.125
Same direction	KABCO	-14.411 (2.166)	1.033 (0.209)	0.502 (0.094)	-	0.236# (0.235)	264.2	272.2	0.582
	KABC	-12.552 (3.234)	0.737 (0.315)	0.504 (0.151)	-	0.539# (0.698)	160.1	168.1	0.336
	KAB (13 crashes)	-8.279# (5.067)	-	-	0.510# (0.542)	3.000# (3.15)	88.9	94.9	0.161
	KA (2 crashes)	-9.971# (13.798)	-	-	0.491# (1.461)	5.32×10^{-6}# (0.609)	21.0	27	0.027
Intersecting direction	KABCO	-12.652 (2.242)	0.746 (0.215)	0.651 (0.110)	-	0.602 (0.321)	275.5	283.5	0.658
	KABC	-14.356 (2.789)	0.728 (0.268)	0.833 (0.150)	-	0.435# (0.343)	183.5	191.5	0.394
	KAB	-13.058 (3.304)	0.575 (0.320)	0.774 (0.182)	-	0.365# (0.540)	135.5	143.5	0.272
	KA (6 crashes)	4.137# (8.354)	-	-	-0.934# (0.928)	8.496# (10.579)	48.1	54.1	0.079
Opposite direction	KABCO (10 crashes)	-6.978# (8.480)	-	-	0.345# (0.902)	9.029×10^{-7}# (0.029)	72.5	78.5	0.126
	KABC (4 crashes)	-6.946# (9.189)	-	-	0.236# (0.985)	8.69×10^{-7}# (0.032)	37.6	43.6	0.051
	KAB (2 crashes)	3.758# (10.847)	-	-	-1.019# (1.228)	1.43×10^{-4}# (10.729)	20.5	26.5	0.026
	KA (1 crash)	42.737# (68.643)	-	-	-5.782# (8.396)	4.734# (19.475)	7.9	13.9	0.016
Single vehicle	KABCO	-7.259 (1.796)	-	-	0.663 (0.192)	0.826 (0.251)	405.2	411.2	1.026
	KABC	-7.837 (2.281)	-	-	0.608 (0.242)	0.256# (0.394)	219.3	225.3	0.469
	KAB (18 crashes)	-2.295# (4.036)	-	-	-0.097# (0.441)	2.927×10^{-7}# (0.007)	114.4	120.4	0.218
	KA (1 crash)	1.116# (15.288)	-	-	-0.798# (1.714)	0.002# (77.506)	11.8	17.8	0.013

* Moore-Penrose inverse matrix used.

Not significant at 90th percentile confidence interval.

Table 4-24: Base Condition SPFs, Multilane Four-Leg Stop-Controlled (4ST) Intersections

Crash Type Minnesota (N = 139)	Severity	b_0	b_1	b_2	b_3	k	–2LL	AIC	MAD
Total	KABCO	-9.561 (1.353)	0.773 (0.140)	0.383 (0.073)	-	0.410 (0.099)	553.9	561.9	1.886
	KABC	-10.411 (1.814)	0.711 (0.185)	0.475 (0.102)	-	0.433 (0.162)	381.2	389.2	1.064
	KAB	-8.843 (2.404)	0.441 (0.249)	0.509 (0.142)	-	0.683 (0.323)	269.4	277.4	0.668
	KA (14 crashes)	-13.245 (6.064)	-	-	1.053[#] (0.645)	5.857[#] (4.216)	85.5	91.5	0.184
Same direction	KABCO	-14.343 (2.160)	1.158 (0.218)	0.345 (0.115)	-	0.362 (0.210)	284.8	292.8	0.696
	KABC	-13.190 (3.713)	-	-	1.118 (0.391)	0.619[#] (0.845)	139.1	145.1	0.307
	KAB (12 crashes)	-9.502 (5.235)	-	-	0.641[#] (0.558)	0.877[#] (1.989)	82.5	88.5	0.158
	KA (1 crash)	9.504[#] (23.757)	-	-	-1.733[#] (2.708)	2.100 $\times10^{-5}$[#] (3.983)	11.2	17.2	0.014
Intersecting direction	KABCO	-11.531 (2.333)	0.496 (0.234)	0.939 (0.148)	-	0.942 (0.271)	342.2	350.2	1.048
	KABC	-8.626 (2.967)	-	-	0.757 (0.319)	1.867 (0.578)	286.3	292.3	0.805
	KAB	-9.196 (4.102)	-	-	0.740 (0.441)	3.498 (1.428)	181.8	187.8	0.483
	KA	-10.886 (7.863)	-	-	0.770 (0.843)	11.215 (8.632)	66.2	72.2	0.137
Opposite direction	KABCO (15 crashes)	-6.868[#] (4.439)	-	-	0.383[#] (0.476)	0.191[#] (1.253)	97.5	103.5	0.193
	KABC (9 crashes)	-6.939[#] (6.289)	-	-	0.338[#] (0.676)	8.545 $\times10^{-7}$[#] (0.023)	65.8	71.8	0.121
	KAB (5 crashes)	-6.969[#] (8.465)	-	-	0.276[#] (0.911)	3.638 $\times10^{-6}$[#] (0.125)	43.0	49	0.069
	KA (1 crash)	-77.373[#] (114.01)	-	-	7.178[#] (11.168)	0.776[#] (23.934)	7.1	13.1	0.013
Single vehicle	KABCO	-9.855 (1.848)	-	-	0.929 (0.196)	0.337 (0.188)	343.6	349.6	0.849
	KABC	-10.416 (2.743)	-	-	0.876 (0.290)	0.154[#] (0.449)	191.8	197.8	0.442
	KAB (22 crashes)	-7.161 (3.664)	-	-	0.456[#] (0.392)	0.092[#] (0.827)	126.5	132.5	0.268
	KA (2 crashes)	-27.071 (15.661)	-	-	2.284[#] (1.591)	2.778 $\times10^{-6}$[#] (0.108)	15.9	21.9	0.027

[*] Moore-Penrose inverse matrix used.

[#] Not significant at 90th percentile confidence interval.

Table 4-25: Base Condition SPFs, Multilane Four-Leg Signal-Controlled (4SG) Intersections

Crash Type Ohio (N = 53)	Severity	b_0	b_1	b_2	b_3	k	–2LL	AIC	MAD
Total	KABCO	-7.741 (2.037)	-	-	0.932 (0.232)	0.443 (0.151)	264.8	270.8	2.792
	KABC	-14.318 (4.593)	-	-	1.442 (0.515)	0.775# (0.567)	105.7	111.7	0.728
	KAB	-14.662 (5.772)	-	-	1.399 (0.644)	0.499# (0.829)	69.5	75.5	0.434
	KA (4 crashes)	-0.930* (14.969)	-	-	-0.318* (1.734)	4.932×10^{-6}* (0.676)	28.8*	34.8*	0.140*
Same direction	KABCO	-12.709 (2.805)	-	-	1.391 (0.316)	0.443 (0.213)	183.3	189.3	1.567
	KABC	-17.140 (6.798)	-	-	1.659 (0.756)	0.786# (1.067)	62.7	68.7	0.390
	KAB	-14.669# (11.669)	-	-	1.242# (1.305)	5.026# (8.002)	26.7	32.7	0.137
	KA (1 crash)	-0.050* (.)	-	-	-0.592* (0.114)	3.879×10^{-7}* (.)	9.2*	15.2*	0.036*
Intersecting direction	KABCO	-9.724 (3.021)	-	-	1.024 (0.342)	0.561 (0.268)	166.5	172.5	1.123
	KABC	-14.965 (7.444)	-	-	1.412 (0.837)	1.925# (1.787)	61.2	67.2	0.396
	KAB	-20.048 (9.846)	-	-	1.921 (1.095)	2.105# (2.486)	42.6	48.6	0.256
	KA (1 crash)	-0.088* (0.006)	-	-	-0.575* (0.107)	2.146×10^{-7}* (.)	10.3*	16.3*	0.038*
Opposite direction	KABCO	-9.904 (4.564)	-	-	0.965 (0.518)	0.999# (2.342)	119.3	125.3	0.673
	KABC (4 crashes)	0.813* (.)	-	-	-0.519* (0.060)	1.042×10^{-7}* (.)	29.4*	35.4*	0.143*
	KAB (3 crashes)	-1.476* (.)	-	-	-0.288* (.)	1.245×10^{-6}* (.)	23.4*	29.4*	0.108*
	KA (1 crash)	0.778* (12.892)	-	-	-0.682* (1.522)	5.146×10^{-6}* (.)	9.9*	15.9*	0.037*
Single vehicle	KABCO (14 crashes)	1.557# (3.382)	-	-	-0.378# (0.393)	0.573# (0.550)	105.3	111.3	0.685
	KABC (2 crashes)	-4.075* (0.350)	-	-	-0.0269* (0.0611)	1.623×10^{-7}* (.)	17.2*	23.2*	0.075*
	KAB (1 crash)	-5.352* (12.126)	-	-	0.039* (1.393)	4.470×10^{-7}* (.)	9.9*	15.9*	0.038*
	KA (1 crash)	-5.352* (12.126)	-	-	0.039* (1.393)	4.470×10^{-7}* (.)	9.9*	15.9*	0.038*

* Moore-Penrose inverse matrix used.
# Not significant at 90th percentile confidence interval.

Table 4-26: Validation of the Minnesota (2009–11) Safety Performance Functions Using the Ohio (2009–11) Data, Three-Leg Stop-Controlled (3ST) Intersections

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM, 2010)				Calibration Function (Srinivasan et al., 2016)					
					CF[+]	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	157	256.174	1.636	4.814	0.613	157	1.286	3.938	0.643 (0.142)	0.939 (0.236)	1.062 (0.294)	155.74	1.284	3.952
KABC	63	111.594	0.898	1.261	0.565	63	0.709	1.048	0.557 (0.099)	1.240 (0.378)	1.544 (0.620)	63.817	0.709	1.057
KAB	46	51.146	0.587	0.699	0.899	46	0.566	0.697	1.163 (0.497)	1.344 (0.482)	1.808 (0.832)	46.980	0.566	0.708
KA	15	9.408	0.197	0.207	1.594	15	0.237	0.211	0.002 (0.004)	-1.573 (0.836)	2.891[#] (2.499)	14.667	0.219	0.184
SD KABCO	61	66.245	0.697	1.237	0.921	61	0.675	1.224	0.752 (0.178)	0.534 (0.224)	1.797 (0.686)	60.399	0.677	1.219
SD KABC	21	28.462	0.327	0.201	0.738	21	0.288	0.192	0.691 (0.370)	0.943 (0.380)	0.356[#] (0.859)	21.094	0.289	0.192
SD KAB	15	9.097	0.187	0.147	1.649	15	0.226	0.144	2.542[#] (6.035)	1.173[#] (0.941)	1.043[#] (1.605)	14.996	0.225	0.144
SD KA	4	1.403	0.045	0.034	2.851	4	0.066	0.033	0.006[#] (0.046)	-0.404[#] (1.821)	9.400×10^{-5}[#] (2.894)	4.000	0.066	0.033
ID KABCO	47	77.587	0.741	1.105	0.606	47	0.592	0.916	0.558 (0.152)	0.637 (0.306)	2.890 (1.154)	47.174	0.597	0.879
ID KABC	23	43.927	0.464	0.502	0.524	23	0.331	0.383	0.311 (0.161)	0.378[#] (0.344)	5.613 (2.984)	22.937	0.34	0.358
ID KAB	17	26.069	0.305	0.253	0.652	17	0.249	0.231	0.367[#] (0.267)	0.566[#] (0.414)	4.200[#] (2.899)	17.002	0.254	0.223
ID KA	8	6.103	0.116	0.068	1.311	8	0.131	0.07	4.590×10^{-4}[#] (0.002)	-1.490[#] (1.474)	9.104×10^{-7}[#] (0.024)	8.000	0.119	0.059
OD KABCO	12	7.406	0.155	0.160	1.620	12	0.187	0.157	27682* (2.142×10^{-6})	4.651* (0.133)	2.976* (2.787)	11.947*	0.177*	0.151*
OD KABC	5	2.846	0.065	0.041	1.757	5	0.081	0.041	1.864* (.)	1.010* (.)	6.469×10^{-8}* (.)	5.119*	0.082*	0.040*
OD KAB	3	2.021	0.042	0.026	1.485	3	0.051	0.026	2.200×10^{-5}* (1.400×10^{-5})	-1.552* (.)	9.786×10^{-8}* (8.320×10^{-4})	3.000*	0.049*	0.025*
OD KA	0	100.834	0.862	60.951	0	0	0	0	Failed to Converge					
SV KABCO	37	102.663	0.747	0.707	0.360	37	0.459	0.351	0.361 (0.064)	1.034 (0.460)	0.361[#] (0.536)	36.955	0.458	0.351
SV KABC	14	34.665	0.347	0.168	0.404	14	0.211	0.136	0.706[#] (0.703)	1.498 (0.856)	1.075[#] (1.663)	13.986	0.21	0.135
SV KAB	11	14.895	0.202	0.121	0.739	11	0.174	0.120	8.350×10^{-13}* (0)	-12.217* (4.010×10^{-4})	2.236* (2.724)	10.973*	0.168*	0.115*
SV KA	3	0.950	0.034	0.043	3.158	3	0.051	0.043	5.402×10^{-6}* (.)	-1.645* (.)	1.135* (2.192)	2.903*	0.048*	0.040*

[+]Calibration Factor; * Moore-Penrose inverse matrix used; [#] Not significant at 90th percentile confidence interval.

[+]Calibration Factor
* Moore-Penrose inverse matrix used.
[#] Not significant at 90th percentile confidence interval.

Table 4-27: Calibration of the Minnesota (2009–11) Safety Performance Functions Using the Ohio (2009–11) Data, Four-Leg Stop-Controlled (4ST) Intersections

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM)				Calibration Function (Srinivasan et al. 2016)					
					CF+	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	199	285.409	2.572	13.306	0.697	199	2.114	9.81	0.922 (0.234)	0.817 (0.196)	0.694 (0.184)	202.66	2.098	8.833
KABC	83	132.870	1.381	4.001	0.625	83	1.122	2.688	0.764 (0.143)	0.717 (0.240)	1.124 (0.419)	83.856	1.096	2.348
KAB	69	66.733	0.974	1.941	1.034	69	0.985	1.959	1.030 (0.204)	0.720 (0.274)	1.329 (0.532)	69.392	0.968	1.778
KA	23	8.143	0.334	0.637	2.824	23	0.438	0.612	2.638# (3.970)	0.957# (0.632)	3.014 (1.663)	23.588	0.443	0.611
SD KABCO	42	67.011	0.747	1.138	0.627	42	0.620	0.815	0.634 (0.122)	0.699 (0.257)	0.673# (0.476)	42.536	0.623	0.725
SD KABC	17	15.813	0.307	0.185	1.075	17	0.314	0.186	0.219# (0.230)	0.036# (0.540)	2.706×10^{-8}# (4.650×10^{-4})	17.004	0.33	0.186
SD KAB	12	7.120	0.201	0.123	1.685	12	0.239	0.118	16.042# (50.849)	1.957# (1.356)	7.206×10^{-6}# (0.404)	12.000	0.234	0.120
SD KA	2	0.603	0.031	0.024	3.316	2	0.048	0.025	0.002# (0.010)	-0.459# (0.927)	5.835×10^{-6}# (0.375)	2.000	0.047	0.024
ID KABCO	96	201.074	2.408	34.202	0.477	96	1.61	10.946	1.056 (0.202)	0.295 (0.166)	1.979 (0.567)	96.532	1.373	4.559
ID KABC	47	50.749	0.78	1.452	0.926	47	0.764	1.451	1.279 (0.497)	1.786 (0.689)	2.205 (0.892)	48.808	0.760	1.525
ID KAB	40	24.419	0.608	1.114	1.638	40	0.688	1.068	4.304# (4.039)	1.842 (0.765)	2.439 (1.072)	41.749	0.681	1.129
ID KA	19	5.952	0.275	0.487	3.192	19	0.377	0.460	5.757# (13.421)	1.226# (0.884)	3.166# (1.954)	19.301	0.377	0.463
OD KABCO	18	8.929	0.279	0.226	2.016	18	0.342	0.211	58.693# (159.81)	2.546 (1.259)	0.004# (0.776)	18.000	0.330	0.210
OD KABC	8	5.477	0.149	0.110	1.461	8	0.174	0.109	1000# (2396.4)	3.459 (0.916)	0.718# (1.882)	7.771	0.166	0.106
OD KAB	8	2.998	0.126	0.114	2.668	8	0.174	0.109	1000# (2467.560)	2.811 (0.760)	0.871# (2.081)	7.722	0.168	0.106
OD KA	1	1.076	0.025	0.024	0.930	1	0.024	0.023	0.034# (0.098)	0.108# (0.306)	8.483×10^{-6}# (0.638)	1.000	0.024	0.012
SV KABCO	43	74.346	0.734	0.95	0.578	43	0.608	0.749	0.582 (0.100)	0.957 (0.386)	0.422# (0.348)	43.248	0.609	0.747
SV KABC	11	25.943	0.352	0.157	0.424	11	0.225	0.114	0.294# (0.316)	0.663# (0.903)	1.464×10^{-6}# (0.038)	11.016	0.227	0.114
SV KAB	9	13.087	0.229	0.098	0.688	9	0.191	0.095	1.097# (2.053)	1.277# (1.047)	1.553×10^{-6}# (0.052)	8.690	0.188	0.095
SV KA	1	1.099	0.025	0.012	0.910	1	0.024	0.012	0.050# (0.206)	0.296# (0.881)	9.103×10^{-6}# (0.552)	1.000	0.024	0.012

+Calibration Factor; * Moore-Penrose inverse matrix used; # Not significant at 90th percentile confidence interval.

Table 4-28: Calibration of the Ohio (2009–11) Safety Performance Functions Using the Minnesota (2009–11) Data, Four-Leg Signal-Controlled (4SG) Intersections

Crash Type	Observed Crashes	HSM Pred.	MAD	MSPE	Calibration Factor (HSM)				Calibration Function (Srinivasan et al. 2016)					
					CF^{+}	N Fitted	MAD	MSPE	a	b	k	N Fitted	MAD	MSPE
KABCO	202	238.99	5.664	45.653	0.845	202	5.239	44.93	0.227$^{\#}$ (0.249)	1.551 (0.473)	0.358 (0.138)	200.025	4.929	39.730
KABC	57	45.671	1.491	5.074	1.248	57	1.591	4.718	1.289 (0.395)	0.944 (0.376)	0.280$^{\#}$ (0.208)	56.522	1.581	4.771
KAB	16	21.439	0.667	0.590	0.746	16	0.596	0.514	0.697* (0.025)	0.724* (0.593)	$8.5\times10^{-8*}$ (.)	15.135*	0.602*	0.513*
KA	2	1.466	0.140	0.162	1.364	2	0.161	0.162	Failed to Converge					
SD KABCO	96	139.79	3.133	15.321	0.687	96	2.791	12.42	0.477$^{\#}$ (0.363)	1.189 (0.422)	0.547 (0.263)	95.170	2.746	11.927
SD KABC	24	22.243	1.063	1.672	1.079	24	1.082	1.660	1.071 (0.314)	0.875$^{\#}$ (0.540)	0.920$^{\#}$ (0.751)	23.710	1.081	1.683
SD KAB	5	4.637	0.316	0.166	1.078	5	0.324	0.166	0.400* (.)	0.387* (.)	0* (.)	5.000*	0.328*	0.164*
SD KA	1	0.045	0.043	0.042	22.40	1	0.082	0.043	0* (.)	-1.287* (.)	0.794* (.)	1.009*	0.072*	0.033*
ID KABCO	65	79.477	1.871	6.012	0.818	65	1.727	5.685	0.767* (0.504)	1.048* (0.525)	0.334* (0.210)	64.778*	1.724*	5.662*
ID KABC	25	17.941	0.795	1.233	1.393	25	0.807	1.118	1.372 (0.318)	0.933 (0.481)	0.059$^{\#}$ (0.273)	24.966	0.807	1.126
ID KAB	10	15.474	0.498	0.380	0.646	10	0.405	0.293	0.562* (.)	0.587* (.)	0* (.)	10.000*	0.452*	0.296*
ID KA	1	1.093	0.080	0.037	0.915	1	0.077	0.037	2.661* (0.617)	1.411* (0.323)	$9.6\times10^{-7*}$ (.)	0.846*	0.071*	0.037*
OD KABCO	16	34.491	1.105	1.512	0.464	16	0.699	0.954	0.279 (0.164)	2.158 (1.156)	0.471$^{\#}$ (0.630)	16.044	0.671	0.918
OD KABC	5	2.454	0.273	0.257	2.037	5	0.340	0.243	13021.250* (.)	4.897* (.)	0.562* (.)	4.978*	0.309*	0.230*
OD KAB	0	1.541	0.064	0.004	0	0	0	0	0* (.)	1.588* (0.073)	0.979* (0)	$1.869\times10^{-8*}$	$7.78\times10^{-10*}$	$6.097\times10^{-19*}$
OD KA	0	0.293	0.012	0.000	0	0	0	0	0* (.)	2.022* (0.048)	0.996* (0)	$1.242\times10^{-11*}$	$5.177\times10^{-13*}$	$2.934\times10^{-25*}$
SV KABCO	25	9.301	1.043	3.109	2.688	25	1.109	2.788	0.014$^{\#}$ (0.032)	-4.310 (2.210)	0.747$^{\#}$ (0.622)	24.235	1.032	2.024
SV KABC	3	0.709	0.148	0.119	4.229	3	0.220	0.111	$2.69\times10^{-7*}$ (.)	-3.673* (.)	0* (.)	3.000*	0.212*	0.106*
SV KAB	1	3.165	0.164	0.062	0.316	1	0.080	0.042	0.084* (.)	0.313* (.)	0* (.)	1.000*	0.080*	0.040*
SV KA	0	3.165	0.132	0.030	0	0	0	0	0* (.)	1.553* (0.226)	1.000* (0)	$1.242\times10^{-11*}$	$5.177\times10^{-13*}$	$2.934\times10^{-25*}$

$^{+}$Calibration Factor; * Moore-Penrose inverse matrix used; $^{\#}$ Not significant at 90th percentile confidence interval.

5 MODELS FOR URBAN AND SUBURBAN ARTERIALS

5.1 ROADWAY SEGMENTS

5.1.1 Estimation Data

The process of developing models for urban and suburban arterial road segments involved developing an initial set of models and then validating them with a second dataset obtained later. Following the successful validation, we combined the two datasets into a larger dataset to re-estimate the models.

We attempted three sets of models:

1. Base Condition SPFs—using only those sites meeting the HSM base conditions
2. Average Condition Exposure SPFs—using all available sites but only including exposure-related variables in the models
3. Average Condition Multi-Variable SPFs—using all available sites and including non-exposure variables where possible

The data for the initial models developed for urban and suburban arterial segments came from Ohio. The Ohio Department of Transportation (DOT) collected all of the data necessary for calibrating and applying the SPFs and CMFs in the current HSM chapter on urban and suburban arterials and made them available to the research team. No variables beyond the necessary ones were available in this database.

The Ohio data were provided for all site types in the HSM chapter, including the following:

- Two-lane undivided (2U)
- Two-lane plus two-way left-turn lane (3T)
- Four-lane divided (4D)
- Four-lane undivided (4U)
- Four-lane plus two-way left-turn lane (5T)

Table 5-1 shows the total mileage for each site type in Ohio and the sum of crashes by crash type in the data used for base condition models. Traffic volume and crash data for all sites include the years 2007–11. The crash data do not include intersection-related crashes. The Ohio DOT defines intersection crashes as any crash within 250 feet of an intersection. This definition has been adopted in part because of the perceived unreliability of the police reports in properly identifying intersection-related crashes. The queried crash types for the initial models included the following:

- Pedestrian–vehicle (PED)
- Bicycle–vehicle (BIKE)
- Multiple-vehicle driveway related (MVD)
- Rear end (RE)
- Head-on (HO)
- Right angle (ANG)
- Sideswipe same direction (SSD)
- Sideswipe opposite direction (SOD)
- Multiple-vehicle non-driveway other (MVN OTHER)
- Single vehicle (SV)

- Single-vehicle run-off-road (SV ROR)
- Single-vehicle fixed object (SV FIXEDOBJ)
- Single-vehicle other object (SV OTHEROBJ)
- Single-vehicle other (SV OTHER)
- Animal-related (ANIMAL)
- Nighttime (NIGHT)

In the course of developing the SPFs, some crash types were combined. Additionally, we decided not to include animal crashes, although they remain for the purposes of the data description, as they are informative with respect to crash type occurrence in Ohio, which provided the initial calibration data.

Table 5-2 shows the number of sites (N), minimum, maximum, mean, and standard deviation for the crash counts for the five-year period for each site type in Ohio for the base condition sites.

Table 5-3 shows the number of sites, minimum, maximum, mean, and standard deviation for the continuous explanatory variables for each site type in Ohio for the base condition sites. The explanatory variable definitions are identical to those in the current HSM urban and suburban arterial chapter.

Table 5-4 shows the total mileage for the discrete explanatory variables for each site type for base condition sites in Ohio. Again, the variable definitions are identical to those in the current HSM chapter.

Table 5-1: Ohio Segment Length and Crash Type Totals for Five-Year Period for Base Condition Sites (Urban/Suburban Arterial Segments)

Site Type	Length (mi.)	PED	BIKE	MVD	RE	HO	ANG	SSD	SOD	MVN OTHER	SV	ROR	ANIMAL	FO	MO	SV OTHER	NIGHT
2U	447.256	25	10	370	1255	43	77	140	240	311	3582	1360	2112	1289	36	145	2516
3T	62.174	3	2	108	259	4	37	21	24	43	364	94	253	95	5	11	298
4D	160.595	17	5	134	1151	4	58	320	38	174	1611	488	1045	443	44	79	1368
4U	97.7	19	8	232	657	17	109	247	63	112	462	158	278	159	12	13	528
5T	74.99	12	8	396	1205	12	146	322	59	165	674	204	420	206	18	30	793

Table 5-2: Ohio Segment Crash Type Statistics for Five-Year Period for Base Condition Sites (Urban/Suburban Arterial Segments)

Site Type	Stat.	PED	BIKE	MVD	RE	HO	ANG	SSD	SOD	MVN OTHER	SV	SV ROR	ANIM AL	FO	MO	SV OTHER	Night
2U	N	760	760	760	760	760	760	760	760	760	760	760	760	760	760	760	760
2U	MIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2U	MAX	2	1	13	66	2	7	7	8	8	109	36	81	32	3	8	78
2U	MEAN	0.03	0.01	0.49	1.65	0.06	0.10	0.18	0.32	0.41	4.71	1.79	2.78	1.70	0.05	0.19	3.31
2U	STD	0.19	0.11	1.30	4.21	0.25	0.47	0.62	0.88	1.01	8.75	3.84	5.79	3.64	0.25	0.60	6.28
3T	N	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182	182
3T	MIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3T	MAX	1	1	10	34	1	6	2	4	7	56	13	42	12	1	2	36
3T	MEAN	0.02	0.01	0.59	1.42	0.02	0.20	0.12	0.13	0.24	2	0.52	1.39	0.52	0.03	0.06	1.64
3T	STD	0.13	0.10	1.54	3.70	0.15	0.63	0.37	0.52	0.70	5.23	1.42	3.92	1.41	0.16	0.26	3.73
4D	N	358	358	358	358	358	358	358	358	358	358	358	358	358	358	358	358
4D	MIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4D	MAX	2	2	19	172	1	7	28	4	19	98	37	61	36	4	8	85
4D	MEAN	0.05	0.01	0.37	3.22	0.01	0.16	0.89	0.11	0.49	4.50	1.36	2.92	1.24	0.12	0.22	3.82
4D	STD	0.24	0.14	1.64	13.98	0.11	0.67	2.64	0.44	1.62	10.76	3.51	7.32	3.19	0.45	0.86	9.13
4U	N	348	348	348	348	348	348	348	348	348	348	348	348	348	348	348	348
4U	MIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4U	MAX	2	3	34	78	2	21	32	7	13	35	15	25	15	1	2	28
4U	MEAN	0.05	0.02	0.67	1.89	0.05	0.31	0.71	0.18	0.32	1.33	0.45	0.80	0.46	0.03	0.04	1.52
4U	STD	0.26	0.20	2.55	5.91	0.23	1.34	2.38	0.68	1.14	3.26	1.31	2.25	1.32	0.18	0.22	3.61
5T	N	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180	180
5T	MIN	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5T	MAX	2	2	29	115	2	11	24	4	18	46	16	29	16	2	4	43
5T	MEAN	0.07	0.04	2.20	6.69	0.07	0.81	1.79	0.33	0.92	3.74	1.13	2.33	1.14	0.10	0.17	4.41
5T	STD	0.27	0.23	4.37	15.99	0.27	1.75	3.48	0.76	2.16	6.75	2.34	4.66	2.39	0.35	0.50	7.54

Table 5-3: OH Segment Continuous Variable Statistics for Base Condition Sites (Urban/Suburban Arterial Segments)

Site Type	Stat.	Length	AADT	Med Width	Parking Prop	FODensity	Offset FO	Maj Comm	Min Comm	Maj Ind	Min Ind	Maj Res	Min Res	Other Dwy
2U	N	760	760	760	760	760	760	760	760	760	760	760	760	760
2U	MIN	0.01	100	0	0	25	2	0	0	0	0	0	0	0
2U	MAX	6.29	23,028	0	0	75	20	10	41	6	28	4	193	5
2U	MEAN	0.59	6975	0	0	37.76	8.95	0.19	2.34	0.08	0.97	0.03	10.96	0.03
2U	STD	0.72	3978	0	0	12.95	3.80	0.77	4.70	0.48	2.63	0.26	20.03	0.24
3T	N	182	182	182	182	182	182	182	182	182	182	182	182	182
3T	MIN	0.02	1356	0	0	25	2	0	0	0	0	0	0	0
3T	MAX	3.29	23780	0	0	75	20	11	49	12	10	2	65	1
3T	MEAN	0.34	1022	0	0	41.87	8.11	0.98	5.24	0.29	0.41	0.05	4.82	0.02
3T	STD	0.44	4034	0	0	13.75	4.09	1.89	9.12	1.17	1.27	0.27	8.99	0.15
4D	N	358	358	358	358	358	358	358	358	358	358	358	358	358
4D	MIN	0.01	256	10	0	25	10	0	0	0	0	0	0	0
4D	MAX	4.81	45,874	100	0	75	30	33	47	8	5	2	64	2
4D	MEAN	0.45	14,384	33.27	0	34.32	21.63	0.42	1.11	0.14	0.13	0.01	0.87	0.02
4D	STD	0.67	8758	29.14	0	11.67	4.11	2.11	4.60	0.69	0.61	0.14	4.24	0.18
4U	N	348	348	348	348	348	348	348	348	348	348	348	348	348
4U	MIN	0.01	1150	0	0	25	2	0	0	0	0	0	0	0
4U	MAX	5.96	41,418	0	0	75	25	11	57	11	16	3	78	4
4U	MEAN	0.28	14,281	0	0	43.09	7.84	0.48	3.56	0.27	0.52	0.04	3.45	0.05
4U	STD	0.47	7350	0	0	13.84	4.54	1.32	7.05	0.93	1.93	0.25	9.06	0.34
5T	N	180	180	180	180	180	180	180	180	180	180	180	180	180
5T	MIN	0.01	5356	0	0	25	2	0	0	0	0	0	0	0
5T	MAX	2.91	50,553	0	0	75	20	23	75	9	16	2	46	2
5T	MEAN	0.42	19,422	0	0	38.97	8.47	2.23	8.09	0.42	0.52	0.07	3.05	0.08
5T	STD	0.51	83456	0	0	10.74	4.37	4.06	12.22	1.20	1.92	0.30	7.02	0.31

Table 5-4: OH Segment Categorical Variable Total Mileage (mi.) for Base Condition Sites (Urban/Suburban Arterial Segments)

Variable	2U	3T	4U	4D	5T
Lighting	Yes – 0.000 No – 447.256	Yes – 0.000 No – 62.174	Yes – 0.000 No – 97.700	Yes – 0.000 No – 160.595	Yes – 0.000 No – 74.990
Automated Enforcement	Yes – 0.000 No – 447.256	Yes – 0.000 No – 62.174	Yes – 0.000 No – 97.700	Yes – 0.000 No – 160.595	Yes – 0.000 No – 74.990
Speed Limit (mph)	<=30 – 5.693 >30 – 441.563	<=30 – 2.939 >30 – 59.235	<=30 – 7.844 >30 – 89.856	<=30 – 1.151 >30 – 159.444	<=30 – 0.498 >30 – 74.492
Parking	Yes – 0.000 No – 447.256	Yes – 0.000 No – 62.174	Yes – 0.000 No – 97.700	Yes – 0.000 No – 160.595	Yes – 0.000 No – 74.990
Parking Type	None – 447.256	None – 62.174	None – 97.700	None – 160.595	None – 74.990

5.1.2 Estimated Models

Documented in this section are the base condition models intended for use in the HSM predictive chapter for urban and suburban arterials. Models developed for all sites representing the average site conditions are found in Appendix A. These could be applied for network screening or other safety management tasks where models for average site conditions are desired.

The model development process involved using the Ohio data to estimate a set of initial models, which we subsequently validated where possible using a dataset that later became available from Minnesota. Following the validation, we combined the two datasets to re-estimate the final models.

We developed the initial models using the same base conditions as those in the current HSM chapter for urban and suburban arterials:

- No on-street parking
- No roadside fixed objects
- A 15-foot median width for divided roads
- No lighting
- No automated speed enforcement

The model predictions do not include intersection-related or animal crashes. The initial base condition models were estimated for the following crash types:

- Total (TOT)
- Multiple-vehicle driveway related (MVD)
- Multiple-vehicle non-driveway related (MVN)
- Rear end (RE)
- Sideswipe same direction (SSD)
- Head-on plus sideswipe opposite direction (HO+SOD)
- Multiple-vehicle non-driveway other (MVN OTHER)
- Single-vehicle (SV)
- Nighttime (NIGHT)

To develop the base condition SPFs, we used only sites with no lighting or parking or automated enforcement. Because no sites had zero roadside fixed objects and few divided roadways had a median width of exactly 15 feet, these variables were included in the models only if considered appropriate for that crash type and if the variable was statistically significant in the model and with the expected direction of effect. For TOT, MVD, SV, and NIGHT crashes, the number of driveways was also directly included in the models where warranted. If the variables for fixed objects or median width were included, they would have been set to the base condition for application. The number of driveways in a segment should be entered in those models where it is included—that is, there is no base condition for the number of driveways in a segment.

Note that some parameter estimates are not statistically significant at the 95 percent confidence level but are consistent across site types and/or crash types in the direction of effect and magnitude. In these cases, we deemed the estimates acceptable.

For TOT, MVD, and NIGHT crashes, parameter estimates for driveway count variables were inconsistent in their levels of statistical significance and whether one driveway type was associated with fewer or more crashes. In light of these findings, we considered two options. Option 1 used the same driveway definitions and model form for considering driveways as in the current HSM chapter. Option 2 used the total driveway density (driveways per mile) as an alternate variable. For TOT and NIGHT crashes, the model form for Option 1 did not include length, as the inclusion of this variable created poor parameter estimates for the relationship between average annual daily traffic (AADT) and driveways.

Table 5-5 to Table 5-16 document all initial base models developed using the Ohio data. The model form is provided below each table for each crash type, along with the parameter estimates and standard error (in brackets). For most site type/crash type combinations, a model was successfully calibrated; the tables note where they were not.

5.1.3 Validation of Models

5.1.3.1 Initial model validation

We validated the initial models that used Ohio data with data from Minnesota covering the years 2010–14. The sample size of crashes was not large enough, however, to validate all crash type models. Table 5-17 provides the segment length and crash type totals for the Minnesota base condition sites. For validation, we calculated the calibration factors for each SPF using the current HSM procedure, which is to calculate a simple ratio of the sum of observed crashes divided by the sum of predicted crashes prior to calibration. We used “The Calibrator,” a spreadsheet tool developed by the Federal Highway Administration (FHWA), for the validation exercise.

Table 5-5: Total (TOT) for Base Conditions Option 1 (Urban/Suburban Arterial Segments)

Site Type	a1	b1	a2	b2	e	f	g	h	j	k	l	dispersion
2U	-10.4000 (1.6710)	1.0210 (0.1885)	-10.4800 (1.6700)	0.6210 (0.1872)	21.3300 (13.8400)	5.6580 (3.6670)	47.6600 (18.6600)	26.5800 (12.8800)	15.3400 (18.0800)	7.0580 (3.6000)	41.5400 (20.4400)	0.6276 (0.0471)
3T	-11.8600 (4.5250)	1.1340 (0.4928)	-11.8600 (3.7500)	0.6005 (0.4125)	33.5800 (21.3500)	27.1700 (16.8600)	28.6500 (22.4100)	11.9700 (19.5800)	25.6500 (28.1000)	8.8680 (9.306)	-1.3600 (31.7500)	0.6635 (0.1085)
4U	-13.1800 (1.7260)	1.2010 (0.1808)	-19.9900 (3.0130)	1.5070 (0.3086)	39.4400 (19.2000)	26.8600 (13.6100)	24.5800 (18.6800)	28.9600 (16.1000)	18.3300 (25.4600)	13.7500 (8.4100)	-2.1840 (26.0900)	0.4030 (0.0481)
4D	-17.7500 (1.6420)	1.8030 (0.1742)	-20.7000 (5.1830)	1.6240 (0.5345)	24.0600 (19.6500)	18.8500 (15.5800)	23.9900 (21.4900)	26.0800 (23.1600)	15.5500 (29.7900)	24.5300 (16.5300)	18.9500 (28.0700)	0.4158 (0.0444)
5T	-23.2900 (4.2650)	2.2750 (0.4321)	-11.4400 (2.0540)	0.6492 (0.2119)	37.0800 (17.5300)	30.4900 (14.6400)	34.1100 (17.4900)	11.4100 (19.9800)	2.3670 (29.8500)	9.3020 (8.1580)	0.0370 (30.6800)	0.5754 (0.0749)

Crashes per year = $\exp(a1)(AADT)^{b1}$ +
$\exp(a2)(AADT)^{b2}$(e*MajComm+f*MinComm+g*MajInd+h*MinInd+j*MajRes+k*MinRes+l*OtherDwy)

Dispersion is modeled as a constant

Table 5-6: Total (TOT) for Base Conditions Option 2 (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-6.0860 (0.5892)	0.7511 (0.0670)	-0.5430 (0.0919)	-0.6938 (0.0820)	n/a	n/a	n/a
3T	-10.2221 (2.2704)	1.1790 (0.2454)	-0.3919 (0.2694)	-0.2997 (0.1898)	n/a	n/a	n/a
4U	-14.6786 (1.7359)	1.6114 (0.1817)	-0.0053 (0.1939)	-0.4560 (0.1231)	0.0089 (0.0030)	n/a	n/a
4D	-11.9469 (1.1524)	1.3272 (0.1179)	-0.6179 (0.1560)	-0.5502 (0.1208)	0.0182 (0.0050)	n/a	-0.0054 (0.0032)
5T	-11.6621 (1.7458)	1.3068 (0.1767)	-0.7018 (0.2128)	-0.7834 (0.1490)	n/a	0.0162 (0.0096)	n/a

Crashes per year = $(length)\exp^{(Alpha1)} AADT^{(Beta1)} \exp^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = $\exp^{(Alpha2)}(length)^{(Beta2)}$

Table 5-7: Multiple-Vehicle Non-Driveway (MVN) for Base Conditions (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4
2U	-12.0774 (0.7334)	1.3440 (0.0839)	-0.5669 (0.1097)	-0.7129 (0.0950)	0.0050 (0.0040)	n/a
3T	-14.9031 (2.8001)	1.6288 (0.3017)	-0.2586 (0.2970)	-0.4172 (0.2068)	n/a	n/a
4U	-18.2639 (2.0110)	1.9781 (0.2096)	-0.1077 (0.1961)	-0.6026 (0.1195)	n/a	n/a
4D	-16.2885 (1.4294)	1.6796 (0.1388)	-0.4327 (0.1634)	-0.6772 (0.1221)	0.0223 (0.0079)	-0.0053 (0.0039)
5T	-14.0029 (1.8790)	1.5117 (0.1899)	-0.6424 (0.2155)	-0.8478 (0.1544)	0.0208 (0.0104)	n/a

Crashes per year = $(length)exp^{(Alpha1)}AADT^{(Beta1)}\ exp^{(Beta3*fodensity+Beta4*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = $exp^{(Alpha2)}(length)^{(Beta2)}$

Table 5-8: Rear End (RE) for Base Conditions (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4
2U	-16.9150 (1.0174)	1.8217 (0.1136)	-0.2968 (0.1320)	-0.6680 (0.1181)	n/a	n/a
3T	-19.8717 (3.7237)	2.1175 (0.4000)	0.2335 (0.2991)	-0.2750 (0.2398)	n/a	n/a
4U	-20.3644 (2.3837)	2.1262 (0.2471)	0.1069 (0.2222)	-0.6057 (0.1436)	n/a	n/a
4D	-23.9555 (2.0377)	2.4546 (0.2086)	0.1034 (0.1749)	-0.5155 (0.1390)	n/a	n/a
5T	-18.0852 (2.3200)	1.9245 (0.2344)	-0.1383 (0.2344)	-0.6228 (0.1740)	n/a	n/a

Crashes per year = $(length)exp^{(Alpha1)}AADT^{(Beta1)}\ exp^{(Beta3*fodensity+Beta4*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = $exp^{(Alpha2)}(length)^{(Beta2)}$

Table 5-9: Sideswipe-Same-Direction (SSD) for Base Conditions (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4
2U	-14.9454 (2.0400)	1.3668 (0.2269)	0.3567 (0.3159)	-0.3906 (0.3335)	n/a	n/a
3T	-12.7872 (6.3104)	1.0883 (0.6755)	-1.0868 (1.1962)	-1.0000 (n/a)	n/a	n/a
4U	-21.7028 (3.1316)	2.1713 (0.3240)	0.0021 (0.3140)	-0.7319 (0.2117)	n/a	n/a
4D	-10.8456 (1.6374)	1.0164 (0.1676)	-0.4498 (0.2638)	-0.6528 (0.2570)	n/a	n/a
5T	-13.8498 (2.4777)	1.3813 (0.2499)	-0.4941 (0.3007)	-0.3073 (0.2718)	n/a	n/a

Crashes per year = (length)exp$^{(Alpha1)}$AADT$^{(Beta1)}$ exp$^{(Beta3*fodensity+Beta4*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = exp$^{(Alpha2)}$(length)$^{(Beta2)}$

Table 5-10: Head-On + Sideswipe-Opposite-Direction (HO+SOD) for Base Conditions (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4
2U	-8.2884 (1.3195)	0.7028 (0.1488)	0.0027 (0.2386)	-0.4152 (0.3140)	n/a	n/a
3T	-15.5567 (6.9161)	1.4489 (0.7412)	-0.5288 (1.5562)	-2.8313 (1.1828)	n/a	n/a
4U	No model calibrated					
4D	-10.7128 (3.3027)	0.7981 (0.3360)	0.0622 (0.7457)	-1.2091 (0.6002)	n/a	n/a
5T	-10.3196 (3.6247)	0.8767 (0.3658)	-0.4212 (0.5618)	-0.0014 (0.7216)	n/a	n/a

Crashes per year = (length)exp$^{(Alpha1)}$AADT$^{(Beta1)}$ exp$^{(Beta3*fodensity+Beta4*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = exp$^{(Alpha2)}$(length)$^{(Beta2)}$

Table 5-11: Multiple-Vehicle Non-Driveway Other (MVN OTHER) for Base Conditions (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4
2U	-9.9584 (1.2413)	0.8413 (0.1391)	-0.5762 (0.3268)	-1.4801 (0.2629)	0.0153 (0.0063)	n/a
3T	-8.9257 (3.4592)	0.8182 (0.3723)	-1.0071 (0.7971)	-0.1620 (0.7198)	n/a	n/a
4U	-16.1223 (3.1953)	1.3978 (0.3222)	0.4561 (0.3452)	-0.3153 (0.3047)	0.0277 (0.0117)	n/a
4D	-11.2681 (2.1597)	0.9584 (0.2055)	-0.5306 (0.3606)	-0.7167 (0.3160)	0.0199 (0.0111)	-0.0109 (0.0065)
5T	-13.4898 (2.9205)	1.1930 (0.2901)	-0.6420 (0.4135)	-0.6885 (0.4026)	0.0221 (0.0121)	n/a

Crashes per year = (length)exp$^{(Alpha1)}$AADT$^{(Beta1)}$ exp$^{(Beta3*fodensity+Beta4*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = exp$^{(Alpha2)}$(length)$^{(Beta2)}$

Table 5-12: Single-Vehicle (SV) for Base Conditions (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4
2U	No model calibrated					
3T	-4.0996 (3.7906)	0.2682 (0.4107)	0.2009 (0.5641)	0.5802 (0.7384)	n/a	n/a
4U	-7.5399 (2.8850)	0.6261 (0.3014)	-0.5967 (0.6644)	-1.1117 (0.5442)	n/a	n/a
4D	-7.6387 (1.4666)	0.6832 (0.1514)	-0.7085 (0.2950)	-0.5342 (0.3962)	n/a	n/a
5T	-2.8316 (3.1881)	0.1865 (0.3242)	-0.0347 (0.3416)	-0.7949 (0.3089)	n/a	n/a

Crashes per year = (length)exp$^{(Alpha1)}$AADT$^{(Beta1)}$ exp$^{(Beta3*fodensity+Beta4*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = exp$^{(Alpha2)}$(length)$^{(Beta2}$

Table 5-13: Nighttime (NIGHT) for Base Conditions (Option 1) (Urban/Suburban Arterial Segments)

Site Type	a1	b1	a2	b2	e	f	g	h	j	k	l	dispersion
2U	-8.184 (1.176)	0.5827 (0.131)	-9.2390 (1.456)	0.2255 (0.161)	19.360 (17.170)	3.3640 (4.3750)	50.5000 (21.560)	31.4600 (16.610)	15.3100 (22.270)	10.1500 (5.9840)	32.5600 (22.720)	0.6245 (0.0938)
3T	No model calibrated											
4U	No model											
4D	-18.98 (2.427)	1.720 (0.252)	-20.150 (6.542)	1.3740 (0.664)	30.350 (20.480)	9.9470 (13.090)	19.8100 (20.280)	13.1600 (22.960)	14.5500 (30.460)	27.2200 (17.700)	28.7600 (28.4000)	0.4411 (0.0923)
5T	No model calibrated											

Crashes per year = $\exp(a1)(AADT)^{b1} + \exp(a2)(AADT)^{b2}(e*MajComm+f*MinComm+g*MajInd+h*MinInd+j*MajRes+k*MinRes+l*OtherDwy)$

Dispersion is modeled as a constant

Table 5-14: Nighttime (NIGHT) for Base Conditions (Option 2) (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-4.5003 (0.9316)	0.3562 (0.1060)	-0.2500 (0.1657)	-0.7635 (0.2024)	n/a	n/a	n/a
3T	-9.8418 (3.7771)	0.8976 (0.4063)	-0.2694 (0.6161)	-0.3113 (0.4739)	n/a	n/a	n/a
4U	-15.4899 (2.7560)	1.5044 (0.2859)	-0.1265 (0.3565)	-0.7128 (0.2904)	n/a	n/a	n/a
4D	-10.5500 (1.7363)	1.0089 (0.1750)	-0.5876 (0.2844)	-0.5028 (0.2706)	n/a	n/a	-0.0119 (0.0051)
5T	-12.0353 (2.2854)	1.1821 (0.2302)	-0.8642 (0.3749)	-0.7848 (0.2764)	n/a	n/a	n/a

Crashes per year = $(length)\exp^{(Alpha1)} AADT^{(Beta1)} \exp^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = $\exp^{(Alpha2)}(length)^{(Beta2)}$

Table 5-15: Multi-Vehicle Driveway (MVD) for Base Conditions (Option 1) (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	e	f	g	h	j	k	l	dispersion
2U	-14.1100 (3.6350)	0.7840 (0.3716)	41.2600 (18.2600)	18.2400 (9.5630)	29.5900 (18.4800)	21.2900 (12.1000)	17.0300 (14.9600)	6.6180 (4.0100)	24.9200 (20.2400)	0.8656 (0.1913)
3T	-11.1600 (4.4080)	0.4170 (0.4767)	33.5600 (18.0600)	25.6400 (14.0800)	37.4600 (19.5500)	15.9200 (13.9300)	32.1400 (24.3300)	9.1550 (7.0810)	-2.9530 (30.8800)	0.5281 (0.1680)
4U	No model calibrated									
4D	-15.8200 (3.8660)	0.8799 (0.3761)	32.5700 (17.3800)	19.4700 (11.5200)	19.4100 (15.9700)	30.0500 (16.8000)	10.8000 (28.9100)	20.6400 (12.8300)	32.4000 (18.5600)	0.4986 (0.1910)
5T	-13.7000 (2.4940)	0.7159 (0.2525)	0.4131 (17.3400)	30.6700 (14.5400)	30.7600 (17.5500)	15.1000 (20.9200)	3.1770 (29.4200)	4.2310 (3.9390)	-0.2437 (30.9100)	0.7528 (0.1609)

Crashes/year =exp(Alpha1)$AADT^{Beta1}$(e*MajComm+f*MinComm+g*MajInd+h*MinInd+j*MajRes+k*MinRes+l*OtherDwy)

Dispersion is held constant.

Table 5-16: Multi-Vehicle Driveway (MVD) for Base Conditions (Option 2) (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-12.6511 (1.4075)	1.1705 (0.1567)	0.1267 (0.1841)	-0.7488 (0.1874)	0.0165 (0.0040)	n/a	n/a
3T	-8.9552 (4.2685)	0.8526 (0.4686)	0.5432 (0.3483)	-0.1579 (0.3117)	0.0022 (0.0076)	n/a	n/a
4U	-18.4926 (3.2011)	1.7592 (0.3312)	0.3468 (0.2954)	-0.6666 (0.2208)	0.0199 (0.0057)	n/a	n/a
4D	-11.0400 (3.3545)	0.9786 (0.3428)	0.4003 (0.3425)	-0.7067 (0.2770)	0.0567 (0.0113)	n/a	-0.0558 (0.0145)
5T	-11.7476 (2.5826)	1.0318 (0.2599)	-0.5609 (0.2950)	-1.1911 (0.2496)	0.0131 (0.0070)	0.0298 (0.0148)	n/a

Crashes per year = (length)$exp^{(Alpha1)}AADT^{(Beta1)}$ $exp^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = $exp^{(Alpha2)}(length)^{(Beta2)}$

Table 5-17: MN Segment Length and Crash Type Totals for 5 Year Period for Base Condition Sites (Urban/Suburban Arterial Segments)

Facility Type	No. of segments	No. of miles	Average AADT	TOT	Multi-Veh Driveway	Rear End	Head-On + SOD	SSD	Multi-Veh Other	Single-Vehicle	Night
2U	236	33.86	9,511	320	21	118	25	16	36	115	69
3T	63	7.23	10,841	76	4	39	6	7	17	7	18
4D	92	14.95	22,150	308	4	182	15	31	25	57	63
4U	113	11.72	10,386	160	15	53	9	24	41	22	29
5T	15	1.60	15,753	20	2	6	2	2	6	2	3

Table 5-18 to Table 5-21 provide goodness-of-fit measures for the initial base condition models validated with the Minnesota data. The cumulative residuals (CURE) plot measures are for CURE plots using the predicted number of crashes on the x-axis. Only site/crash types with a reasonable number of crashes (approximately 100) are included.

The interpretation of the goodness-of-fit measures is challenging, since the implications can vary significantly depending on the measure used. Some results are not promising but not surprising, and quite typical when assessing the transferability of SPFs to other jurisdictions. In general, the calibrated dispersion parameters are less than 2, indicating a reasonable prediction accuracy, and the values of MAD are not unreasonable. The modified R^2 and CURE plot statistics are more inconsistent across site and crash types. Note, though, that such biases within ranges of model predictions are not uncommon. On the whole, we conclude that the validation exercise does not provide evidence that the chosen model forms are not working.

For the models for total crashes where we investigated two options for handling driveway counts, the results indicate that Option 2 (summing all driveways together and dividing by segment length) was more successful overall. We applied this model form for developing the final models.

Table 5-18: Validation of Initial 2U Base Condition Models

Crash Type	Observed Crashes	Calibration Factor (coefficient of variation)	Dispersion	MAD	Modified R^2	CURE max dev	CURE % dev
TOT Option 1	320	1.02 (0.15)	1.27	1.33	0.28	65.52	71
TOT Option 2	320	0.92 (0.14)	1.08	1.31	0.04	58.83	16
MVN	195	0.64 (0.19)	1.41	0.96	0.00	95.17	97
RE	118	0.71 (0.23)	1.82	0.67	0.00	68.20	96

Table 5-19: Validation of 3T Base Condition Models

Crash Type	Observed Crashes	Calibration Factor (coefficient of variation)	Dispersion	MAD	Modified R^2	CURE max dev	CURE % dev
TOT Option 1	76	1.34 (0.18)	0.46	0.97	0.51	29.49	86
TOT Option 2	76	1.03 (0.17)	0.34	0.97	0.61	26.55	89
MVN	69	1.50 (0.17)	0.32	0.87	0.64	23.29	89

Table 5-20: Validation of 4D Base Condition Models

Crash Type	Observed Crashes	Calibration Factor (coefficient of variation)	Dispersion	MAD	Modified R^2	CURE max dev	CURE % dev
TOT Option 1	308	2.07 (0.27)	1.21	3.02	0.28	103.08	86
TOT Option 2	308	1.13 (0.22)	0.75	2.52	0.40	88.49	92
MVN	253	1.07 (0.30)	1.16	2.56	0.10	63.49	85

Table 5-21: Validation of 4U Base Condition Models

Crash Type	Observed Crashes	Calibration Factor (coefficient of variation)	Dispersion	MAD	Modified R^2	CURE max dev	CURE % dev
TOT Option 1	160	1.57 (0.25)	1.79	1.37	0.10	66.36	93
TOT Option 2	160	1.54 (0.21)	1.17	1.34	0.02	72.04	95
MVN	127	2.21 (0.24)	1.46	1.20	0.00	61.90	96

5.1.3.2 Final Models

Given the reasonable results of the validation exercise and the paucity of the initial model estimation data, we re-estimated all base condition models using the combined Ohio and Minnesota base condition sites. We re-estimated all initial crash type models, as well as estimating models by crash severity (all crash types combined).

As with the initial base model development, we used only sites with no lighting, parking, or automated enforcement to develop the base condition SPFs. Because no sites had zero roadside fixed objects and few divided roadways had a median width of exactly 15 feet, we included these variables in the models only if we considered them appropriate for the crash type, and if the variable was statistically significant in the model and with the expected direction of effect. If we included a variable, we would set it to the base condition for application. We attempted to include driveway density in all models developed, acknowledging that driveway presence might affect different crash types in different ways. We entered the number of driveways in a segment in those models in which it was included—that is, where there was no base condition for the number of driveways in a segment.

Final base condition models were calibrated for the following crash types (as defined previously):

- TOT
- KABC
- KAB
- KA
- MVN
- RE
- SSD
- HO+SOD
- MVN OTHER
- SV
- NIGHT
- MVD

Table 5-22 provides the ranges of the AADT and the number of driveways per mile (DWYDENS) variables by site type for the combined Ohio and Minnesota data that we used to estimate the base condition models. Table 5-23 to Table 5-34 show the estimated coefficients for the final models.

Table 5-22: Range of Base Model Variables for Final Urban/Suburban Segment Models

Site Type	AADT	DWYDENS
2U	100 to 23,028	0.00 to 154.76
3T	1,356 to 23,780	0.00 to 150.00
4U	1,150 to 41,418	0.00 to 320.00
4D	256 to 52,830	0.00 to 170.21
5T	5,356 to 50,553	0.00 to 115.38

Table 5-23: Total for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-6.2667 (0.5588)	0.0275 (0.1142)	0.7566 (0.0605)	-0.5377 (0.0933)	-0.5136 (0.0707)	0.0038 (0.0019)	-	-
3T	-11.6510 (2.0472)	0.0161 (0.2024)	1.3322 (0.2182)	-0.4150 (0.2627)	-0.1955 (0.1712)	-	-	-
4U	-13.9903 (1.5040)	-0.5241 (0.1935)	1.5943 (0.1624)	0.0671 (0.1891)	-0.2778 (0.1056)	0.0087 (0.0023)	-	-
4D	-11.0228 (1.0556)	-0.1656 (0.1557)	1.2544 (0.1043)	-0.6630 (0.1534)	-0.4471 (0.1051)	0.0168 (0.0045)	-	-0.0072 (0.0029)
5T	-12.0014 (1.7516)	0.6171 (0.5269)	1.2897 (0.1764)	-0.7017 (0.2111)	-0.7813 (0.1451)	-	0.0132 (0.0085)	-

Crashes per year = $(length)exp^{(Alpha1+Ohio)}AADT^{(Beta1)}\ exp^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = $exp^{(Alpha2)}(length)^{(Beta2)}$

Table 5-24: KABC for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-5.7643 (0.6944)	-0.0920 (0.1450)	0.6064 (0.0754)	-0.4355 (0.1201)	-0.5853 (0.0992)	-	-	-
3T	-13.1435 (2.8117)	-0.2039 (0.2670)	1.3799 (0.3002)	-0.1284 (0.3599)	-0.0128 (0.2931)	-	-	-
4U	-15.1227 (1.8518)	-0.5914 (0.2498)	1.6901 (0.1997)	-0.5771 (0.2726)	-0.6791 (0.1551)	0.0046 (0.0029)	-0.0184 (0.0054)	-
4D	-10.6728 (1.2521)	-0.3482 (0.1812)	1.1237 (0.1230)	-0.7472 (0.2009)	-0.5053 (0.1587)	0.0103 (0.0055)	-	-0.0110 (0.0037)
5T	-13.3013 (1.9432)	0.2684 (0.6352)	1.3873 (0.1944)	-0.7945 (0.2605)	-0.8925 (0.1924)	-	-	-

Crashes per year = $(length)exp^{(Alpha1+Ohio)}AADT^{(Beta1)}\ exp^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = $exp^{(Alpha2)}(length)^{(Beta2)}$

Table 5-25: KAB for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-5.6569 (0.8238)	0.5840 (0.2042)	0.4691 (0.0883)	-0.2961 (0.1357)	-0.6501 (0.1446)	-	-	-
3T	-14.3398 (3.5217)	-0.0055 (0.3474)	1.4226 (0.3751)	-0.2135 (0.4482)	0.0325 (0.4309)	-	-	-
4U	-15.5357 (2.2757)	0.0737 (0.3448)	1.5342 (0.2437)	-0.4129 (0.3578)	-0.6175 (0.2534)	-	-	-
4D	-11.2921 (1.4475)	0.1675 (0.2328)	1.0759 (0.1405)	-0.8647 (0.2650)	-0.6098 (0.2221)	0.0092 (0.0062)	-	-0.0074 (0.0042)
5T	-14.9040 (2.3332)	0.9934 (0.8272)	1.4140 (0.2278)	-0.6481 (0.2976)	-0.6560 (0.2486)	-	-	-

Crashes per year = $(length)exp^{(Alpha1+Ohio)}AADT^{(Beta1)}\ exp^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = $exp^{(Alpha2)}(length)^{(Beta2)}$

For KA crashes, no model was calibrated for four-lane divided (4D) or four-lane plus two-way left-turn-lane (5T) sites. For these, the appropriate proportion of KA crashes could be applied as a multiplicative factor with the TOT, KABC, or KAB model. Alternatively, the KA model for average conditions presented in the appendices could be applied.

Table 5-26: KA for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-6.1199 (1.3024)	0.6025 (0.3516)	0.3677 (0.1387)	-0.2131 (0.2868)	-0.7058 (0.3038)	-	-	-
3T	-20.3972 (7.7770)	0.4340 (0.8545)	1.8815 (0.8258)	0.7946 (0.6013)	-0.5334 (0.6767)	-	-	-
4U	-17.4579 (3.8002)	1.4088 (1.1221)	1.4692 (0.4016)	-1.3176 (1.2690)	-1.7818 (0.8893)	-	-	-
4D	-	-	-	-	-	-	-	-
5T	-	-	-	-	-	-	-	-

Crashes per year = $(\text{length})\exp^{(\text{Alpha1+Ohio})}\text{AADT}^{(\text{Beta1})}\ \exp^{(\text{Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid})}$

The dispersion parameter is modeled as: Dispersion parameter = $\exp^{(\text{Alpha2})}(\text{length})^{(\text{Beta2})}$

Table 5-27: MVN for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-13.1934 (0.7523)	0.0825 (0.1329)	1.4403 (0.0802)	-0.6182 (0.1179)	-0.5753 (0.0849)	0.0069 (0.0020)	-	-
3T	-15.4009 (2.4601)	-0.3760 (0.2229)	1.7234 (0.2626)	-0.2706 (0.2918)	-0.2234 (0.1849)	-	-	-
4U	-16.3961 (1.7182)	-0.8820 (0.2158)	1.8756 (0.1866)	-0.0044 (0.1954)	-0.3995 (0.1068)	-	-	-
4D	-14.9113 (1.3107)	-0.2680 (0.1919)	1.5965 (0.1269)	-0.4376 (0.1604)	-0.4917 (0.1092)	-	0.0112 (0.0046)	-0.0059 (0.0035)
5T	-15.6856 (1.9567)	0.6385 (0.5866)	1.6077 (0.1925)	-0.6279 (0.2182)	-0.8216 (0.1552)	-	0.0177 (0.0096)	-

Crashes per year = $(\text{length})\exp^{(\text{Alpha1+Ohio})}\text{AADT}^{(\text{Beta1})}\ \exp^{(\text{Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid})}$

The dispersion parameter is modeled as: Dispersion parameter = $\exp^{(\text{Alpha2})}(\text{length})^{(\text{Beta2})}$

Table 5-28: RE for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-17.2835 (1.0139)	0.1802 (0.1628)	1.8433 (0.1077)	-0.2692 (0.1303)	-0.5029 (0.0957)	-	-	-
3T	-19.7152 (3.3256)	-0.2954 (0.2896)	2.1326 (0.3546)	0.1870 (0.2908)	-0.2297 (0.1936)	-	-	-
4U	-19.9318 (2.0767)	-0.6741 (0.2526)	2.1519 (0.2226)	0.1871 (0.2237)	-0.3413 (0.1299)	-	-	-
4D	-22.5573 (1.8509)	-0.1243 (0.2366)	2.3241 (0.1842)	0.0222 (0.1695)	-0.5113 (0.1182)	-	-	-
5T	-18.8071 (2.3175)	0.7287 (0.6390)	1.9239 (0.2318)	-0.1217 (0.2342)	-0.5654 (0.1705)	-	-	-

Crashes per year = $(\text{length})\exp^{(\text{Alpha1+Ohio})}\text{AADT}^{(\text{Beta1})}\ \exp^{(\text{Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid})}$

The dispersion parameter is modeled as: Dispersion parameter = $\exp^{(\text{Alpha2})}(\text{length})^{(\text{Beta2})}$

For SSD crashes, a base condition model was not successfully calibrated for 3T sites. We calibrated the recommended model for 3T sites by using all sites and representing average conditions.

Table 5-29: SSD for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-14.1461 (1.9684)	-0.1494 (0.3208)	1.2943 (0.2097)	0.4303 (0.2930)	-0.4248 (0.2587)	-	-	-
3T	-14.4915 (3.2424)	-0.7704 (0.3049)	1.3985 (0.3481)	-0.5623 (0.5474)	-0.7902 (0.3238)	-	-	-
4U	-20.2134 (2.7017)	-0.7956 (0.3278)	2.0999 (0.2903)	0.0841 (0.3067)	-0.5012 (0.1882)	-	-	-
4D	-10.1287 (1.5856)	0.1939 (0.2382)	0.9255 (0.1557)	-0.3942 (0.2456)	0	-	-	-
5T	-14.7441 (2.5240)	0.7764 (0.7739)	1.3932 (0.2492)	-0.4866 (0.3002)	-0.2846 (0.2685)	-	-	-

Crashes per year = $(\text{length})\exp^{(\text{Alpha1+Ohio})}\text{AADT}^{(\text{Beta1})}\ \exp^{(\text{Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid})}$

The dispersion parameter is modeled as: Dispersion parameter = $\exp^{(\text{Alpha2})}(\text{length})^{(\text{Beta2})}$

For SOD+HO crashes, a base condition model was not successfully calibrated for 4D sites. We calibrated the recommended model for 4D sites by using all sites and representing average conditions.

Table 5-30: SOD+HO for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-8.1147 (1.2688)	-0.0461 (0.2454)	0.6884 (0.1363)	-0.0349 (0.2264)	-0.4037 (0.2497)	-	-	-
3T	-18.2308 (6.0061)	-0.7286 (0.4845)	1.7639 (0.6379)	-0.6224 (1.1807)	0.4519 (1.2972)	-	-	-
4U	-12.2573 (2.8768)	-0.4853 (0.4100)	1.1343 (0.3100)	-0.4689 (0.5951)	-0.4739 (0.4805)	-	-	-
4D	-8.5679 (2.0962)	-1.0946 (0.2278)	0.7000 (0.2100)	-0.2926 (0.5993)	-0.5178 (0.3654)	-	-	-
5T	-9.6385 (3.4905)	-0.5459 (0.7628)	0.8631 (0.3551)	-0.4456 (0.5536)	0	-	-	-

Crashes per year = $(\text{length})\exp^{(\text{Alpha1+Ohio})}\text{AADT}^{(\text{Beta1})}\ \exp^{(\text{Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid})}$

The dispersion parameter is modeled as: Dispersion parameter = $\exp^{(\text{Alpha2})}(\text{length})^{(\text{Beta2})}$

Table 5-31: MVN OTHER for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-11.8140 (1.2722)	0.4828 (0.2885)	1.0308 (0.1330)	-0.2403 (0.2129)	-1.0218 (0.1557)	0.0125 (0.0043)	-	-
3T	-9.9762 (3.1817)	-0.5783 (0.2896)	0.9931 (0.3402)	-1.1242 (0.7883)	0.0000 (n/a)	-	-	-
4U	-12.9084 (2.5712)	-1.1735 (0.3325)	1.3778 (0.2820)	0.4001 (0.2544)	-0.5018 (0.1620)	-	-	-
4D	-9.7538 (1.8069)	-0.0008 (0.2730)	0.8329 (0.1708)	-0.7641 (0.3445)	-0.4188 (0.2943)	0.0228 (0.0059)	-0.0156 (0.0055)	-
5T	-10.7916 (2.6197)	0.3563 (0.8276)	0.9049 (0.2556)	-0.6876 (0.3087)	-1.0932 (0.2338)	0.0351 (0.0121)	-	-

Crashes per year = (length)exp$^{(Alpha1+Ohio)}$AADT$^{(Beta1)}$ exp$^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = exp$^{(Alpha2)}$(length)$^{(Beta2)}$

For SV crashes, a base condition model was not successfully calibrated for 2U sites, nor was an SV model calibrated for average condition sites. For 2U, we recommend applying the total crash model, with the proportion of SV crashes applied as a multiplier.

Table 5-32: SV for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	Use Total model with proportion of SV crashes applied							
3T	-5.8853 (3.1236)	0.5923 (0.4389)	0.4605 (0.3329)	-0.1893 (0.4213)	-0.2883 (0.3431)	-	-	-
4U	-8.0865 (2.0681)	-0.3745 (0.2859)	0.7804 (0.2230)	-0.1446 (0.3820)	-0.2903 (0.2575)	-	-	-
4D	-6.4017 (1.1787)	0.0008 (0.1957)	0.6158 (0.1174)	-0.7961 (0.2256)	-0.4715 (0.2323)	-	-	-
5T	-2.5500 (2.5756)	1.0495 (0.8577)	0.1118 (0.2599)	-0.2065 (0.2853)	-0.5637 (0.2571)	-	-	-

Crashes per year = (length)exp$^{(Alpha1+Ohio)}$AADT$^{(Beta1)}$ exp$^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = exp$^{(Alpha2)}$(length)$^{(Beta2)}$

Table 5-33: Nighttime for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-3.5624 (0.8332)	-0.4718 (0.1697)	0.3012 (0.0913)	-0.2936 (0.1619)	-0.5305 (0.1519)	-	-	-
3T	-11.6068 (3.6625)	-0.8093 (0.3293)	1.1744 (0.3908)	-0.0771 (0.5609)	-0.1357 (0.3601)	-	-	-
4U	-15.2872 (2.4074)	-1.0078 (0.2853)	1.5836 (0.2584)	0.0294 (0.3401)	-0.3047 (0.2428)	-	-	-
4D	-9.3164 (1.5701)	-0.5921 (0.2011)	0.9517 (0.1531)	-0.6972 (0.2836)	-0.3866 (0.2297)	-0.0154 (0.0047)	-	-
5T	-12.0075 (2.2805)	0.2310 (0.7202)	1.1560 (0.2267)	-0.8875 (0.3773)	-0.7748 (0.2775)	-	-	-

Crashes per year = (length)exp$^{(Alpha1+Ohio)}$AADT$^{(Beta1)}$ exp$^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = exp$^{(Alpha2)}$(length)$^{(Beta2)}$

Table 5-34: MVD for Base Conditions Combined Data (Urban/Suburban Arterial Segments)

Site Type	Alpha1	Ohio	Beta1	Alpha2	Beta2	Beta3	Beta4	Beta5
2U	-13.6423 (1.4577)	0.5823 (0.2970)	1.2126 (0.1530)	0.1486 (0.1816)	-0.7353 (0.1621)	0.0177 (0.0038)	-	-
3T	-12.6716 (4.2048)	1.2663 (0.5717)	1.1270 (0.4444)	0.5494 (0.3448)	-0.1044 (0.3036)	-	-	-
4U	-18.7490 (2.9281)	0.0410 (0.4100)	1.7873 (0.3133)	0.3580 (0.2914)	-0.5745 (0.2016)	0.0183 (0.0049)	-	-
4D	-11.5730 (3.3464)	0.3881 (0.6447)	0.9784 (0.3263)	0.3945 (0.3450)	-0.7538 (0.2827)	0.0562 (0.0112)	-0.0472 (0.0130)	-
5T	-12.4484 (2.7823)	1.4306 (1.0530)	1.0472 (0.2702)	-0.4074 (0.2826)	-1.0384 (0.2437)	0.0186 (0.0063)	-	-

Crashes per year = (length)exp$^{(Alpha1+Ohio)}$AADT$^{(Beta1)}$ exp$^{(Beta3*dwydens+ +Beta4*fodensity+Beta5*medwid)}$

The dispersion parameter is modeled as: Dispersion parameter = exp$^{(Alpha2)}$(length)$^{(Beta2)}$

5.2 INTERSECTIONS

5.2.1 Estimation Data

Models have been estimated for four-leg signal-controlled (4SG) intersections, three-leg signal-controlled (3SG) intersections, four-leg stop-controlled (4ST) intersections, and three-leg stop-controlled (3ST) intersections on urban and suburban arterials, based on Ohio data for 2009–11. We estimated three sets of models:

- **Average condition models** use data from all the sites but use only AADT as the independent variable.
- **Base condition models** use data only from base conditions and use only AADT as the independent variable.
- **Fully specified models** use AADT and other intersection characteristics as independent variables.

Only the base condition models are reported here; for the average condition models, see Appendix A.

Ohio DOT provided the traffic volume, crash data, and other site characteristics they had compiled to calibrate the prediction models from the HSM. While the major roads leading to the intersections were predominantly state maintained, the minor roads often were not. For the quality of minor-road AADT, Ohio DOT provided two confidence levels. If the minor-road AADT was based on actual counts, the confidence level was considered high. If it was based on the functional class, then the confidence level was considered medium. To have a sufficient sample of sites for estimating the prediction models, we included intersections with both high and medium confidence levels for the minor-road AADT.

We estimated models for the following crash types:

- Total (TOT)
- Same direction (SD)
- Opposite direction (OD)
- Intersecting direction (ID)
- Single vehicle (SV)

For these individual crash types, we estimated models for KABCO, KABC, KAB, and KA severity levels. Models for pedestrian and bicycle crashes were not estimated because pedestrian and bicycle volumes were lacking at most of the intersections.

Table 5-35 shows the distribution of categorical variables (for example, number of turn lanes) for the four intersection types. Table 5-36 to Table 5-39 show the summary statistics for the data used for estimating the prediction models, which are based on all three years of data for each site.

In setting base conditions, we sought to use the same conditions as in the current HSM chapter for urban and suburban arterial road intersections.

For signal-controlled intersections, we defined the base conditions as follows:

- No left-turn lanes
- No right-turn lanes

- No right-turn-on-red prohibition (that is, right-turn-on-red is allowed on all legs)
- No red light cameras
- Lighting is present. (Note: This is different from what is currently in the HSM. As shown in Table 5-35, most of the signal-controlled intersections had lighting, which left us with an insufficient sample of intersections without lighting.)

For stop-controlled intersections, we defined the base conditions as follows:

- No left-turn lanes
- No right-turn lanes
- No lighting
- No schools within 1,000 feet
- No bus stops within 1,000 feet
- No alcohol sales establishments within 1,000 feet

Table 5-35: Distribution of categorical variables by intersection type (urban/suburban arterials)

Variable		3SG	3ST	4SG	4ST
Number of legs with left-turn lanes	0	485	7214	803	2342
	1	301	315	210	74
	2	189	48	692	106
	3	0	0	323	11
	4	0	0	734	2
Number of legs with right-turn lanes	0	721	7470	1985	2466
	1	204	101	430	59
	2	50	6	243	8
	3	0	0	68	2
	4	0	0	36	0
Number of legs with left-turn lanes on major road	0	619	7282	998	2374
	1	323	282	331	69
	2	33	13	1433	92
Number of legs with right-turn lanes on major road	0	865	7523	2286	2496
	1	105	54	359	38
	2	5	0	117	1
Number of legs with left-turn lanes on minor road	0	703	7481	1396	2474
	1	254	89	430	48
	2	18	7	936	13
Number of legs with right-turn lanes on minor road	0	792	7518	2221	2498
	1	177	59	411	33
	2	6	0	130	4
Lighting	Not Present	91	2407	278	680
	Present	884	5170	2484	1855
Number of approaches prohibiting right-turn-on-red	0	852	7574	2454	2532
	1	84	0	98	0
	2	39	0	79	0
	3	0	0	35	0
	4	0	0	96	0
Red light camera	Not Present	963	7576	2708	2535
	Present	12	0	54	0
Schools within 1000 feet	Not Present	849	6961	2420	2289
	Present	126	616	342	246
Number of liquor stores within 1000 feet	0	937	7341	2559	2437
	1 to 8	38	236	203	98
Number of bus stops within 1000 feet	0	707	6615	2322	2318
	1 or 2	32	179	101	50
	3 or more	236	783	339	167

Table 5-36: Descriptive statistics for base condition SPFs (3SG: 345 intersections)

Variable	Number of crashes	Mean	Std. Dev	Minimum	Maximum
Major road AADT		12,363	4949	3050	32,109
Minor road AADT		4077	3026	110	18,415
Total Intersection AADT		16,440	5989	4440	44,345
Ratio of Minor to Total Intersection AADT		0.25	0.13	0.02	0.5
KA	62	0.18	0.42	0	2
KAB	375	1.09	1.37	0	9
KABC	854	2.48	2.47	0	13
KABCO	4026	11.67	9.14	0	52
SV_KA	13	0.04	0.19	0	1
SV_KAB	47	0.14	0.38	0	3
SV_KABC	67	0.19	0.47	0	4
SV_KABCO	253	0.73	1.21	0	15
SD_KA	22	0.06	0.24	0	1
SD_KAB	158	0.46	0.75	0	4
SD_KABC	424	1.23	1.38	0	6
SD_KABCO	2302	6.67	5.8	0	39
OD_KA	10	0.03	0.17	0	1
OD_KAB	60	0.17	0.53	0	6
OD_KABC	99	0.29	0.68	0	7
OD_KABCO	369	1.07	1.47	0	11
ID_KA	17	0.05	0.24	0	2
ID_KAB	108	0.31	0.73	0	6
ID_KABC	253	0.73	1.36	0	10
ID_KABCO	974	2.82	3.52	0	24

Table 5-37: Descriptive statistics for base condition SPFs (3ST: 2082 intersections)

Variable	Number of crashes	Mean	Std Dev	Minimum	Maximum
Major road AADT		8187	5221	270	38,460
Minor road AADT		2137	1400	33	18,460
Total Intersection AADT		10,324	5810	540	56,920
Ratio of Minor to Total Intersection AADT		0.23	0.12	0	0.5
KA	198	0.1	0.32	0	3
KAB	840	0.4	0.78	0	7
KABC	1422	0.68	1.15	0	11
KABCO	4756	2.28	3.47	0	49
SV_KA	59	0.03	0.18	0	2
SV_KAB	222	0.11	0.36	0	5
SV_KABC	297	0.14	0.43	0	6
SV_KABCO	952	0.46	0.87	0	13
SD_KA	52	0.02	0.16	0	2
SD_KAB	323	0.16	0.48	0	6
SD_KABC	661	0.32	0.75	0	9
SD_KABCO	2390	1.15	2.37	0	31
OD_KA	43	0.02	0.14	0	1
OD_KAB	128	0.06	0.25	0	2
OD_KABC	184	0.09	0.3	0	3
OD_KABCO	453	0.22	0.54	0	4
ID_KA	43	0.02	0.16	0	3
ID_KAB	163	0.08	0.33	0	4
ID_KABC	272	0.13	0.49	0	9
ID_KABCO	885	0.43	1.08	0	14

Table 5-38: Descriptive statistics for base condition SPFs (4SG: 589 intersections)

Variable	Number of crashes	Mean	Std. Dev	Minimum	Maximum
Major road AADT		11,067	5650	1810	34,960
Minor road AADT		3803	3167	72	27,228
Total Intersection AADT		14,870	7344	2061	56,488
Ratio of Minor to Total Intersection AADT		0.25	0.13	0.01	0.5
KA	148	0.25	0.56	0	4
KAB	767	1.3	1.79	0	14
KABC	1798	3.05	3.67	0	35
KABCO	7253	12.31	12.7	0	109
SV_KA	16	0.03	0.16	0	1
SV_KAB	73	0.12	0.37	0	3
SV_KABC	112	0.19	0.48	0	3
SV_KABCO	409	0.69	1.05	0	9
SD_KA	53	0.09	0.33	0	3
SD_KAB	283	0.48	0.92	0	8
SD_KABC	868	1.47	2.15	0	18
SD_KABCO	3964	6.73	8.32	0	76
OD_KA	27	0.05	0.23	0	2
OD_KAB	167	0.28	0.74	0	6
OD_KABC	309	0.52	1.14	0	10
OD_KABCO	1021	1.73	2.81	0	25
ID_KA	51	0.09	0.29	0	2
ID_KAB	239	0.41	0.83	0	8
ID_KABC	483	0.82	1.34	0	8
ID_KABCO	1671	2.84	3.36	0	26

Table 5-39: Descriptive statistics for base condition SPFs (4ST: 551 intersections)

Variable	Number of crashes	Mean	Std Dev	Minimum	Maximum
Major road AADT		8251	6179	450	37,301
Minor road AADT		2088	1459	50	13,773
Total Intersection AADT		10,339	6658	810	40,111
Ratio of Minor to Total Intersection AADT		0.23	0.13	0	0.5
KA	120	0.22	0.56	0	4
KAB	432	0.78	1.39	0	9
KABC	706	1.28	1.99	0	16
KABCO	1931	3.5	4.58	0	51
SV_KA	20	0.04	0.2	0	2
SV_KAB	61	0.11	0.37	0	2
SV_KABC	72	0.13	0.39	0	2
SV_KABCO	265	0.48	0.81	0	5
SD_KA	15	0.03	0.18	0	2
SD_KAB	84	0.15	0.46	0	4
SD_KABC	219	0.4	1.05	0	14
SD_KABCO	720	1.31	2.91	0	46
OD_KA	21	0.04	0.2	0	2
OD_KAB	60	0.11	0.38	0	3
OD_KABC	83	0.15	0.44	0	3
OD_KABCO	214	0.39	0.82	0	6
ID_KA	64	0.12	0.41	0	4
ID_KAB	225	0.41	0.99	0	7
ID_KABC	328	0.6	1.31	0	9
ID_KABCO	705	1.28	2.24	0	15

5.2.2 Estimated Models

We estimated all the models using negative binomial regression with a constant overdispersion parameter and the traditional log-linear framework. Most previous studies on this topic have used a *power* function, which provides limited flexibility in the functional form. In this section, we used the *Hoerl* function to provide more flexibility in the functional form (Hauer, 2015). The Hoerl function allows the relationship to have a convex/concave shape with inflection points. With it, the dependent variable (Y) is related to the independent variable (X) in the following way:

$$Y = \exp[a_1 + a_2 X + a_3 \ln(X)] \text{ or } Y = e^{a_1} e^{a_2 X} X^{a_3}, \quad (5\text{-}1)$$

where a_1, a_2, and a_3 are parameters to be estimated. We examined two functional forms, Model A and Model B.

Model A included as the starting point the following independent variables in the following form:

$$Y = e^{a} \times e^{b \times \left(\frac{AADT_{maj}}{10000}\right)} \times (AADT_{maj})^{c} \times e^{d \times \left(\frac{AADT_{min}}{10000}\right)} \times (AADT_{min})^{e} \tag{5-2}$$

Model B included as the starting point the following independent variables in the following form:

$$Y = e^{a} \times e^{b \times \left(\frac{AADT_{tot}}{10000}\right)} \times (AADT_{tot})^{c} \times e^{d \times \left(\frac{AADT_{min}}{AADT_{tot}}\right)} \times \left(\frac{AADT_{min}}{AADT_{tot}}\right)^{e} \tag{5-3}$$

where Y is the predicted number of crashes in one year, and a, b, c, d, and e are parameters to be estimated.

For both model forms A and B, the estimation started with all the variables presented above, and, through backward elimination, variables that were not statistically significant were removed. The final models from both forms (that is, forms A and B) are presented below.

Table 5-40 to Table 5-43 provide the model results, including parameter estimates (that is, coefficients a, b, c, d, and e), standard errors (in parentheses), and the overdispersion parameter (k). For some crash types, models could not be estimated, or they did not converge. In some cases, the overdispersion parameter was quite high (exceeding 2). These models should be used with caution. For the crash types for which models could not be estimated, we recommend using the prediction for the next closest model and multiplying the prediction by the proportion of that crash type. For example, if a prediction model for KA crashes is not available, but one for KAB crashes is, the prediction for KA crashes can be obtained by the following equation:

$$Predicted\ KA\ crashes = Predicted\ KAB\ crashes \times \left(\frac{Number\ of\ KA\ crashes\ in\ the\ data\ set}{Number\ of\ KAB\ crashes\ in\ the\ data\ set}\right) \tag{5-4}$$

Table 5-40: Prediction Models for 3SG Intersections

Crash Type	Severity	Model Form	a (S.E.)	b (S.E.)	c (S.E.)	d (S.E.)	e (S.E.)	k (S.E.)
All	KABCO	B	-4.5704 (1.0173)		.6366 (.1051)		.1519 (.0618)	.4669 (.0432)
All	KABC	A	-6.7956 (1.3109)		.4799 (.1274)		.2585 (.0725)	.5344 (.0805)
All	KAB	A	-8.0554 (1.6897)		.5062 (.1660)		.2814 (.0916)	.5745 (.1336)
All	KA	A	Model did not converge					
SV	KABCO	A	-2.3447 (.2106)	.3894 (.1448)		.9168 (.1964)		.4113 (.1444)
SV	KABC		Model did not converge					
SV	KAB		Model did not converge					
SV	KA		Model did not converge					
SD	KABCO	A	-6.2255 (1.0384)		.5414 (.1034)		.2390 (.0575)	.4615 (.0486)
SD	KABC	A	-9.1985 (1.5615)		0.6682 (0.1517)		0.2495 (0.0828)	0.3761 (0.1029)
SD	KAB	A	-9.3282 (2.2931)		.5844 (.2243)		.2413 (.1223)	.4521 (.2344)
SD	KA		Model did not converge					
OD	KABCO	B	-10.0017 (1.9351)		0.9248 (0.1991)			0.7486 (0.1512)
OD	KABC	A	-17.9744 (3.5381)		1.3504 (0.3371)		0.3523 (0.1741)	1.1826 (0.4533)
OD	KAB	A	-21.2395 (4.5175)		1.4846 (.4257)		.5304 (.2182)	1.4144 (.6955)
OD	KA		Model did not converge					
ID	KABCO	B	-2.3636 (1.6005)		0.2385 (0.1658)			1.0859 (0.1221)
ID	KABC	B	-2.9487 (2.3752)		0.1596 (0.2460)			1.7788 (0.3222)
ID	KAB	B	-4.1873 (3.2704)		0.1997 (0.3385)			2.2076 (0.6438)
ID	KA	B	-8.9578 (7.2945)		0.5014 (0.7521)			4.3772 (4.1273)

Table 5-41: Prediction models for 3ST intersections

Crash Type	Severity	Model Form	a (S.E.)	b (S.E.)	c (S.E.)	d (S.E.)	e (S.E.)	k (S.E.)
All	KABCO	A	-3.1275 (0.8631)	0.6210 (0.1209)	0.2319 (0.1083)	0.7280 (0.1754)		0.8087 (0.0442)
All	KABC	B	-2.3919 (.0706)	.7690 (.0514)				.7615 (.0838)
All	KAB	B	-2.7900 (.0829)	.6705 (.0593)				.8031 (.1266)
All	KA	B	-4.0506 (.1482)	.5272 (.1044)				.8594 (.4459)
SV	KABCO	B	1.0576 (1.3977)	0.3861 (0.1658)	-0.3676 (0.1711)			1.0986 (0.1316)
SV	KABC	B	-0.0628 (2.1687)	0.2730 (0.2626)	-0.3596 (0.2659)			1.4458 (0.3783)
SV	KAB	B	-1.8441 (1.1345)		-0.1647 (0.1252)			1.6069 (0.5129)
SV	KA	B	-2.0986 (2.1063)		-0.2835 (0.2334)			4.0221 (2.6334)
SD	KABCO	B	-14.8383 (.6258)		1.4636 (.0689)		-.1385 (.0539)	1.1428 (.0803)
SD	KABC	B	-14.2585 (.9668)		1.3176 (.0989)	-1.1784 (.4720)		.9478 (.1552)
SD	KAB	B	-14.1222 (1.3007)		1.2298 (.1327)	-1.2920 (.6521)		1.3071 (.3124)
SD	KA	B	-11.5340 (2.5011)		.7330 (.2685)			1.7962 (2.0336)
OD	KABCO	B	-3.3353 (.1100)	0.6177 (.0794)				1.1523 (.2364)
OD	KABC	B	-4.1549 (.1470)	.5514 (.1003)				.2950 (.3737)
OD	KAB	B	-4.4870 (.1778)	.5270 (.1231)				.6168 (.6314)
OD	KA	B	Model did not converge					
ID	KABCO	A	-11.1651 (0.8035)		0.8094 (0.0849)		0.2535 (0.0719)	2.2740 (0.2200)
ID	KABC	A	-12.7177 (1.2614)		0.8967 (0.1316)		0.1974 (0.1118)	3.3635 (0.6204)
ID	KAB	A	-14.4692 (1.5707)		0.9112 (0.1608)		0.3412 (0.1432)	3.0787 (0.8626)
ID	KA	A	-17.4865 (3.1341)		1.0812 (0.3051)		0.3558 (0.3044)	7.9899 (4.2722)

Table 5-42: Prediction models for 4SG intersections

Crash Type	Severity	Model Form	a (S.E.)	b (S.E.)	c (S.E.)	d (S.E.)	e (S.E.)	k (S.E.)
All	KABCO	B	-7.4359 (.6531)		.9218 (.0685)			.5514 (.0381)
All	KABC	B	-10.5443 (.8782)		1.0989 (.0915)			.6386 (.0628)
All	KAB	B	-9.9857 (1.0898)		.9535 (.1134)			.7440 (.1031)
All	KA	B	-9.6739 (1.8404)		.7511 (.1907)			.6997 (.3296)
SV	KABCO	B	-4.3216 (1.2070)		.3000 (.1264)			.7818 (.1586)
SV	KABC	B	-7.7339 (2.0833)		0.5209 (0.2169)			0.9105 (0.4749)
SV	KAB	B	-8.9332 (2.5001)		0.6011 (0.2597)			0.8532 (0.6746)
SV	KA		Model did not converge					
SD	KABCO	A	-8.2447 (0.6774)		0.9424 (0.0744)	0.6264 (0.1243)		0.5800 (0.0448)
SD	KABC	A	-14.2230 (1.0584)		1.2127 (0.1114)		0.2693 (0.0657)	0.6257 (0.0848)
SD	KAB	A	-15.2404 (1.6168)		1.2210 (0.1695)		0.2476 (0.1001)	0.8485 (0.2019)
SD	KA	A	-14.3865 (3.1409)		.7926 (.3312)		.4313 (.2196)	1.7976 (1.1221)
OD	KABCO	A	-9.7053 (1.0998)		.7364 (.1173)		.2867 (.0767)	1.1587 (.1190)
OD	KABC	B	-13.5030 (1.8439)		1.2228 (0.1914)			1.8372 (0.3049)
OD	KAB	B	-12.7760 (2.1772)		1.0838 (.2256)			2.2166 (.5129)
OD	KA	B	-12.8419 (4.4636)		.9029 (.4616)			3.8585 (3.0410)
ID	KABCO	A	-1.5214 (.4578)	.3492 (.0863)			.1316 (.0594)	.8944 (.08100)
ID	KABC	A	-5.6212 (1.1967)		.2977 (.1319)		.1958 (.0860)	1.2440 (.1843)
ID	KAB	A	-6.1898 (1.6067)		.4390 (.1679)			1.4297 (.3125)
ID	KA	A	Model did not converge					

Table 5-43: Prediction models for 4ST crashes

Crash Type	Severity	Model Form	a (S.E.)	b (S.E.)	c (S.E.)	d (S.E.)	e (S.E.)	k (S.E.)
All	KABCO	A	-3.6743 (.6256)		.4071 (.0726)	.9208 (.3490)		1.0155 (.0867)
All	KABC	B	-3.9675 (.9735)		.3417 (.1067)			1.6020 (.1834)
All	KAB		Could not obtain useful model					
All	KA		Could not obtain useful model					
SV	KABCO	B	-2.2170 (.1253)	.3435 (.0889)				.5835 (.1913)
SV	KABC	B	-13.8618 (5.6055)	-1.0741 (0.6352)	1.3021 (0.6814)			1.5358 (0.8553)
SV	KAB	B	-12.0750 (5.9194)	-0.9275 (0.6816)	1.0714 (0.7211)			2.8132 (1.3122)
SV	KA	B	-8.3241 (3.5390)		0.4279 (0.3835)			1.7361 (2.8313)
SD	KABCO	A	-12.4690 (1.0120)		1.0633 (.1013)		.2661 (.0875)	1.1504 (.1471)
SD	KABC	A	-5.9134 (.9920)	.8168 (.1394)			.4033 (.1322)	1.8464 (.3910)
SD	KAB	A	-5.8261 (1.3016)	.3579 (.1865)			.3360 (.1742)	1.8506 (.7743)
SD	KA		Model did not converge					
OD	KABCO	B	-6.0829 (1.2859)		0.4417 (0.1399)			1.4996 (0.3422)
OD	KABC	B	-5.5548 (1.9029)		0.2814 (0.2078)			2.3460 (0.9251)
OD	KAB	B	-5.5578 (2.1901)		0.2462 (0.2393)			3.0857 (1.3915)
OD	KA		Model did not converge					
ID	KABCO		Could not obtain useful model					
ID	KABC		Could not obtain useful model					
ID	KAB		Could not obtain useful model					
ID	KA		Could not obtain useful model					

5.2.3 Validation of Models

To calibrate and validate the models estimated using the data from Ohio, we used six years of data (2010–15) from North Carolina. Some of the data for calibration were compiled as part of project funded by the North Carolina Department of Transportation (Smith et al. 2016). It was extremely difficult to find intersections that matched the base conditions. Among the 102 four-leg signal-controlled intersections that were identified, for example, only three matched the base conditions used to estimate the original SPFs. Similarly, among the 33 three-leg signal-controlled intersections identified, none matched the base conditions. Since the number of intersections with the "base conditions" were very small, the calibration/validation sample included all intersections (including both intersections whose characteristics matched the base conditions, and intersections whose characteristics did not match the base conditions). For those intersections whose characteristics did not match the base condition, the CMFs from the 1st edition of the HSM were used to adjust the predictions from the base models. As mentioned earlier, for the base models that were estimated using the Ohio data for signalized intersections, the base condition was "lighting present". However, the base condition in the 1st edition of the HSM was "lighting not present". Hence, for signalized intersections, the inverse of the CMF from the HSM was used.

We conducted calibration and validation for the following crash types:

- All crashes (KABCO)
- SV_KABCO
- SD_KABCO
- OD_KABCO
- ID_KABCO

The sample of crashes was limited for the other crash types, especially in the case of the stop-controlled intersections. Table 5-44 to Table 5-47 provide the summary of the data used in the calibration and validation.

Table 5-44: Summary of Calibration/Validation data set from North Carolina for 3SG intersections (33 intersections; 2010 to 2015)

Variable	Sum	Mean	Standard deviation	Minimum	Maximum
KABCO	839	25.42	29.40	0	171
SV_KABCO	83	2.52	2.27	0	7
SD_KABCO	466	14.12	19.98	0	116
OD_KABCO	119	3.61	5.67	0	31
ID_KABCO	179	5.42	5.24	0	18
Major AADT		14,935	7,801	3,198	32,208
Minor AADT		7,124	4,245	267	16,683
Total AADT		22,059	10,519	6,935	48,542

Table 5-45: Summary of Calibration/Validation data from North Carolina for 3ST intersections (52 intersections: 2010 to 2015 data)

Variable	Sum	Mean	Standard deviation	Minimum	Maximum
KABCO	304	5.85	8.10	0	36
SV_KABCO	55	1.06	1.55	0	7
SD_KABCO	124	2.38	4.43	0	19
OD_KABCO	48	0.92	1.59	0	8
ID_KABCO	87	1.67	2.38	0	10
Major AADT		7,682	7,232	67	45,733
Minor AADT		1,764	2,227	18	9,500
Total AADT		9,446	7,832	117	45,803

Table 5-46: Summary of Calibration/Validation data from North Carolina for 4SG intersections (102 intersections: 2010 to 2015)

Variable	Sum	Mean	Standard deviation	Minimum	Maximum
KABCO	6049	59.3	51.4	0	275
SV_KABCO	285	2.8	2.6	0	13
SD_KABCO	3605	35.3	39.4	0	229
OD_KABCO	752	7.4	6.4	0	35
ID_KABCO	1436	14.1	10.3	0	50
Major AADT		19,574	10,047	3,500	47,063
Minor AADT		10,053	6,492	15	33,625
Total AADT		29,627	14,308	6,200	69,433

Table 5-47: Summary of the Calibration/Validation data from North Carolina for 4ST intersections (55 intersections: 2010 to 2015 data)

Variable	Sum	Mean	Standard deviation	Minimum	Maximum
KABCO	464	8.92	7.48	0	35
SV_KABCO	48	0.92	1.12	0	4
SD_KABCO	156	3.00	2.69	0	15
OD_KABCO	51	0.98	1.30	0	6
ID_KABCO	210	4.04	5.24	0	29
Major AADT		7101	5339	335	29,375
Minor AADT		1122	1214	5	8,625
Total AADT		8223	5372	363	38,000

We considered two basic options for calibration and validation. The first was to estimate a calibration factor following the approach outline in the HSM. The second was to estimate a calibration function (Srinivasan et al. 2016). As discussed in Srinivasan et al. (2016), calibration functions can take many forms. For this effort, we used the following form:

$$Y = a \times \prod_{i=1}^{n} CMF_i \times (Base\ Pred)^b \tag{5-5}$$

where Y is the expected number of crashes, *Base Pred* is the prediction from the base model, and $\prod_{i=1}^{n} CMF_i$ represents the product of the CMFs from the HSM. This equation can also written as follows:

$$Y = \prod_{i=1}^{n} CMF_i \times \exp[\ln(a) + b \times ln(Base\ Pred)] \tag{5-6}$$

If b is close to 1, the calibration function will not provide any advantages over a simple calibration factor (if b = 1, the calibration factor is "a").

To estimate ln(a) and b, we used a negative binomial regression model. The dependent variable was the number of observed crashes at each site, and the independent variables included Base Pred. The product of the CMFs was included as an offset. For some of the intersection types, data were available on the North Carolina region (Coastal, Piedmont, or Mountain), and we included them in the negative binomial model in addition to Base Pred.

There are at least two ways of estimating a calibration function. One is to use the approach followed in Srinivasan et al. (2016), in which the solver tool in Excel is used to estimate the negative binomial regression using maximum likelihood estimation, with a constraint to ensure the total fitted values are equal to the total observed crash counts. The second is to use traditional tools, such as SAS PROC GENMOD, which also use maximum likelihood estimation, but without any constraints regarding the fitted and observed values. The second approach was used here, but before the calibration functions were evaluated using goodness-of-fit measures, they were calibrated in order to ensure that the total fitted values and the total observed crash counts are the same.

Table 5-48 to Table 5-51 provide information on the calibration factors and calibration functions estimated for 3SG, 3ST, 4SG, and 4ST intersections, respectively. After estimating both, we used "The Calibrator," a tool (already mentioned) developed by FHWA for calibrating and assessing SPFs, to obtain the following goodness-of-fit (GOF) measures:

- Modified R^2—higher values indicate better-fitting SPFs.
- MAD (mean absolute deviation)—lower values indicate better-fitting SPFs.
- Maximum absolute CURE deviation (MACD)—lower values indicate better-fitting SPFs.
- Percentage CURE deviation—lower values indicate better-fitting SPFs. The Calibrator recommends this be 5 percent or lower.

The Calibrator tool automatically calibrates a model before producing the GOF measures. For example, if a calibration function predicts a total of 845.5 crashes and the total observed crashes were 839, then a calibration factor of 839/845.5 (=0.992) is applied before the GOF measures are produced. In most cases, the calibration functions provided better GOF measures. This indicates agencies should consider using calibration functions if the calibration factors alone do not provide a reasonable fit for their sample data.

Table 5-48: Calibration/Validation Results for 3SG intersections

Crash Type	Option	Total Observed Crashes	Total Predicted Crashes	ln(a) (S.E.)	b (S.E.)	Calibration Factor	Modified R^2	MAD	MACD	% Cure Deviation
All	HSM calibration	839	765.1			1.097	0.20	13.42	135.8	33%
	Calibration function	839	845.5	-3.761 (1.430)	2.115 (0.426)	0.992	0.28	13.67	81.6	0%
SV	HSM calibration	83	62.1			1.337	0.00	1.98	18.1	18%
	Calibration function	83	82.8	0.984 (0.258)	0.226 (0.287)	1.002	0.05	1.78	6.7	3%
SD	HSM calibration	466	443.8			1.050	0.21	8.60	82.9	52%
	Calibration function	466	471.4	-3.242 (1.148)	2.113 (0.406)	0.989	0.27	8.78	52.2	3%
OD	HSM calibration	119	72.9			1.633	0.18	2.87	25.3	15%
	Calibration function	119	119.2	-0.247 (0.495)	1.636 (0.447)	0.999	0.22	2.99	13.6	3%
ID	HSM calibration	179	158.3			1.131	0.09	3.99	25.0	9%
	Calibration function	179	182.7	-9.498 (2.399)	6.276 (1.321)	0.979	0.45	3.25	10.5	3%

Table 5-49: Calibration/Validation results for 3ST intersections

Crash Type	Option	Total Observed Crashes	Total Predicted Crashes	ln(a) (S.E.)	b (S.E.)	Calibration Factor	Modified R^2	MAD	MACD	% Cure Deviation
All	HSM calibration	304	207.0			1.469	0.00	5.35	49.4	2%
	Calibration function	304	412.4	0.04512 (0.6234)	1.1633 (0.3058)	0.737	0.00	5.77	96.9	13%
SV	HSM calibration	55	48.8			1.126	0.00	1.23	20.9	83%
	Calibration function	55	60.0	0.5732 (0.4092)	-1.6434 (1.1889)	0.917	0.00	1.08	11.5	8%
SD	HSM calibration	124	70.4			1.761	0.04	2.45	16.6	0%
	Calibration function	124	124.3	0.7871 (0.5059)	0.5261 (0.2174)	0.997	0.18	2.51	17.6	0%
OD	HSM calibration	48	20.0			2.406	0.00	1.03	6.2	21%
	Calibration function	48	97.3	1.5399 (0.6837)	1.9334 (0.7517)	0.493	0.00	1.32	25.4	21%
ID	HSM calibration	87	28.0			3.110	0.13	1.64	12.5	2%
	Calibration function	87	92.2	0.3237 (0.4322)	0.7153 (0.2365)	0.943	0.20	1.47	13.6	6%

Table 5-50: Calibration/Validation results for 4SG intersections

Crash type	Option	Total Observed Crashes	Total Predicted Crashes	ln(a) (S.E.)	b (S.E.)	Calibration Factor	Modified R^2	MAD	MACD	% Cure Deviation
All	HSM calibration	6260	2753.1			2.270	0.23	32.91	447.2	25%
	Calibration function	6260	6507.2	-0.6905 (0.5942)	1.4591 (0.1528)	0.960	0.36	30.19	408.8	1%
SV	HSM calibration	285	105.2			2.708	0.00	2.08	38.3	17%
	Calibration function	285	286.8	0.0150 (0.3689)	2.8972 (0.5726)	0.994	0.24	1.85	17.4	1%
SD	HSM calibration	3605	2245.1			1.606	0.44	20.52	204.8	1%
	Calibration function	3605	3891.9	0.4928 (0.4031)	1.0693 (0.1069)	0.926	0.43	19.20	290.2	1%
OD	HSM calibration	752	413.5			1.819	0.02	4.72	78.2	20%
	Calibration function	752	768.0	1.9957 (0.2708)	0.4826 (0.1249)	0.979	0.15	4.54	36.8	1%
ID	HSM calibration	1436	542.3			2.648	0.03	7.54	99.8	12%
	Calibration function	1436	1488.1	1.5775 (0.4119)	0.8396 (0.1790)	0.965	0.11	7.26	85.4	4%

Table 5-51: Calibration/Validation Results for 4ST Intersections

Crash Type	Option	Total Observed Crashes	Total Predicted Crashes	ln(a) (S.E.)	b (S.E.)	Calibration Factor	Modified R^2	MAD	MACD	% Cure Deviation
All	HSM calibration	464	266.6			1.741	0.00	5.88	76.9	10%
	Calibration function	464	551.4	0.01803 (0.5403)	1.3714 (0.3097)	0.842	0.00	5.96	72.9	0%
SV	HSM calibration	48	41.3			1.163	0.00	0.81	4.8	2%
	Calibration function	48	48.6	0.1673 (0.1791)	1.0300 (0.6541)	0.987	0.00	0.81	5.0	2%
SD	HSM calibration	157	68.2			2.300	0.00	2.28	31.3	2%
	Calibration function	157	181.5	1.1655 (0.1535)	0.7739 (0.1651)	0.865	0.00	2.21	20.0	0
OD	HSM calibration	52	31.0			1.678	0.14	0.90	7.1	2%
	Calibration function	52	56.1	0.9375 (0.2568)	2.2908 (0.6165)	0.926	0.00	0.89	8.7	2%
ID	HSM calibration	213	97.3			2.189	0.00	3.62	49.0	44%
	Calibration function	213	250.6	0.9496 (0.3205)	0.9954 (0.4092)	0.850	0.00	3.62	48.9	42%

6 REVISITING THE HSM CALIBRATION APPROACH

6.1 APPROACHES CONSIDERED

6.1.1 Background on HSM Approach

The development of new models for the HSM, taken together with research conducted since its release in 2010 on key issues pertaining to the calibration procedure, provided the need and the opportunity to revisit that procedure in this research project with a view to updating it. The key issues, which are interrelated with others, pertain to the sample size for calibration data and to whether and how to capture the variation of the calibration factor with site characteristics. To address the latter issue, we investigated a procedure based on calibration functions.

A review of the research on establishing minimum sample sizes and estimating calibration functions, along with the results of an empirical investigation in this project, led to the proposed calibration procedure update documented here. The research review suggested that required samples will, indeed, vary across site types, jurisdictions, and crash types and severities. In particular, a consensus seemed apparent that the desirable minimum suggested in the HSM of 30–50 sites with at least 100 crashes a year might not be universally applicable. The research carried out since 2010 has not, however, provided any consistent guidance on what does constitute an appropriate sample. In some cases, recommended sample sizes are so large that a jurisdiction may be better off acquiring (or hiring) personnel with the skill sets required to estimate their own models directly rather than calibrate an external one. The sample size guidance in the procedure recommended here is based on a report by Bahar et al. *(*2014); even so, sample sizes based on that guidance are not directly estimated but, rather, are determined through an iterative assessment of the accuracy of the calibration factor.

The empirical investigation pursued in this project, in essence, evaluates the guidance in Bahar et al. (2014) by using various sample sizes in assessing and comparing combinations of the following three options that include exploration of calibration functions:

A) Estimating a single calibration factor (C)
B) Estimating a calibration function
C) Directly estimating a model using the calibration data

We performed different sets of analyses pertaining to these assessments and comparisons for four representative site types. For all analyses, we estimated a constant calibration factor using the HSM methodology as the sum of the model predictions divided by the sum of the observed crashes for the calibration data. We also estimated a calibration function in the following form, based on research by Srinivasan et al. (2016):

$$N_{\text{predicted}} = a \times (\text{Unadjusted Prediction})^{b} \quad (6\text{-}1)$$

This function, in effect, allows the calibration factor to vary from site to site, depending on site characteristics that affect the crash prediction, most notably traffic volume.

We applied two alternative approaches for this investigation, as described below.

6.1.2 Approach 1

We used three representative site types for this investigation: urban four-lane divided segments; urban two-lane divided segments; and rural two-lane, three-leg stop-controlled intersections. The final models estimated and presented in earlier chapters were calibrated to randomly selected sites from another jurisdiction to increase sample sizes. We also directly estimated models with model forms identical to those being calibrated, with the exception that we used a constant overdispersion parameter.

The logic behind this "iterative" approach was that, at small sample sizes, applying either a calibration factor or function to an original model would prove superior to using a directly estimated model. As sample sizes increased, there would be a point at which a directly calibrated model would perform better. At the other end of the spectrum, there would also be a point at which the sample size would be too small even to estimate a reliable calibration factor.

We evaluated the performance of a calibrated model using several criteria provided by the FHWA Calibrator spreadsheet tool (Lyon et al. 2016). The guidance this tool provided indicated a calibrated model is reasonable if either the coefficient of variation (CV) of the estimated calibration factor is 0.15 or less or if a cumulative residuals (CURE) plot for the fitted values has fewer than 5 percent of the data points outside of the two standard deviation limits. Other goodness-of-fit measures provided by the tool include the mean absolute deviation (MAD), modified R^2, a calibrated constant overdispersion parameter, and the maximum deviation from zero of the CURE plot for the fitted values.

6.1.3 Approach 2

This investigation assessed the temporal and spatial transferability and calibration of the models. In this case, all of the data available for the calibration were used rather than samples of various sizes. The site type investigated was multilane rural highways.

6.2 Approach 1 Results

6.2.1 Urban Four-Lane Divided Segments

We calibrated the model we developed for total crashes and average conditions using data from Ohio to randomly selected sites from Minnesota for increasing sample sizes.

Table 6.1 presents the results of the investigation for urban four-lane divided segments. It shows the number of sites and total crashes used and includes a number of measures, among them the calculated calibration factor (C) and its coefficient of variation (CV), the parameter estimates of the calibration function (a and b in Equation 6.1), and the goodness-of-fit measures, and it compares the three options as provided by the Calibrator tool.

A number of observations can be made from the results in Table 6-1:

1. For the three sample sizes investigated, the goodness-of-fit statistics are reasonably similar.
2. The maximum calibration factor CV value of 0.15 recommended in the Calibrator tool guidance is not reached until a sample size of 100 sites and 271 crashes. Most interesting is that, at smaller sample sizes, a model directly estimated for the Minnesota data was successful, and the goodness-of-fit statistics for all three options were comparable. This would seem to indicate that even if the CV is higher than 0.15, a directly estimated model may still be feasible.

3. The calibration function does perform better in general than a calibration factor, although the differences are not very large for these data.
4. The percentage of data points beyond the two standard deviation limits of the CURE plot for fitted values increases as the sample size increases. This may indicate that at small sample sizes the percentage outside these limits may be small simply due to the small sample.

6.2.2 Urban Two-Lane Undivided Segments

The model for total crashes and average conditions developed using data from Ohio was calibrated to randomly selected sites from Minnesota.

Table 6-2 shows the results of the investigation for urban two-lane undivided segments. From them, the following observations can be made:

1. With 25 sites, the directly estimated model and calibration function models did not converge. Surprisingly, however, the calibration factor of 0.14 had a lower CV than the 50-site sample and would be considered acceptable per the Calibrator tool guidance. All of the goodness-of-fit statistics look impressive at first glance, but this is deceptive, as the sample size is only 59 crashes. A modified R^2 of 0.96, for example, is unrealistically high.
2. With 50 sites, although the CV is greater than the 0.15 threshold, the calibration function measures are slightly better than those for the directly estimated model, except for the CURE plot measure of data points outside the two standard deviation limits for the predicted values, for which there is a tie.
3. With 75 sites, the calibration factor and function perform better than the directly estimated model, with the exception of the overdispersion parameter measure. The goodness-of-fit statistics are worse than for the 50-site sample, and the CV of 0.16 is just over the 0.15 threshold.
4. The results for 100 sites are similar to those for 75 sites.
5. As was seen for urban four-lane divided segments, the percentage of data points outside the two standard deviation limits of the CURE plot for the fitted values increases as the sample size increases.

6.2.3 Rural Two-Lane, Three-Leg Stop-Controlled Intersections

We calibrated the model for total crashes and base conditions developed using data from Minnesota (a total of seven years of crash data) to randomly selected sites from Ohio (a total of five years of crash data).

Table 6-3 shows the results of the investigation for rural two-lane, three-leg stop-controlled intersections. A number of observations can be made from these results:

1. For all four sample sizes, the goodness-of-fit statistics are reasonably similar.
2. The maximum calibration factor CV value of 0.15 recommended in the Calibrator tool guidance is not reached until a sample size of 125 sites and 247 crashes is reached. Most interesting is that a model directly estimated for the Ohio data was successful at smaller sample sizes, and the goodness-of-fit statistics for all three options were comparable. This would seem to indicate that even if the CV is higher than the 0.15 threshold, a directly estimated model may still be feasible even with smaller sample sizes.
3. The calibration function does perform better in general than a calibration factor, although the differences are not very large for these data.
4. The percentage of data points beyond the two standard deviation limits of the CURE plot for fitted values increases as the sample size increases for the calibration factor option (Option A). This may

indicate that at small sample sizes the percentage outside these limits may be small simply due to the small sample.

6.3 Approach 2 Results

The investigation for this approach and site type involved an assessment of the temporal and spatial transferability and calibration of the models based on the CV of the calibration factor. In this case, we used all of the data available for the calibration rather than samples of various sizes.

First, we applied Texas 2012 data for calibration of the SPFs, using Texas 2009–11 data for undivided highway segments. Table 6-4 shows the results. Then, we used Ohio 2009–11, Washington 2009–11, and Illinois 2009–10 data for calibration of the California SPFs for divided highway segments. The results are shown in Table 6.5.

The results in Table 6-5 indicate that, for Ohio and Illinois, the calibration function would provide predictions similar to those provided by a single calibration factor, since parameter b (Equation 6.1) was close to 1.0. No insights could be obtained on sample sizes of sites and crashes, as the results were not only inconsistent but very jurisdiction-specific. The lowest MAD value, for example, was for the data with the largest number of sites but the fewest crashes and the highest value of the CV of the calibration factor.

The temporal calibration results in Table 6-4 show parameter b of the calibration function was also close to 1.0, but even with a relatively large sample of sites and crashes for the same state, the CV of the calibration factor was beyond the threshold of 0.15 recommended for a successful calibration.

Table 6-1: Results for Urban Four-Lane Divided Segments

No. Sites	Observed Crashes	C (CV)	Calibration Function Parameters		MAD Option*			Modified R^2 Option			overdispersion parameter Option			CURE max dev Option			CURE % dev Option		
			a (s.e.)	b (s.e.)	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
50	140	1.48 (0.24)	0.4296 (0.2277)	0.9859 (0.2735)	2.28	2.28	2.10	0.16	0.16	0.26	0.97	0.97	0.96	19.02	18.61	20.57	4	4	2
75	161	1.11 (0.17)	0.0222 (0.1811)	1.1183 (0.1974)	1.63	1.66	1.66	0.32	0.29	0.29	0.56	0.57	0.56	9.99	14.38	18.75	17	7	5
100	271	1.25 (0.15)	0.4178 (0.1324)	0.8995 (0.1415)	2.22	2.13	2.20	0.00	0.12	0.00	0.66	0.63	0.64	52.83	41.28	50.13	21	15	20

*A) Estimating a single calibration factor (C); B) Estimating a calibration function; C) Directly estimating a model using the calibration data

Table 6-2: Results for Urban Two-Lane Undivided Segments

No. Sites	Observed Crashes	C (CV)	Calibration Function Parameters		MAD Option*			Modified R^2 Option			overdispersion parameter Option			CURE max dev Option			CURE % dev Option		
			a (s.e.)	b (s.e.)	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
25	59	1.39 (0.14)	n/a	n/a	1.20	n/a	n/a	0.96	n/a	n/a	0.02	n/a	n/a	3.95	n/a	n/a	4	n/a	n/a
50	137	1.15 (0.20)	0.3374 (0.2004)	0.8394 (0.1949)	1.96	1.92	1.97	0.43	0.48	0.43	0.60	0.57	0.60	21.13	13.68	21.25	4	2	2
75	186	1.15 (0.16)	0.3786 (0.1641)	0.7642 (0.1740)	1.82	1.74	1.85	0.35	0.44	0.14	0.68	0.63	0.59	17.76	11.57	31.24	9	3	31
100	232	1.15 (0.16)	0.1846 (0.1513)	0.9945 (0.1749)	1.69	1.69	1.73	0.31	0.31	0.18	0.69	0.69	0.62	22.86	22.44	38.74	16	16	26

*A) Estimating a single calibration factor (C); B) Estimating a calibration function; C) Directly estimating a model using the calibration data

Table 6-3: Results for Rural Two-Lane Three-Leg Stop-controlled Intersections

No. Sites	Observed Crashes	C (CV)	Calibration Function Parameters		MAD			Modified R^2			overdispersion parameter			CURE max dev			CURE % dev		
					Option*			Option			Option			Option			Option		
			a (s.e.)	b (s.e.)	A	B	C	A	B	C	A	B	C	A	B	C	A	B	C
50	97	1.31 (0.23)	1.7636 (0.1637)	0.4575 (0.1848)	1.71	1.63	1.64	0.00	0.19	0.18	0.94	0.71	0.72	19.93	4.73	4.71	46	1	2
75	145	1.22 (0.21)	1.5785 (0.1452)	0.5806 (0.16)	1.74	1.68	1.69	0.22	0.22	0.24	0.88	0.79	0.79	21.45	41.28	16.53	46	1	1
100	194	1.23 (0.18)	1.63 (0.123)	0.5482 (0.1413)	1.77	1.67	1.67	0.11	0.19	0.26	0.91	0.77	0.72	37.17	15.33	13.63	49	1	2
125	247	1.29 (0.15)	1.6714 (0.1051)	0.5797 (0.1249)	1.74	1.65	1.65	0.09	0.22	0.24	0.80	0.68	0.66	44.67	16.09	17.74	58	1	1

*A) Estimating a single calibration factor (C); B) Estimating a calibration function; C) Directly estimating a model using the calibration data

Table 6-4: Calibration results using Texas 2012 data for calibration of SPFs using Texas 2009-2011 data for undivided highway segments

Data	Crash Type	Observed Crashes	HSM Pred.	MAD	Calibration Factor (C) (HSM, 2010)			Calibration Function $N_{predicted} = a \times (\text{Unadjusted Prediction})^b$			
					C (CV)	N Fitted	MAD	a (SE)	b (SE)	N Fitted	MAD
TX 2012 (n=402)	Total KABCO	195	233.28	0.583	0.836 (0.211)	195	0.542	0.825 (0.089)	0.838 (0.084)	188.964	0.554

Table 6-5: Calibration results using Ohio, Washington, and Illinois data for calibration of California SPFs for divided highway segments.

Data	Crash Type	Observed Crashes	HSM Pred.	MAD	Calibration Factor (C) (HSM, 2010)			Calibration Function $N_{predicted} = a \times (\text{Unadjusted Prediction})^b$			
					C (CV)	N Fitted	MAD	a (SE)	b (SE)	N Fitted	MAD
OH (n=407)	Total KABCO	856	866.12	1.348	0.988 (0.100)	856	1.345	0.991 (0.066)	1.003 (0.058)	861.064	1.346
WA (n=216)	Total KABCO	730	441.15	2.007	1.655 (0.144)	730	1.978	1.969 (0.141)	0.848 (0.065)	733.060	1.939
IL (n=592)	Total KABCO	170	233.75	0.463	0.727 (0.210)	170	0.413	0.747 (0.102)	1.046 (0.131)	169.461	0.411

6.4 Conclusions on Calibration Exercise

6.4.1 Summary of Findings

The results of the analyses indicate no consistency with regard to which option (calibration factor, calibration function, or directly estimated model) will perform best for a given sample size. For some cases, a small sample that is estimated using some criterion (for example, maximum CV of the calibration factor) may work; for others, it may not. What sample size will work is also highly variable, and dependent on factors including the average crash rate and amount of variation of site characteristics in the data.

It is concluded that, at present, the required sample size for any of the calibration options can only be determined by trial and error, and the current HSM sample size guidance and subsequently developed resources (Bahar et al. 2014) can provide reasonable practical limits for the amount of data that may practically be collected for the start of a calibration exercise.

Other key calibration issues were investigated but could not be resolved in this research. They included the following:

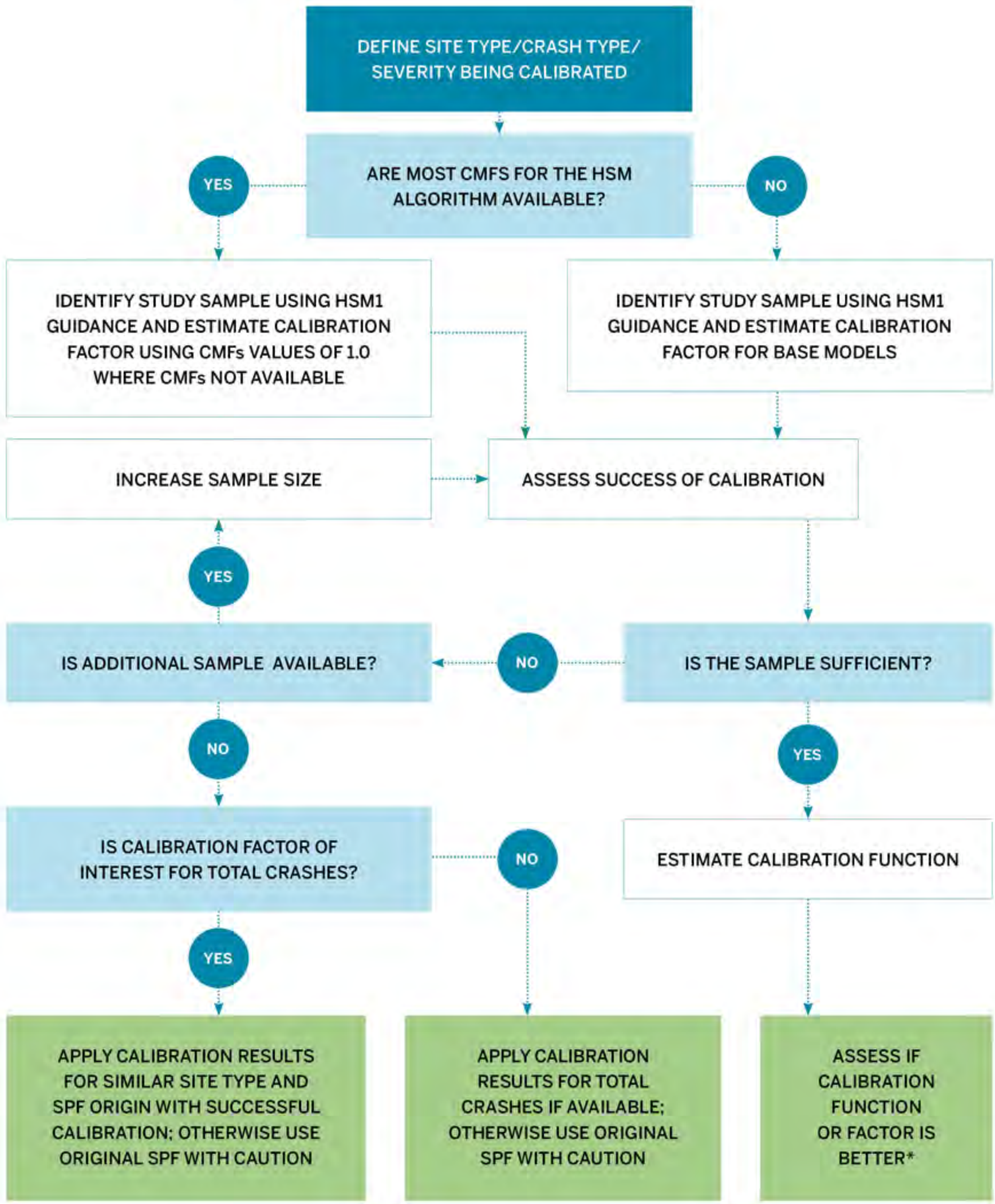
1) *Should the calibration factor be estimated for the base models rather than for the HSM algorithm as a whole (that is, applying CMFs to the base models), as is the case at the moment?* The recommendation is to maintain the status quo for site types, crash types, and crash severities for which there are enough CMFs to apply the algorithm, and to conduct further research on this topic. For situations in which there are few or no CMFs, the recommendation is to estimate the calibration factor from the base models.
2) *Should the overdispersion parameter be calibrated?* The current HSM methodology does not suggest this. It is recommended that future research consider basing this decision on an estimate of the standard deviation of the calibrated overdispersion parameter. Future research will also need to consider how the overdispersion parameter should be calibrated for a calibration function.

6.4.2 Recommended Calibration Procedure Update

On the basis of these conclusions, the following is recommended as an updated calibration procedure as depicted in Figure 6-1.

1. For site types, crash types, and crash severities for which there are enough CMFs to apply the HSM algorithm, perform the calibration for the algorithm as a whole (that is, by applying CMFs to the base models). For other situations, perform the calibration for the base models.
2. Start with an available sample that is desirably random and at least as large as that recommended in the HSM.
3. Perform the calibration first with a constant calibration factor. The FHWA Calibrator tool can be used.
4. Assess the success of the calibration. The user guide for the FHWA Calibrator tool provides guidance on how success can be assessed with CURE plots and the CV of the calibration factor. The latter measure is estimated and assessed in the Calibrator tool based on guidance provided in Bahar et al. (2014), Appendix B. That guidance can be used instead of the tool.
5. If the sample is insufficient, then incrementally assemble additional data for additional sites and assess until a successful calibration is achieved.

    - If a successful calibration cannot be achieved with the entire sample available for *total* crashes, then the calibration results for a similar site type (from which a successful calibration was achieved) may be assumed to apply.
    - If a successful calibration cannot be achieved with the entire sample available for a *specific crash type or severity*, then the calibration results for total crashes, however obtained, may be assumed to apply.
6. Estimate a calibration function using the approach in Srinivasan et al. (2016), and adopt it in preference to the calibration factor, if it is successfully estimated and performs better.
7. If appropriate skills are available or could be acquired, it is recommended to try to estimate directly a model with the final calibration dataset and adopt it if it is successfully estimated and performs better than the calibration factor and calibration function. The FHWA Calibrator tool can be used in this performance assessment.



*If calibration is being done for base models and appropriate skills are available or could be acquired, it is recommended to try to directly estimate a model with the final calibration dataset and adopt the model if successfully estimated and performs better than the calibration factor and calibration function.

Figure 6-1: Suggested Calibration Process

7 FINDINGS AND CONCLUSIONS

7.1 PROPOSED MODELS AND PROCEDURES FOR MANUAL

This report presents SPFs that were estimated to predict crashes by type and severity for the facility types covered by the HSM. To optimize the accuracy of the crash predictions, it would have been ideal to estimate SPFs for all of the crash types and severities need as base conditions for applying CMFs. Unfortunately, because the numbers of crashes of various types and severities were limited in the databases available for the project, we could not estimate models for all the specific types—for example, for head-on and rear-end crashes within the opposite- and same-direction crash type categories, respectively. Some models also have overdispersion parameters high enough to cast doubt on their accuracy for prediction or coefficients on the volume predictors (AADT) that are not statistically significant at 90 percent confidence. We address these cases by reporting average proportions of specific crash types within the broader crash type category for which models were estimated. These proportions can be used where the predicted models are not available, or where the analyst chooses not to use them. This approach will still be more accurate than that provided in the current HSM.

While the initial scope of work had proposed to estimate probabilistic models for crash severity, based on theoretical and practical considerations about the application of such models in the HSM procedures, we chose not to use this approach. Predictions of crash severity may be calculated using the count models for severity that have been estimated and presented here.

The data sources for estimation and validation for each model by facility are listed in Table 7-1. The rest of this section identifies the models that were estimated and will be proposed for inclusion in the HSM, by facility type. It also lists the crash types and severity for which models were NOT estimated, for each facility type that might require proportions to be estimated. For these situations, we provide default proportions from the data for each facility type, although individual jurisdictions may calculate proportions from their own data for more accurate local predictions. Finally, the section summarizes findings from revisiting the calibration procedures in the HSM in light of the newly estimated models.

Table 7-1: Data Sources (States) for SPF Estimation and Validation, by Facility Type

Facility Type	Road Segments Estimation	Road Segments Validation	Intersections Estimation	Intersections Validation
Two-lane rural highways	Washington	Ohio	3ST&4ST: Minnesota 4SG: Ohio	3ST&4ST: Ohio 4SG: Minnesota
Multilane rural highways	4U: Texas* 4D: California	4U: Texas* 4D: Illinois, Washington	3ST&4ST: Minnesota 4SG: Ohio	3ST&4ST: Ohio 4SG: Minnesota
Urban/suburban arterials	Ohio	Minnesota	Ohio	North Carolina

*Data from 2009–11 were used for estimation and from 2012 for validation.

7.1.1 Two-Lane Rural Highway Models

Figure 7-1 identifies the crash models estimated for segments on two-lane rural highways. SPFs were estimated for four levels of severity (KABC0, KABC, KAB, and KA) for total, same-direction, opposite-direction, and single-vehicle crashes. Models for intersecting-direction crashes were not estimated, as they were all assigned to intersections.

Figure 7-2 lists the specific crash types included in each of the broader crash type categories that were estimated. Total crashes include all of these crash types. If predictions of crashes of any of these specific types are needed for applying CMFs, proportions of them within the aggregate crash types (those in the left column of the figure) will be required and provided in the proposed HSM content.

Figure 7-3 illustrates the crash models estimated for intersections (all types) on two-lane rural highways. Models were estimated for intersecting-direction crashes as well as for the types estimated for segments and for the same four levels of severity.

Similarly, Figure 7-4 lists the specific crash types included in each of the broader crash type categories that were estimated for two-lane rural highway intersections. Again, total crashes include all of these crash types, and their proportions within the aggregate crash types will be required for estimating them.

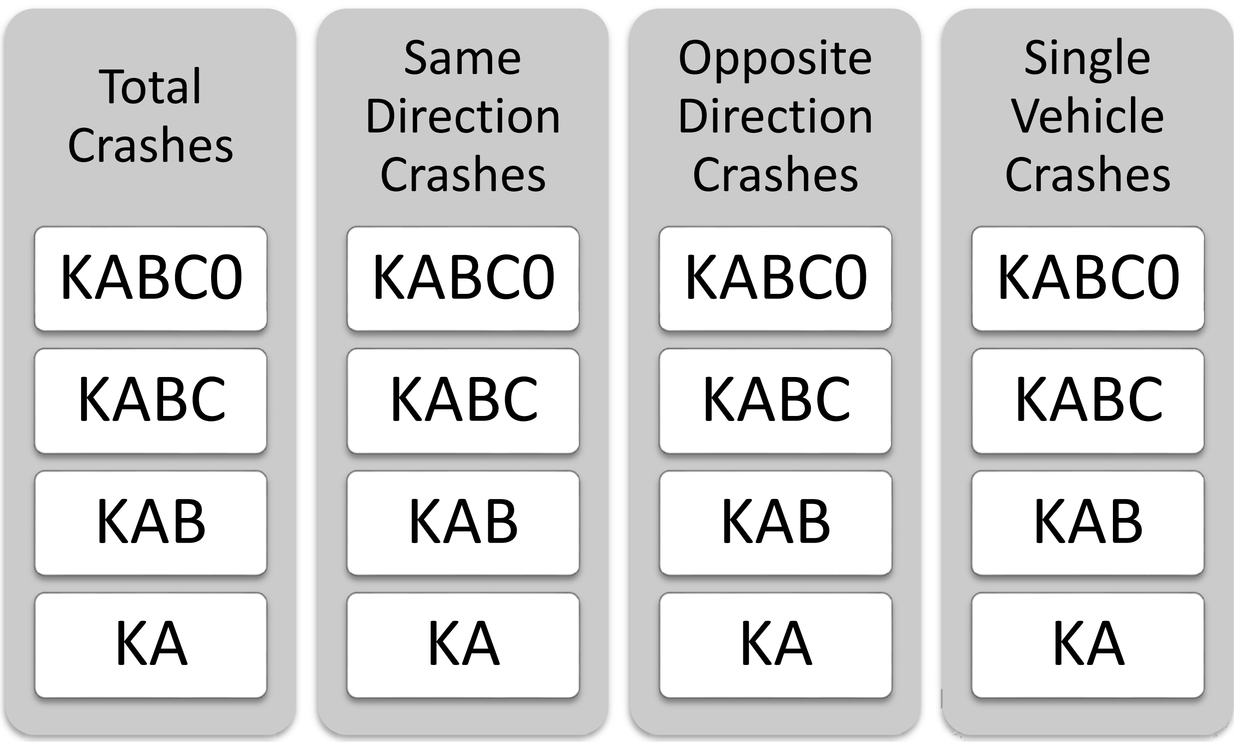


Figure 7-1: Crash Types Estimated for Segments on Two-Lane Rural Highways

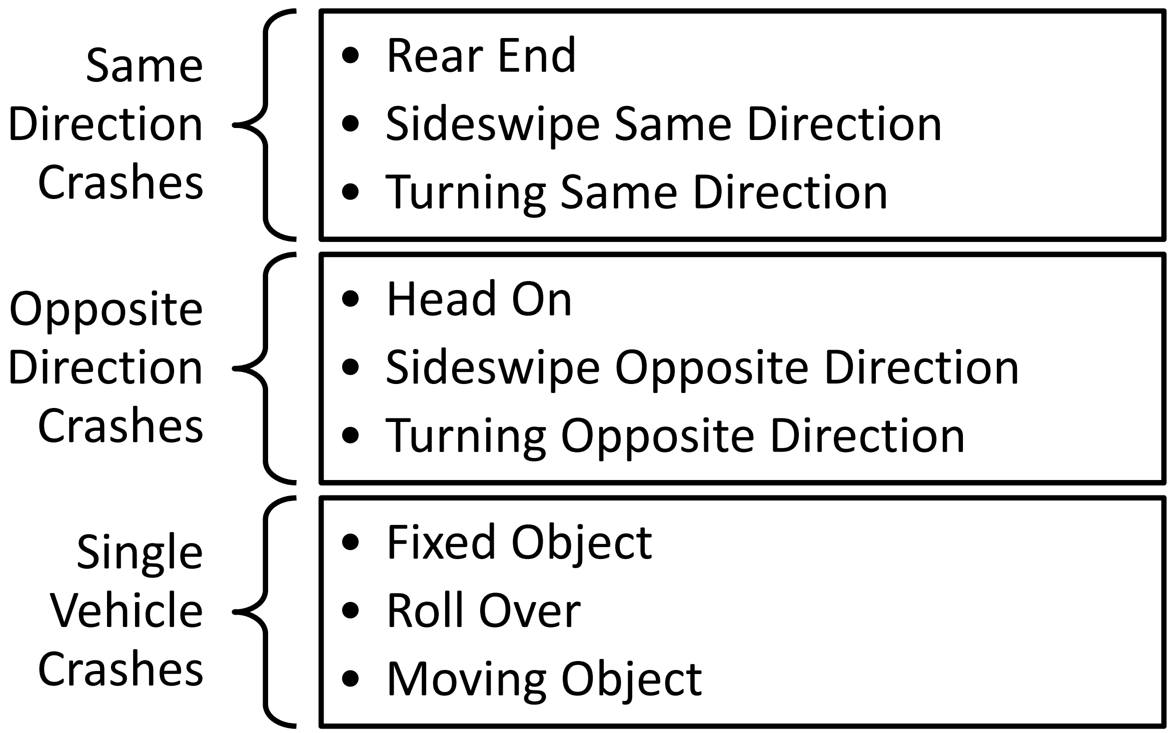


Figure 7-2: Specific Crash Types Included in the Estimated Crash Types (Rural 2U).

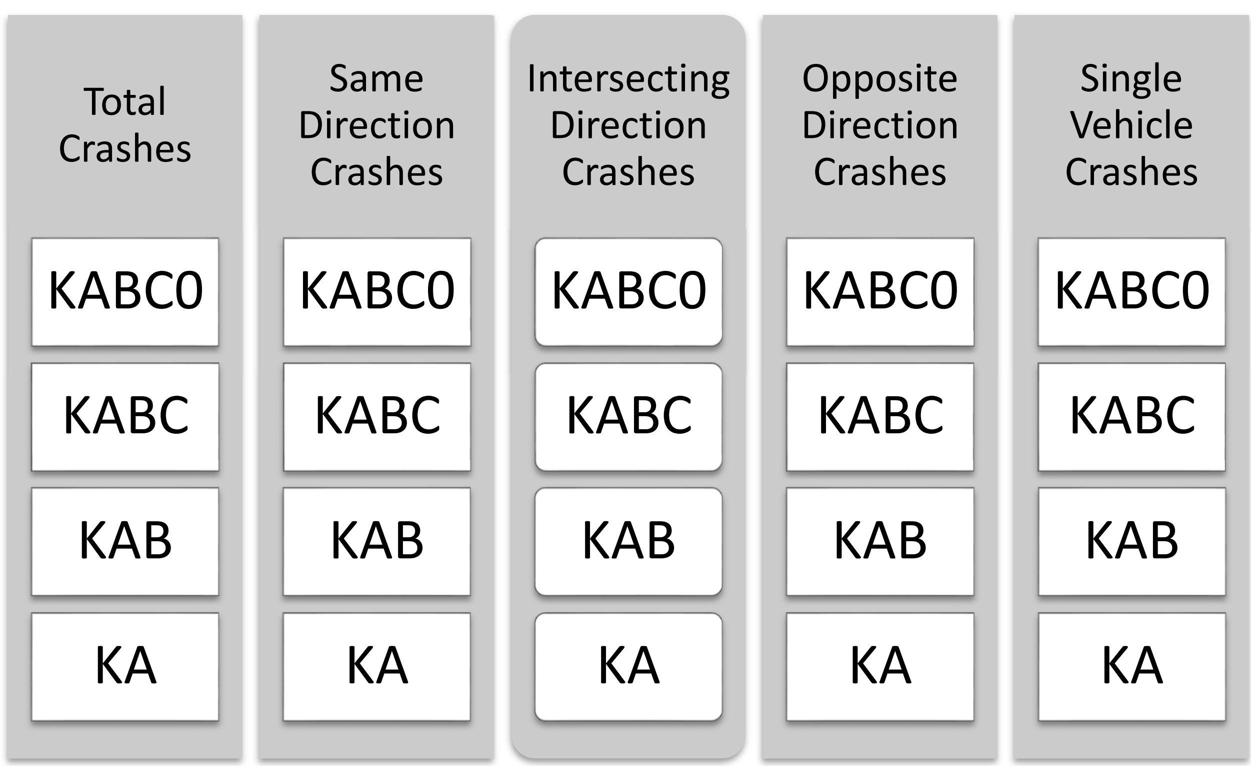


Figure 7-3: Crash Types Estimated for Intersections on Two-Lane Rural Highways.

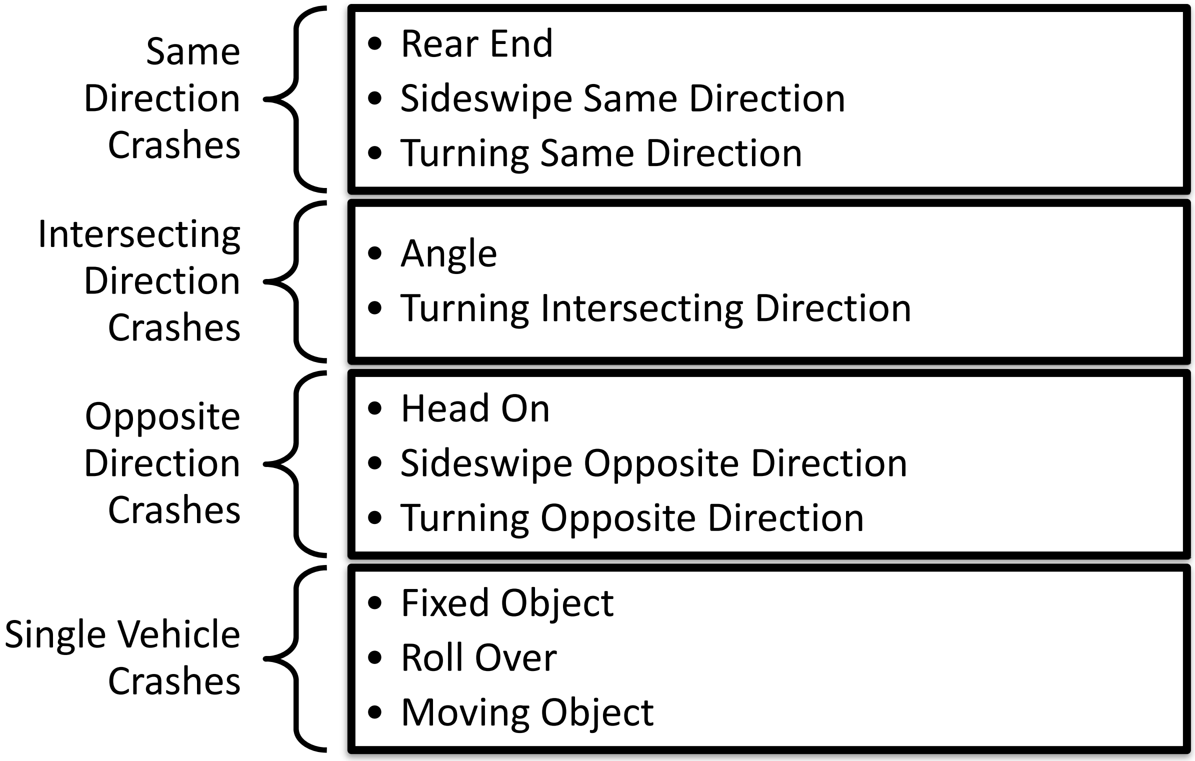


Figure 7-4: Specific Crash Types included in the Estimated Crash Types (Rural 3ST, 4ST and 4SG)

7.1.2 Multilane Rural Highway Models

Figure 7-5 identifies the crash models estimated for divided and undivided segments on multilane rural highways. Models were estimated for the same combinations of type and severity as for two-lane rural highways, with two exceptions. First, due to the small number of same-direction KA crashes, we could not estimate a model for that combination. Second, we attempted models for intersecting-direction crashes for all severity levels, but due to the small number of crashes, only the model for all severity levels was successfully estimated. Figure 7-4 lists the specific crash types included in each of these aggregated types (same as for two-lane rural highway intersection models). Again, the total crash category includes all of the crash types and proportions that will need to be computed for any specific crash types of interest within the broader categories for which models were estimated.

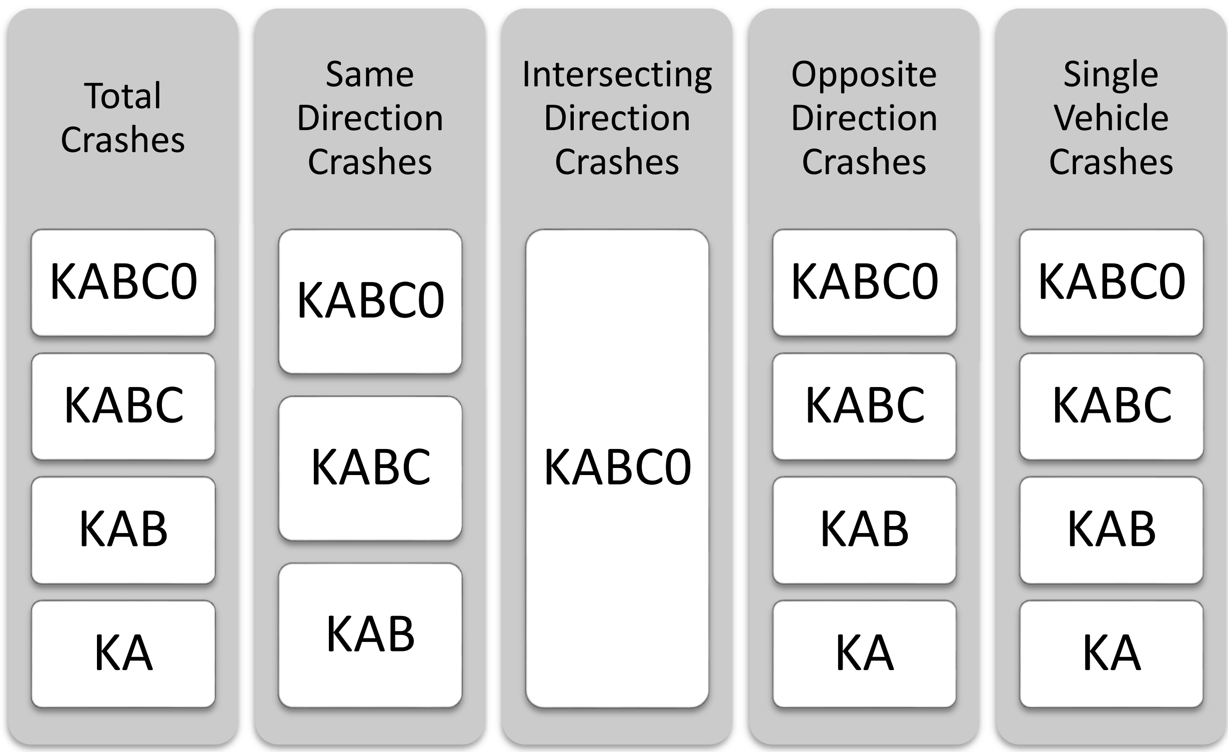


Figure 7-5: Crash Types Estimated for Divided and Undivided Segments on Multilane Rural Highways

Intersection models cover the same crash types as for rural two-lane highway intersections, as depicted in Figure 7-3 and Figure 7-4.

7.1.3 Urban/Suburban Arterial Models

Final base condition models were estimated for urban/suburban arterial segments for the following crash types:

- Total
- KABC
- KAB
- KA
- Multiple-vehicle non-driveway related
- Rear end
- Sideswipe same direction
- Head-on + sideswipe opposite direction
- Multiple-vehicle non-driveway other
- Single vehicle
- Nighttime
- Multiple-vehicle driveway

Note that crashes were estimated by type or severity level, not in combination, as was done for the rural facility types. The reason for this is that, for many combinations of crash type and severity, there simply were not enough crashes to estimate viable models. The combination of crash type and severity is not frequently needed to apply HSM methods, so it is recommended that when predictions of such combinations are needed, proportions may be calculated to allocate crash type predictions among the various severity levels.

For urban/suburban intersections, we estimated models for the same combinations of crash type and severity as for the rural intersection facilities (see Figure 7-3). As noted, however, models for many of the combinations could not be estimated due to small sample sizes or odd estimation results. Figure 7-6 depicts the combinations that could not be estimated for each type of intersection.

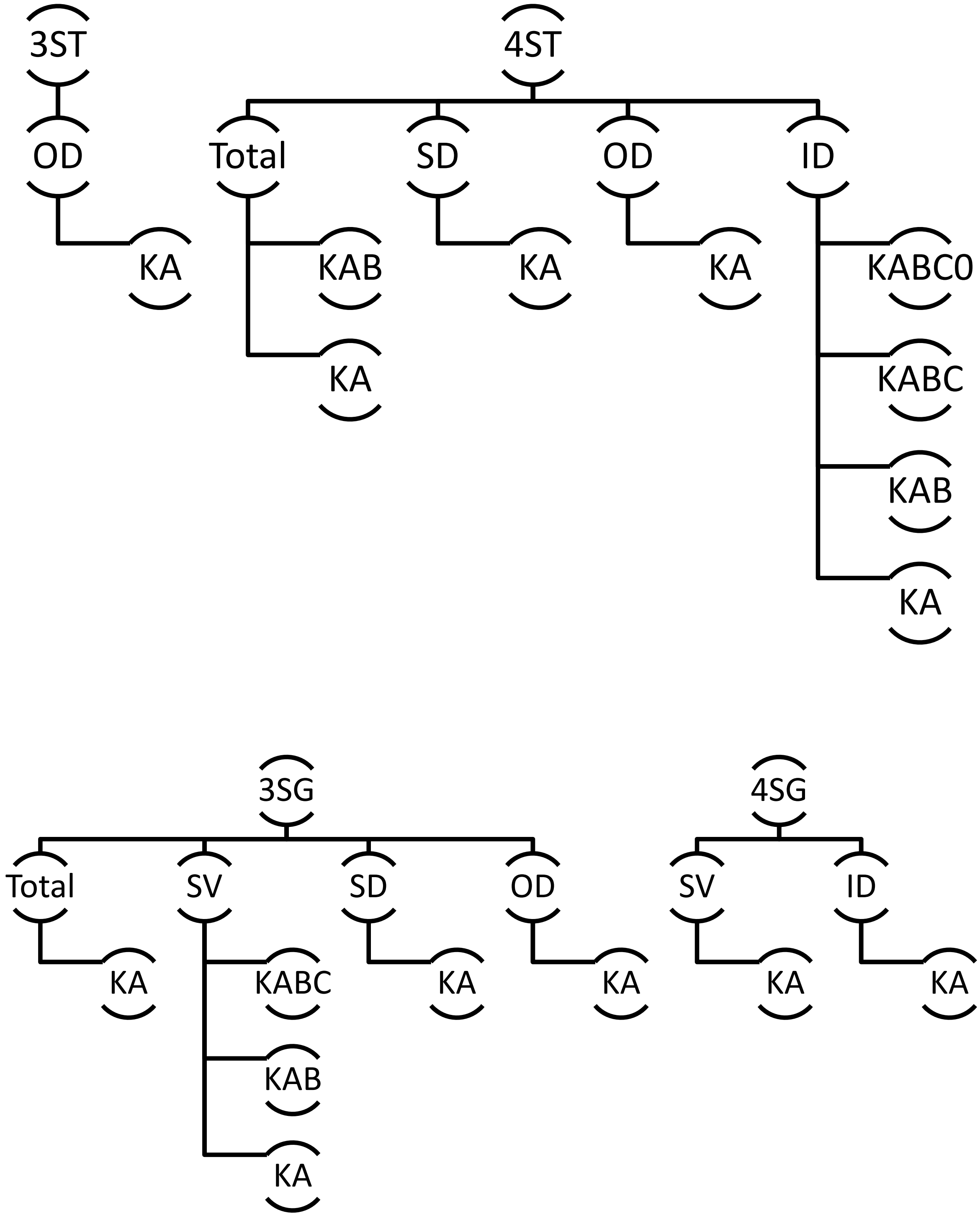


Figure 7-6: Crash Type and Severity SPFs that Were Not Estimated for Urban/Suburban Intersections

7.1.4 Revisit of the Calibration Procedure

The project team revisited the current HSM calibration procedure and evaluated its performance relative to sample size and to using a constant or variable calibration factor (or calibration function). In addition, we considered the issues of calibrating models based on crash predictions with and without CMFs and calibration of the overdispersion parameter. The recommendation is to continue the current procedure in the HSM of calibrating with the CMFs, assuming most of the CMFs for doing so are available. Otherwise, as may be the case for many crash type and severity models at the moment, the calibration may be done without applying CMFs. The issue of calibrating the overdispersion parameter requires further research.

The findings show that the calibration results are definitely sensitive to sample size, but not always in ways that might be expected. The calibration function did not work well with small sample sizes because the optimization procedure to estimate it failed to converge. In general, the more complex the calibration approach, the more data are required to apply it successfully. Unsurprisingly, the sample sizes that resulted in the best calibration results would also be large enough to estimate jurisdiction-specific SPFs. This latter option would be preferred, when possible, to get the most accurate predictions for application in a given jurisdiction. But the findings here show that, in many cases, reasonable predictions are also possible following the HSM procedures, with even a constant calibration factor and modest calibration sample sizes.

7.2 Conclusions

In conclusion, this project has estimated new prediction models for crash types and severity that promise better predictive results than the current HSM-recommended combination of base models for total crashes with proportional factors for allocating among crash type and severity. When sample size permitted, extensive SPFs developed by severity and type are provided for detailed analytics of safety rather than fixed proportions as the current HSM provides. They are estimated with much newer crash data than the models in the HSM, which were estimated using data 10 to 15 years old. These updated models, including ones for total crashes, reflect more current relationships between traffic exposure and crash occurrence, as well as differences in the shape of the SPF viz. the traffic exposure from one crash type or severity level to another. The project has also revisited the predictive method calibration procedure in the HSM and offers refinements to the recommended calibration procedure.

It is noted that estimation and application of crash prediction models is dependent upon having datasets of sufficient size and quality. It was not possible to estimate models for K only crashes for any crash types or in total for any facility type due to the small number of these crashes in any of the data sets. For some crash types, such as same direction crashes, KA crash models also could not be estimated. Some of these crash type and severity combinations are extremely rare due to their nature (e.g., same direction KA crashes), so it is hard to identify future research directions that could overcome this challenge.

It is also noted that many of the roadway characteristic variables that are necessary for estimating and applying these models, for example numbers of driveways of different types and intersection skew angles, are not routinely archived by all transportation agencies. For estimation and validation of these models it was necessary to engage in data collection efforts to augment data provided by the transportation agencies that were used in the project. In order to use these prediction procedures, most agencies will likely need to augment their own data archives with additional roadway characteristics.

This project also provides, in Appendix C, content for incorporating the new estimated models and calibration recommendations into a second edition of the HSM.

8 References

Bahar G., E. Hauer. 2014. User's Guide to Develop Highway Safety Manual Safety Performance Function Calibration Factors. Final Report, NCHRP Project 20-07, Task 332.

Cafiso S., G. Di Silvestro, B. Persaud, and M. Begum. 2010. Revisiting Variability of Dispersion Parameter of Safety Performance for Two-Lane Rural Roads. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2148, Transportation Research Board of the National Academies, Washington, DC, pp. 38-46.

Revisiting Variability of Dispersion Parameter of Safety Performance for Two-Lane Rural Roads

Chang, L. and F. Mannering. 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. *Accident Analysis and Prevention* 31 No. 5, pp. 579-592.

Dziak, J.J., Coffman, D.L., Lanza, S.T., and Li, R. 2012. *Sensitivity and Specificity of Information Criteria*, The Methodology Center, Pennsylvania State University – College of Health and Human Development, Technical Report Series # 12-119, State College, (http://methodology.psu.edu/media/techreports/12-119.pdf).

Farid, A., M. Abdel-Aty, J. Lee, N. Eluru and J.-H. Wang. Exploring the transferability of safety performance functions. *Accident Analysis and Prevention* 94, 2016, pp. 143-152.

Golob T., W. Recker and J. Leonard. 1987. An Analysis of the Severity and Incident Duration of Truck-Involved Freeway Accidents. *Accident Analysis and Prevention* 19, No. 5, pp. 375-395.

Hauer, E. (2015 – in press), *The Art of Regression Modeling in Road Safety*, Springer.

Institute for Digital Research and Education, (IDE-UCLA). Accessed December 18, 2014 at <http://www.ats.ucla.edu/stat/sas/dae/negbinreg.htm>.

Ivan, J. 2004. New Approach for Including Traffic Volumes in Crash Rate Analysis and Forecasting," *Transportation Research Record: Journal of the Transportation Research Board*, No. 1897, Transportation Research Board of the National Academies, Washington, DC, pp. 134-141.

Kockelman, K.M. and Y.J. Kweon. 2002. Driver injury severity: an application of ordered probit models. *Accident Analysis and Prevention* 34 (3), pp. 313-321.

Lyon C., Persaud B., and F. Gross. 2016. The Calibrator—Calibrate, Critique, CURE. An SPF Calibration Tool. FHWA DTFH61-10-D-00022-T-13005.

Papke, L. E., and Wooldridge, J. 1993. Econometric methods for fractional response variables with an application to 401 (k) plan participation rates.

SAS/STAT(R) 9.2 User's Guide, Second Edition. Accessed December 18, 2014 at <http://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#statug_genmod_sect030.htm#statug.genmod.genmodoverdisp>.

Srinivasan, R., M. Colety, G. Bahar, B. Crowther, and M. Farmen. 2016. Estimation of Calibration Functions for Predicting Crashes on Rural Two-Lane Roads in Arizona. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2583, Transportation Research Board, Washington, DC , pp. 17-24.

Wang, J.-H., M. Abdel-Aty, and J. Lee. 2017. Evaluation of Safety Performance Functions Based on Experimental Design Using Cross-Validation Method. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2636, Transportation Research Board, Washington, DC, 2017.

Wang, J.-H., M. Abdel-Aty, and J. Lee. 2016 Examination of the Transferability of Safety Performance Functions for Developing Crash Modification Factors: Using the Empirical Bayes Method. *Transportation Research Record: Journal of the Transportation Research Board Annual Meeting*, No. 2583, Transportation Research Board, Washington, DC, pp. 73-80.

Washington, S., Persaud, B., Lyon, C. and Oh, J. 2005. *Validation of Accident Models for Intersections.* Federal Highway Administration, United States Department of Transportation, Report No. FHWA-RD-03-037.

Yasmin, S., N. Eluru, J. Lee, and M. Abdel-Aty. 2016. Ordered Fractional Split Approach for Aggregate Injury Severity Modeling. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2583, Transportation Research Board, Washington, DC, pp. 119-126.

Zhang, C., J. Ivan and T. Jonsson. 2007. Collision Type Categorization Based on Crash Causality and Severity Analysis. Transportation Research Board 86th Annual Meeting Compendium of Papers, Paper No. 07-2454, Washington, DC.