Appendix A—Specialized Procedures Common to All Part C Chapters

This Appendix presents two specialized procedures intended for use with the predictive method presented in Chapters 10, 11, and 12. These include the procedure for calibrating the predictive models presented in the Part C chapters to local conditions and the Empirical Bayes (EB) Method for combining observed crash frequencies with the estimate provided by the predictive models in Part C. Both of these procedures are an integral part of the predictive method in Chapters 10, 11, and 12, and are presented in this Appendix only to avoid repetition across the chapters.

A.1. CALIBRATION OF THE PART C PREDICTIVE MODELS

The Part C predictive method in Chapters 10, 11, and 12 include predictive models which consist of safety performance functions (SPFs), crash modification factors (CMFs) and calibration factors and have been developed for specific roadway segment and intersection types. The SPF functions are the basis of the predictive models and were developed in HSM-related research from the most complete and consistent available data sets. However, the general level of crash frequencies may vary substantially from one jurisdiction to another for a variety of reasons including climate, driver populations, animal populations, crash reporting thresholds, and crash reporting system procedures. Therefore, for the Part C predictive models to provide results that are meaningful and accurate for each jurisdiction, it is important that the SPFs be calibrated for application in each jurisdiction. A procedure for determining the calibration factors for the Part C predictive models is presented below in Appendix A.1.1.

Some HSM users may prefer to develop SPFs with data from their own jurisdiction for use in the Part C predictive models rather than calibrating the Part C SPFs. Calibration of the Part C SPFs will provide satisfactory results. However, SPFs developed directly with data for a specific jurisdiction may provide more reliable estimates for that jurisdiction than calibration of Part C SPFs. Therefore, jurisdictions that have the capability, and wish to develop their own models, are encouraged to do so. Guidance on development of jurisdiction-specific SPFs that are suitable for use in the Part C predictive method is presented in Appendix A.1.2.

Most of the regression coefficients and distribution values used in the Part C predictive models in Chapters 10, 11, and 12 have been determined through research and, therefore, modification by users is not recommended. However, a few specific quantities, such as the distribution of crashes by collision type or the proportion of crashes occurring during nighttime conditions, are known to vary substantially from jurisdiction to jurisdiction. Where appropriate local data are available, users are encouraged to replace these default values with locally derived values. The values in the predictive models that may be updated by users to fit local conditions are explicitly identified in Chapters 10, 11, and 12. Unless explicitly identified, values in the predictive models should not be modified by the user. A procedure for deriving jurisdiction-specific values to replace these selected parameters is presented below in Appendix A.1.3.

A.1.1. Calibration of Predictive Models

The purpose of the Part C calibration procedure is to adjust the predictive models which were developed with data from one jurisdiction for application in another jurisdiction. Calibration provides a method to account for differences between jurisdictions in factors such as climate, driver populations, animal populations, crash reporting thresholds, and crash reporting system procedures.

The calibration procedure is used to derive the values of the calibration factors for roadway segments and for intersections that are used in the Part C predictive models. The calibration factor for roadway segments, Cr, is used in Equations 10-2, 11-2, 11-3, and 12-2. The calibration factor for intersections, C_i, is used in Equations 10-3, 11-4, and 12-5. The calibration factors, C_r and C_i, are based on the ratio of the total observed crash frequencies for a selected set of sites to the total expected average crash frequency estimated for the same sites, during the same time period, using the applicable Part C predictive method. Thus, the nominal value of the calibration factor, when the observed and predicted crash frequencies happen to be equal, is 1.00. When there are more crashes observed than are predicted by the Part C predictive method, the computed calibration factor will be greater than 1.00. When there are fewer crashes observed than are predicted by the Part C predictive method, the computed calibration factor will be less than 1.00.

It is recommended that new values of the calibration factors be derived at least every two to three years, and some HSM users may prefer to develop calibration factors on an annual basis. The calibration factor for the most recent available period is to be used for all assessment of proposed future projects. If available, calibration factors for the specific time periods included in the evaluation periods before and after a project or treatment implementation are to be used in effectiveness evaluations that use the procedures presented in Chapter 9.

If the procedures in Appendix A.1.3 are used to calibrate any default values in the Part C predictive models to local conditions, the locally-calibrated values should be used in the calibration process described below.

The calibration procedure involves five steps:

- *Step 1*—Identify facility types for which the applicable Part C predictive model is to be calibrated.
- *Step 2*—Select sites for calibration of the predictive model for each facility type.
- *Step 3*—Obtain data for each facility type applicable to a specific calibration period.
- *Step 4*—Apply the applicable Part C predictive model to predict total crash frequency for each site during the calibration period as a whole.
- *Step 5*—Compute calibration factors for use in Part C predictive model.

Each of these steps is described below.

A.1.1.1. Step 1—Identify Facility Types for Which the Applicable Part C SPFs are to be Calibrated.
Calibration is performed separately for each facility type addressed in each Part C chapter. Table A-1 identifies all of the facility types included in the Part C chapters for which calibration factors need to be derived. The Part C SPFs for each of these facility types are to be calibrated before use, but HSM users may choose not to calibrate the SPFs for particular facility types if they do not plan to apply the Part C SPFs for those facility types.

Table A-1. SPFs in the Part C Predictive Models that Need Calibration

	Calibration Factor to be Derived	
Facility, Segment, or Intersection Type	**Symbol**	**Equation Number(s)**
ROADWAY SEGMENTS		
Rural Two-Lane, Two-Way Roads		
Two-lane undivided segments	C_r	10-2
Rural Multilane Highways		
Undivided segments	C_r	11-2
Divided segments	C_r	11-3
Urban and Suburban Arterials		
Two-lane undivided segments	C_r	12-2
Three-lane segments with center two-way left-turn lane	C_r	12-2
Four-lane undivided segments	C_r	12-2
Four-lane divided segments	C_r	12-2
Five-lane segments with center two-way left-turn lane	C_r	12-2
INTERSECTIONS		
Rural Two-Lane, Two-Way Roads		
Three-leg intersections with minor-road stop control	C_i	10-3
Four-leg intersections with minor-road stop control	C_i	10-3
Four-leg signalized intersections	C_i	10-3
Rural Multilane Highways		
Three-leg intersections with minor-road stop control	C_i	11-4
Four-leg intersections with minor-road stop control	C_i	11-4
Four-leg signalized intersections	C_i	11-4
Urban and Suburban Arterials		
Three-leg intersections with minor-road stop control	C_i	12-5
Three-leg signalized intersections	C_i	12-5
Four-leg intersections with minor-road stop control	C_i	12-5
Four-leg signalized intersections	C_i	12-5

A.1.1.2. Step 2—Select Sites for Calibration of the SPF for Each Facility Type.

For each facility type, the desirable minimum sample size for the calibration data set is 30 to 50 sites, with each site long enough to adequately represent physical and safety conditions for the facility. Calibration sites should be selected without regard to the number of crashes on individual sites; in other words, calibration sites should not be selected to intentionally limit the calibration data set to include only sites with either high or low crash frequencies. Where practical, this may be accomplished by selecting calibration sites randomly from a larger set of candidate sites. Following site selection, the entire group of calibration sites should represent a total of at least 100 crashes per year. These calibration sites will be either roadway segments or intersections, as appropriate to the facility type being addressed. If the required data discussed in Step 3 are readily available for a larger number of sites, that larger number of sites should be used for calibration. If a jurisdiction has fewer than 30 sites for a particular facility type, then it is desirable to use all of those available sites for calibration. For large jurisdictions, such as entire states, with a variety of topographical and climate conditions, it may be desirable to assemble a separate set of sites and develop separate calibration factors for each specific terrain type or geographical region. For example, a state with distinct plains and mountains regions, or with distinct dry and wet regions, might choose to develop separate calibration factors for those regions. On the other hand, a state that is relatively uniform in terrain and climate might choose to perform a single calibration for the entire state. Where separate calibration factors are developed by terrain type or region, this needs to be done consistently for all applicable facility types in those regions.

It is desirable that the calibration sites for each facility type be reasonably representative of the range of site characteristics to which the predictive model will be applied. However, no formal stratification by traffic volume or other site characteristics is needed in selecting the calibration sites, so the sites can be selected in a manner to make the data collection needed for Step 3 as efficient as practical. There is no need to develop a new data set if an existing data set with sites suitable for calibration is already available. If no existing data set is available so that a calibration data set consisting entirely of new data needs to be developed, or if some new sites need to be chosen to supplement an existing data set, it is desirable to choose the new calibration sites by random selection from among all sites of the applicable facility type.

Step 2 only needs to be performed the first time that calibration is performed for a given facility type. For calibration in subsequent years, the same sites may be used again.

A.1.1.3. Step 3—Obtain Data for Each Facility Type Applicable to a Specific Calibration Period.
Once the calibration sites have been selected, the next step is to assemble the calibration data set if a suitable data set is not already available. For each site in the calibration data set, the calibration data set should include:

- Total observed crash frequency for a period of one or more years in duration.
- All site characteristics data needed to apply the applicable Part C predictive model.

Observed crashes for all severity levels should be included in calibration. The duration of crash frequency data should correspond to the period for which the resulting calibration factor, C_r or C_i, will be applied in the Part C predictive models. Thus, if an annual calibration factor is being developed, the duration of the calibration period should include just that one year. If the resulting calibration factor will be employed for two or three years, the duration of the calibration period should include only those years. Since crash frequency is likely to change over time, calibration periods longer than three years are not recommended. All calibration periods should have durations that are multiples of 12 months to avoid seasonal effects. For ease of application, it is recommended that the calibration periods consist of one, two, or three full calendar years. It is recommended to use the same calibration period for all sites, but exceptions may be made where necessary.

The observed crash data used for calibration should include all crashes related to each roadway segment or intersection selected for the calibration data set. Crashes should be assigned to specific roadway segments or intersections based on the guidelines presented below in Appendix A.2.3.

Table A-2 identifies the site characteristics data that are needed to apply the Part C predictive models for each facility type. The table classifies each data element as either required or desirable for the calibration procedure. Data for each of the required elements are needed for calibration. If data for some required elements are not readily available, it may be possible to select sites in Step 2 for which these data are available. For example, in calibrating the predictive models for roadway segments on rural two-lane, two-way roads, if data on the radii of horizontal curves are not readily available, the calibration data set could be limited to tangent roadways. Decisions of this type should be made, as needed, to keep the effort required to assemble the calibration data set within reasonable bounds. For the data elements identified in Table A-2 as desirable, but not required, it is recommended that actual data be used if available, but assumptions are suggested in the table for application where data are not available.

Table A-2. Data Needs for Calibration of Part C Predictive Models by Facility Type

Chapter	Data Element	Data Need: Required	Data Need: Desirable	Default Assumption
ROADWAY SEGMENTS				
10—Rural Two-Lane, Two-Way Roads	Segment length	X		Need actual data
	Average annual daily traffic (AADT)	X		Need actual data
	Lengths of horizontal curves and tangents	X		Need actual data
	Radii of horizontal curves	X		Need actual data
	Presence of spiral transition for horizontal curves		X	Base default on agency design policy
	Superelevation variance for horizontal curves		X	No superelevation variance
	Percent grade		X	Base default on terrain[a]
	Lane width	X		Need actual data
	Shoulder type	X		Need actual data
	Shoulder width	X		Need actual data
	Presence of lighting		X	Assume no lighting
	Driveway density		X	Assume 5 driveways per mile
	Presence of passing lane		X	Assume not present
	Presence of short four-lane section		X	Assume not present
	Presence of center two-way left-turn lane	X		Need actual data
	Presence of centerline rumble strip		X	Base default on agency design policy
	Roadside hazard rating		X	Assume roadside hazard rating = 3
	Use of automated speed enforcement		X	Base default on current practice
11—Rural Multilane Highways	*For all rural multilane highways:*			
	Segment length	X		Need actual data
	Average annual daily traffic (AADT)	X		Need actual data
	Lane width	X		Need actual data
	Shoulder width	X		Need actual data
	Presence of lighting	X		Assume no lighting
	Use of automated speed enforcement		X	Base default on current practice
	For undivided highways only:			
	Sideslope	X		Need actual data
	For divided highways only:			
	Median width	X		Need actual data

Table A-2. Continued on next page

Table A-2. Data Needs for Calibration of Part C Predictive Models by Facility Type *continued*

Chapter	Data Element	Data Need: Required	Data Need: Desirable	Default Assumption
12—Urban and Suburban Arterials	Segment length	X		Need actual data
	Number of through traffic lanes	X		Need actual data
	Presence of median	X		Need actual data
	Presence of center two-way left-turn lane	X		Need actual data
	Average annual daily traffic (AADT)	X		Need actual data
	Number of driveways by land-use type	X		Need actual data[b]
	Low-speed vs. intermediate or high speed	X		Need actual data
	Presence of on-street parking	X		Need actual data
	Type of on-street parking	X		Need actual data
	Roadside fixed object density		X	database default on fixed-object offset and density categories[c]
	Presence of lighting		X	Base default on agency practice
	Presence of automated speed enforcement		X	Base default on agency practice
INTERSECTIONS				
10—Rural Two-Lane, Two-Way Roads	Number of intersection legs	X		Need actual data
	Type of traffic control	X		Need actual data
	Average annual daily traffic (AADT) for major road	X		Need actual data
	Average daily traffic (AADT) for minor road	X		Need actual data or best estimate
	Intersection skew angle		X	Assume no skew[d]
	Number of approaches with left-turn lanes	X		Need actual data
	Number of approaches with right-turn lanes	X		Need actual data
	Presence of lighting	X		Need actual data
11—Rural Multilane Highways	*For all rural multilane highways:*			
	Number of intersection legs	X		Need actual data
	Type of traffic control	X		Need actual data
	Average annual daily traffic (AADT) for major road	X		Need actual data
	Average annual daily traffic (AADT) for minor road	X		Need actual data or best estimate
	Presence of lighting	X		Need actual datad
	Intersection skew angle		X	Assume no skew
	Number of approaches with left-turn lanes	X		Need actual data
	Number of approaches with right-turn lanes	X		Need actual data

Table A-2. Continued on next page

Table A-2. Data Needs for Calibration of Part C Predictive Models by Facility Type *continued*

Chapter	Data Element	Data Need		Default Assumption
		Required	Desirable	
12—Urban and Suburban Arterials	*For all intersections on arterials:*			
	Number of intersection legs	X		Need actual data
	Type of traffic control	X		Need actual data
	Average annual daily traffic (AADT) for major road	X		Need actual data
	Average annual daily traffic (AADT) for minor road	X		Need actual data or best estimate
	Number of approaches with left-turn lanes	X		Need actual data
	Number of approaches with right-turn lanes	X		Need actual data
	Presence of lighting	X		Need actual data
	For signalized intersections only:			
	Presence of left-turn phasing	X		Need actual data
	Type of left-turn phasing	X		Prefer actual data, but agency practice may be used as a default
	Use of right-turn-on-red signal operation	X		Need actual data
	Use of red-light cameras	X		Need actual data
	Pedestrian volume		X	Estimate with Table 12-21
	Maximum number of lanes crossed by pedestrians on any approach		X	Estimate from number of lanes and presence of median on major road
	Presence of bus stops within 1,000 ft		X	Assume not present
	Presence of schools within 1,000 ft		X	Assume not present
	Presence of alcohol sales establishments within 1,000 ft		X	Assume not present

[a] Suggested default values for calibration purposes: *CMF* = 1.00 for level terrain; *CMF* = 1.06 for rolling terrain; *CMF* = 1.14 for mountainous terrain
[b] Use actual data for number of driveways, but simplified land-use categories may be used (e.g., commercial and residential only).
[c] CMFs may be estimated based on two categories of fixed-object offset (O_{fo})—either 5 or 20 ft—and three categories of fixed-object density (D_{fo})—0, 50, or 100 objects per mile.
[d] If measurements of intersection skew angles are not available, the calibration should preferably be performed for intersections with no skew.

A.1.1.4. Step 4—Apply the Applicable Part C Predictive Method to Predict Total Crash Frequency for Each Site During the Calibration Period as a Whole

The site characteristics data assembled in Step 3 should be used to apply the applicable predictive method from Chapter 10, 11, or 12 to each site in the calibration data set. For this application, the predictive method should be applied without using the EB Method and, of course, without employing a calibration factor (i.e., a calibration factor of 1.00 is assumed). Using the predictive models, the expected average crash frequency is obtained for either one, two, or three years, depending on the duration of the calibration period selected.

A.1.1.5. Step 5—Compute Calibration Factors for Use in Part C Predictive Models

The final step is to compute the calibration factor as:

$$C_r(\text{or } C_i) = \frac{\sum_{\text{all sites}} \text{observed crashes}}{\sum_{\text{all sites}} \text{predicted crashes}} \tag{A-1}$$

The computation is performed separately for each facility type. The computed calibration factor is rounded to two decimal places for application in the appropriate Part C predictive model.

Example Calibration Factor Calculation

The SPF for four-leg signalized intersections on rural two-lane, two-way roads from Equation 10-18 is:

$$N_{spf\ int} = e^{[-5.73 + 0.60 \times \ln(AADT_{maj}) + 0.20 \times \ln(AADT_{min})]}$$

Where:

$N_{spf\ int}$ = predicted number of total intersection-related crashes per year for base conditions;

$AADT_{maj}$ = average annual daily entering traffic volumes (vehicles/day) on the major road; and

$AADT_{min}$ = average annual daily entering traffic volumes (vehicles/day) on the minor road.

The base conditions are:

- No left-turn lanes on any approach
- No right-turn lanes on any approach

The CMF values from Chapter 10 are:

- CMF for one approach with a left-turn lane = 0.82
- CMF for one approach with a right-turn lane = 0.96
- CMF for two approaches with right-turn lanes = 0.92
- No lighting present (so lighting CMF = 1.00 for all cases)

Typical data for eight intersections is shown in an example calculation shown below. Note that for an actual calibration, the recommended minimum sample size would be 30 to 50 sites that experience at least 100 crashes per year. Thus, the number of sites used here is smaller than recommended, and is intended solely to illustrate the calculations.

For the first intersection in the example the predicted crash frequency for base conditions is:

$$N_{bibase} = e^{(-5.73 + 0.60 \times \ln(4000) + 0.20 \times \ln(2000))} = 2.152 \text{ crashes/year}$$

The intersection has a left-turn lane on the major road, for which CMF_{1i} is 0.67, and a right-turn lane on one approach, a feature for which CMF_{2i} is 0.98. There are three years of data, during which four crashes were observed (shown in Column 10 of Table Ex-1). The predicted average crash frequency from the Chapter 10 for this intersection without calibration is from Equation 10-2:

$$N_{bi} = (N_{bibase}) \times (CMF_{1i}) \times (CMF_{2i}) \times (\text{number of years of data})$$

$$= 2.152 \times 0.67 \times 0.98 \times 3 = 4.240 \text{ crashes in three years, shown in Column 9.}$$

Similar calculations were done for each intersection in the table shown below. The sum of the observed crash frequencies in Column 10 (43) is divided by the sum of the predicted average crash frequencies in Column 9 (45.594) to obtain the calibration factor, C_i, equal to 0.943. It is recommended that calibration factors be rounded to two decimal places, so calibration factor equal to 0.94 should be used in the Chapter 10 predictive model for four-leg signalized intersections.

Table Ex-1. Example of Calibration Factor Computation

1	2	3	4	5	6	7	8	9	10
$AADT_{maj}$	$AADT_{min}$	SPF Prediction	Intersection Approaches with Left-Turn Lanes	CMF_{1i}	Intersection Approaches with Right-Turn Lane	CMF_{2i}	Years of Data	Predicted Average Crash Frequency	Observed Crash Frequency
4000	2000	2.152	1	0.67	1	0.98	3	4.240	4
3000	1500	1.710	0	1.00	2	0.95	2	3.249	5
5000	3400	2.736	0	1.00	2	0.95	3	7.799	10
6500	3000	3.124	0	1.00	2	0.95	3	8.902	5
3600	2300	2.078	1	0.67	1	0.98	3	4.093	2
4600	4500	2.753	0	1.00	2	0.95	3	7.846	8
5700	3300	2.943	1	0.67	1	0.98	3	5.796	5
6800	1500	2.794	1	0.67	1	0.98	2	3.669	4
							Sum	45.594	43
							Calibration Factor (C_i)		0.943

A.1.2. Development of Jurisdiction-Specific Safety Performance Functions for Use in the Part C Predictive Method

Satisfactory results from the Part C predictive method can be obtained by calibrating the predictive model for each facility type, as explained in Appendix A.1.1. However, some users may prefer to develop jurisdiction-specific SPFs using their agency's own data, and this is likely to enhance the reliability of the Part C predictive method. While there is no requirement that this be done, HSM users are welcome to use local data to develop their own SPFs, or if they wish, replace some SPFs with jurisdiction-specific models and retain other SPFs from the Part C chapters. Within the first two to three years after a jurisdiction-specific SPF is developed, calibration of the jurisdiction-specific SPF using the procedure presented in Appendix A.1.1 may not be necessary, particularly if other default values in the Part C models are replaced with locally-derived values, as explained in Appendix A.1.3.

If jurisdiction-specific SPFs are used in the Part C predictive method, they need to be developed with methods that are statistically valid and developed in such a manner that they fit into the applicable Part C predictive method. The following guidelines for development of jurisdiction-specific SPFs that are acceptable for use in Part C include:

- In preparing the crash data to be used for development of jurisdiction-specific SPFs, crashes are assigned to roadway segments and intersections following the definitions explained in Appendix A.2.3 and illustrated in Figure A-1.
- The jurisdiction-specific SPF should be developed with a statistical technique such as negative binomial regression that accounts for the overdispersion typically found in crash data and quantifies an overdispersion parameter so that the model's predictions can be combined with observed crash frequency data using the EB Method.
- The jurisdiction-specific SPF should use the same base conditions as the corresponding SPF in Part C or should be capable of being converted to those base conditions.
- The jurisdiction-specific SPF should include the effects of the following traffic volumes: average annual daily traffic volume for roadway segment and major- and minor-road average annual daily traffic volumes for intersections.
- The jurisdiction-specific SPF for any roadway segment facility type should have a functional form in which predicted average crash frequency is directly proportional to segment length.

These guidelines are not intended to stifle creativity and innovation in model development. However, a model that does not account for overdispersed data or that cannot be integrated with the rest of the Part C predictive method will not be useful.

Two types of data sets may be used for SPF development. First, SPFs may be developed using only data that represent the base conditions, which are defined for each SPF in Chapters 10, 11, and 12. Second, it is also acceptable to develop models using data for a broader set of conditions than the base conditions. In this approach, all variables that are part of the applicable base-condition definition, but have non-base-condition values, should be included in an initial model. Then, the initial model should be made applicable to the base conditions by substituting values that correspond to those base conditions into the model. Several examples of this process are presented in Appendix 10A.

A.1.3. Replacement of Selected Default Values in the Part C Predictive Models to Local Conditions

The Part C predictive models use many default values that have been derived from crash data in HSM-related research. For example, the urban intersection predictive model in Chapter 12 uses pedestrian factors that are based on the proportion of pedestrian crashes compared to total crashes. Replacing these default values with locally derived values will improve the reliability of the Part C predictive models. Table A-3 identifies the specific tables in Part C that may be replaced with locally derived values. In addition to these tables, there is one equation—Equation 10-18—which uses constant values given in the accompanying text in Chapter 10. These constant values may be replaced with locally derived values.

Providing locally-derived values for the data elements identified in Table A-3 is optional. Satisfactory results can be obtained with the Part C predictive models, as they stand, when the predictive model for each facility type is calibrated with the procedure given in Appendix A.1.1. But, more reliable results may be obtained by updating the data elements listed in Table A-3. It is acceptable to replace some, but not all of these data elements, if data to replace all of them are not available. Each element that is updated with locally-derived values should provide a small improvement in the reliability of that specific predictive model. To preserve the integrity of the Part C predictive method, the quantitative values in the predictive models, (other than those listed in Table A-3 and those discussed in Appendices A.1.1 and A.2.2), should not be modified. Any replacement values derived with the procedures presented in this section should be incorporated in the predictive models before the calibration described in Appendix A.1.1 is performed.

Table A-3. Default Crash Distributions Used in Part C Predictive Models Which May Be Calibrated by Users to Local Conditions

Chapter	Table or Equation Number	Type of Roadway Element: Roadway Segments	Type of Roadway Element: Intersections	Data Element or Distribution That May Be Calibrated to Local Conditions
10—Rural Two-Lane, Two-Way Roads	Table 10-3	X		Crash severity by facility type for roadway segments
	Table 10-4	X		Collision type by facility type for roadway segments
	Table 10-5		X	Crash severity by facility type for intersections
	Table 10-6		X	Collision type by facility type for intersections
	Equation 10-18	X		Driveway-related crashes as a proportion of total crashes (p_{dwy})
	Table 10-12	X		Nighttime crashes as a proportion of total crashes by severity level
	Table 10-15		X	Nighttime crashes as a proportion of total crashes by severity level and by intersection type
11—Rural Multilane Highways	Table 11-4	X		Crash severity and collision type for undivided segments
	Table 11-6	X		Crash severity and collision type for divided segments
	Table 11-9		X	Crash severity and collision type by intersection type
	Table 11-15	X		Nighttime crashes as a proportion of total crashes by severity level and by roadway segment type for undivided roadway segments
	Table 11-19		X	Nighttime crashes as a proportion of total crashes by severity level and by roadway segment type for divided roadway segments
	Table 11-24		X	Nighttime crashes as a proportion of total crashes by severity level and by intersection type
12—Urban and Suburban Arterials	Table 12-4	X		Crash severity and collision type for multiple-vehicle nondriveway collisions by roadway segment type
	Table 12-6	X		Crash severity and collision type for single-vehicle crashes by roadway segment type
	Table 12-7	X		Crash severity for driveway-related collisions by roadway segment type[a]
	Table 12-8	X		Pedestrian crash adjustment factor by roadway segment type
	Table 12-9	X		Bicycle crash adjustment factor by roadway segment type
	Table 12-11		X	Crash severity and collision type for multiple-vehicle collisions by intersection type
	Table 12-13		X	Crash severity and collision type for single-vehicle crashes by intersection type
	Table 12-16		X	Pedestrian crash adjustment factor by intersection type for stop-controlled intersections
	Table 12-17		X	Bicycle crash adjustment factor by intersection type
	Table 12-23	X		Nighttime crashes as a proportion of total crashes by severity level and by roadway segment type
	Table 12-27		X	Nighttime crashes as a proportion of total crashes by severity level and by intersection type

[a] The only portion of Table 12-7 that should be modified by the user are the crash severity proportions.
Note: No quantitative values in the Part C predictive models, other than those listed here and those discussed in Appendices A.1.1 and A.1.2, should be modified by HSM users.

Procedures for developing replacement values for each data element identified in Table A-3 are presented below. Most of the data elements to be replaced are proportions of crash severity levels and/or crash types that are part of a specific distribution. Each replacement value for a given facility type should be derived from data for a set of sites that, as a group, includes at least 100 crashes and preferably more. The duration of the study period for a given set of sites may be as long as necessary to include at least 100 crashes. In the following discussion, the term "sufficient data" refers to a data set including a sufficient number of sites to meet this criterion for total crashes. In a few cases, explicitly identified below, the definition of sufficient data will be expressed in terms of a crash category other than total crashes. In assembling data for developing replacements for default values, crashes are to be assigned to specific roadway segments or intersections following the definitions explained in Appendix A.2.3 and illustrated in Figure A-1.

A.1.3.1. Replacement of Default Values for Rural Two-Lane, Two-Way Roads

Five specific sets of default values for rural two-lane, two-way roads may be updated with locally-derived replacement values by HSM users. Procedures to develop each of these replacement values are presented below.

Crash Severity by Facility Type

Tables 10-3 and 10-5 present the distribution of crashes by five crash severity levels for roadway segments and intersections, respectively, on rural two-lane, two-way roads. If sufficient data, including these five severity levels (fatal, incapacitating injury, nonincapacitating injury, possible injury, and property damage only), are available for a given facility type, the values in Tables 10-3 and 10-5 for that facility type may be updated. If sufficient data are available only for the three standard crash severity levels (fatal, injury, and property damage only), the existing values in Tables 10-3 and 10-5 may be used to allocate the injury crashes to specific injury severity levels (incapacitating injury, nonincapacitating injury, and possible injury).

Collision Type by Facility Type

Table 10-4 presents the distribution of crashes by collision type for seven specific types of single-vehicle crashes and six specific types of multiple-vehicle crashes for roadway segments, and Table 10-6 presents the distribution of crashes by collision type for three intersection types on rural two-lane, two-way roads. If sufficient data are available for a given facility type, the values in Tables 10-4 and 10-6 for that facility type may be updated.

Driveway-Related Crashes as a Proportion of Total Crashes for Roadway Segments

Equation 10-18 includes a factor, p_{dwy}, which represents the proportion of total crashes represented by driveway-related crashes. A value for p_{dwy} based on research is presented in the accompanying text. This value may be replaced with a locally-derived value, if data are available for a set for sites that, as a group, have experienced at least 100 driveway-related crashes.

Nighttime Crashes as a Proportion of Total Crashes for Roadway Segments

Table 10-12 presents the proportions of total nighttime crashes by severity level and the proportion of total crashes that occur at night for roadway segments on rural two-lane, two-way roads. These values may be replaced with locally-derived values for a given facility type, if data are available for a set of sites that, as a group, have experienced at least 100 nighttime crashes.

Nighttime Crashes as a Proportion of Total Crashes for Intersections

Table 10-15 presents the proportion of total crashes that occur at night for intersections on rural two-lane, two-way roads. These values may be replaced with locally-derived values for a given facility type, if data are available for a set of sites that, as a group, have experienced at least 100 nighttime crashes.

A.1.3.2. Replacement of Default Values for Rural Multilane Highways

Five specific sets of default values for rural multilane highways may be updated with locally-derived replacement values by HSM users. Procedures to develop each of these replacement values are presented below.

Crash Severity and Collision Type for Undivided Roadway Segments

Table 11-4 presents the combined distribution of crashes for four crash severity levels and six collision types. If sufficient data are available for undivided roadway segments, the values in Table 11-4 for this facility type may be updated. Given that this is a joint distribution of two variables, sufficient data for this application requires a set of sites of a given type that, as a group, have experienced at least 200 crashes in the time period for which data are available.

Crash Severity and Collision Type for Divided Roadway Segments

Table 11-6 presents the combined distribution of crashes for four crash severity levels and six collision types. If sufficient data are available for divided roadway segments, the values in Table 11-6 for this facility type may be updated. Given that this is a joint distribution of two variables, sufficient data for this application requires sites that have experienced at least 200 crashes in the time period for which data are available.

Crash Severity and Collision Type by Intersection Type

Table 11-9 presents the combined distribution of crashes at intersections for four crash severity levels and six collision types. If sufficient data are available for a given intersection type, the values in Table 11-9 for that intersection type may be updated. Given that this is a joint distribution of two variables, sufficient data for this application requires a set of sites of a given type that, as a group, have experienced at least 200 crashes in the time period for which data are available.

Nighttime Crashes as a Proportion of Total Crashes for Roadway Segments

Tables 11-15 and 11-19 present the proportions of total nighttime crashes by severity level and the proportion of total crashes that occur at night for undivided and divided roadway segments, respectively, on rural multilane highways. These values may be replaced with locally-derived values for a given facility type, if data are available for a set of sites that, as a group, have experienced at least 100 nighttime crashes.

Nighttime Crashes as a Proportion of Total Crashes for Intersections

Table 11-24 presents the proportion of total crashes that occur at night for intersections on rural multilane highways. These values may be replaced with locally-derived values for a given facility type, if data are available for a set of sites that, as a group, have experienced at least 100 nighttime crashes.

A.1.3.3. Replacement of Default Values for Urban and Suburban Arterials

Eleven specific sets of default values for urban and suburban arterial highways may be updated with locally-derived replacement values by HSM users. Procedures to develop each of these replacement values are presented below.

Crash Severity and Collision Type for Multiple-Vehicle Nondriveway Crashes by Roadway Segment Type

Table 12-4 presents the combined distribution of crashes for two crash severity levels and six collision types. If sufficient data are available for a given facility type, the values in Table 12-4 for that facility type may be updated. Given that this is a joint distribution of two variables, sufficient data for this application requires a set of sites of a given type that, as a group, have experienced at least 200 crashes in the time period for which data are available.

Crash Severity and Collision Type for Single-Vehicle Crashes by Roadway Segment Type

Table 12-6 presents the combined distribution of crashes for two crash severity levels and six collision types. If sufficient data are available for a given facility type, the values in Table 12-6 for that facility type may be updated. Given that this is a joint distribution of two variables, sufficient data for this application requires a set of sites of a given type that, as a group, have experienced at least 200 crashes in the time period for which data are available.

Crash Severity for Driveway-Related Collision by Roadway Segment Type

Table 12-7 includes data on the proportions of driveway-related crashes for two crash severity levels (fatal-and-injury and property-damage-only crashes) by facility type for roadway segments. If sufficient data are available for a given facility type, these specific severity-related values in Table 12-7 for that facility type may be updated. The rest of Table 12-7, other than the last two rows of data which are related to crash severity, should not be modified.

Pedestrian Crash Adjustment Factor by Roadway Segment Type

Table 12-8 presents a pedestrian crash adjustment factor for specific roadway segment facility types and for two speed categories: low speed (traffic speeds or posted speed limits of 30 mph or less) and intermediate or high speed (traffic speeds or posted speed limits greater than 30 mph). For a given facility type and speed category, the pedestrian crash adjustment factor is computed as:

$$f_{pedr} = \frac{K_{ped}}{K_{non}} \tag{A-2}$$

Where:

f_{pedr} = pedestrian crash adjustment factor;

K_{ped} = observed vehicle-pedestrian crash frequency; and

K_{non} = observed frequency for all crashes not including vehicle-pedestrian and vehicle-bicycle crash.

The pedestrian crash adjustment factor for a given facility type should be determined with a set of sites of that speed type that, as a group, includes at least 20 vehicle-pedestrian collisions.

Bicycle Crash Adjustment Factor by Roadway Segment Type

Table 12-9 presents a bicycle crash adjustment factor for specific roadway segment facility types and for two speed categories: low speed (traffic speeds or posted speed limits of 30 mph or less) and intermediate or high speed (traffic speeds or posted speed limits greater than 30 mph). For a given facility type and speed category, the bicycle crash adjustment factor is computed as:

$$f_{biker} = \frac{K_{bike}}{K_{non}} \tag{A-3}$$

Where:

f_{biker} = bicycle crash adjustment factor;

K_{bike} = observed vehicle-bicycle crash frequency; and

K_{non} = observed frequency for all crashes not including vehicle-pedestrian and vehicle-bicycle crashes.

The bicycle crash adjustment factor for a given facility type should be determined with a set of sites of that speed type that, as a group, includes at least 20 vehicle-bicycle collisions.

Crash Severity and Collision Type for Multiple-Vehicle Crashes by Intersection Type

Table 12-11 presents the combined distribution of crashes for two crash severity levels and six collision types. If sufficient data are available for a given facility type, the values in Table 12-11 for that facility type may be updated. Given that this is a joint distribution of two variables, sufficient data for this application requires a set of sites of a given type that, as a group, have experienced at least 200 crashes in the time period for which data are available.

Crash Severity and Collision Type for Single-Vehicle Crashes by Intersection Type

Table 12-13 presents the combined distribution of crashes for two crash severity levels and six collision types. If sufficient data are available for a given facility type, the values in Table 12-13 for that facility type may be updated. Given that this is a joint distribution of two variables, sufficient data for this application requires a set of sites of a given type that, as a group, have experienced at least 200 crashes in the time period for which data are available.

Pedestrian Crash Adjustment Factor by Intersection Type

Table 12-16 presents a pedestrian crash adjustment factor for two specific types of intersections with stop control on the minor road. For a given facility type and speed category, the pedestrian crash adjustment factor is computed using Equation A-2. The pedestrian crash adjustment factor for a given facility type is determined with a set of sites that, as a group, have experienced at least 20 vehicle-pedestrian collisions.

Bicycle Crash Adjustment Factor by Intersection Type

Table 12-17 presents a bicycle crash adjustment factor for four specific intersection facility types. For a given facility type and speed category, the bicycle crash adjustment factor is computed using Equation A-3. The bicycle crash adjustment factor for a given facility type is determined with a set of sites that, as a group, have experienced at least 20 vehicle-bicycle collisions.

Nighttime Crashes as a Proportion of Total Crashes for Roadway Segments

Table 12-23 presents the proportions of total nighttime crashes by severity level for specific facility types for roadway segments and the proportion of total crashes that occur at night. These values may be replaced with locally-derived values for a given facility type, if data are available for a set of sites that, as a group, have experienced at least 100 nighttime crashes.

Nighttime Crashes as a Proportion of Total Crashes for Intersections

Table 12-27 presents the proportions of total nighttime crashes by severity level for specific facility types for intersections and the proportion of total crashes that occur at night. These values may be replaced with locally-derived values for a given facility type, if data are available for a set of sites that, as a group, have experienced at least 100 nighttime crashes.

A.2. USE OF THE EMPIRICAL BAYES METHOD TO COMBINE PREDICTED AVERAGE CRASH FREQUENCY AND OBSERVED CRASH FREQUENCY

Application of the EB Method provides a method to combine the estimate using a Part C predictive model and observed crash frequencies to obtain a more reliable estimate of expected average crash frequency. The EB Method is a key tool to compensate for the potential bias due to regression-to-the-mean. Crash frequencies vary naturally from one time period to the next. When a site has a higher than average frequency for a particular time period, the site is likely to have lower crash frequency in subsequent time periods. Statistical methods can help to assure that this natural decrease in crash frequency following a high observed value is not mistaken for the effect of a project or for a true shift in the long-term expected crash frequency.

There are several statistical methods that can be employed to compensate for regression-to-the-mean. The EB Method is used in the HSM because it is best suited to the context of the HSM. The Part C predictive models include negative binomial regression models that were developed before the publication of the HSM by researchers who had no data on the specific sites to which HSM users would later apply those predictive models. The HSM users are generally engineers and planners, without formal statistical training, who would not generally be capable of developing custom models for each set of the sites they wish to apply the HSM to and, even if there were, would have no wish to spend the time and effort needed for model development each time they apply the HSM. The EB Method provides the most suitable tool for compensating for regression-to-the-mean that works in this context.

Each of the Part C chapters presents a four-step process for applying the EB Method. The EB Method assumes that the appropriate Part C predictive model (see Section 10.3.1 for rural two-lane, two-way roads, Section 11.3.1 for rural multilane highways, or Section 12.3.1 for urban and suburban arterials) has been applied to determine the predicted crash frequency for the sites that make up a particular project or facility for a particular past time period of interest. The steps in applying the EB Method are:

- Determine whether the EB Method is applicable, as explained in Appendix A.2.1.
- Determine whether observed crash frequency data are available for the project or facility for the time period for which the predictive model was applied and, if so, obtain those crash frequency data, as explained in Appendix A.2.2. Assign each crash instance to individual roadway segments and intersections, as explained in Appendix A.2.3.
- Apply the EB Method to estimate the expected crash frequency by combining the predicted and observed crash frequencies for the time period of interest. The site-specific EB Method, applicable when observed crash frequency data are available for the individual roadway segments and intersections that make up a project or facility, is presented in Appendix A.2.4. The project-level EB Method, applicable when observed crash frequency data are available only for the project or facility as a whole, is presented in Appendix A.2.5.
- Adjust the estimated value of expected crash frequency to a future time period, if appropriate, as explained in Appendix A.2.6.

Consideration of observed crash history data in the Part C predictive method increases the reliability of the estimate of the expected crash frequencies. When at least two years of observed crash history data are available for the facility or project being evaluated, and when the facility or project meets certain criteria discussed below, the observed crash data should be used. When considering observed crash history data, the procedure must consider both the existing geometric design and traffic control for the facility or project (i.e., the conditions that existed during the before period while the observed crash history was accumulated) and the proposed geometric design and traffic control for the project (i.e., the conditions that will exist during the after period, the period for which crash predictions are being made). In estimating the expected crash frequency for an existing arterial facility in a future time period where no improvement project is planned, only the traffic volumes should differ between the before and after periods. For an arterial on which an improvement project is planned, traffic volumes, geometric design features, and traffic control features may all change between the before and after periods. The EB Method presented below provides a method to combine predicted and observed crash frequencies.

A.2.1 Determine whether the EB Method is Applicable

The applicability of the EB Method to a particular project or facility depends on the type of analysis being performed and the type of future project work that is anticipated. If the analysis is being performed to assess the expected average crash frequency of a specific highway facility, but is not part of the analysis of a planned future project, then the EB Method should be applied. If a future project is being planned, then the nature of that future project should be considered in deciding whether to apply the EB Method.

The EB Method should be applied for the analyses involving the following future project types:

- Sites at which the roadway geometrics and traffic control are not being changed (e.g., the “do-nothing” alternative);
- Projects in which the roadway cross section is modified but the basic number of through lanes remains the same (This would include, for example, projects for which lanes or shoulders were widened or the roadside was improved, but the roadway remained a rural two-lane highway);
- Projects in which minor changes in alignment are made, such as flattening individual horizontal curves while leaving most of the alignment intact;
- Projects in which a passing lane or a short four-lane section is added to a rural two-lane, two-way road to increase passing opportunities; and
- Any combination of the above improvements.

The EB Method is not applicable to the following types of improvements:

- Projects in which a new alignment is developed for a substantial proportion of the project length; and
- Intersections at which the basic number of intersection legs or type of traffic control is changed as part of a project.

The reason that the EB Method is not used for these project types is that the observed crash data for a previous time period is not necessarily indicative of the crash experience that is likely to occur in the future after such a major geometric improvement. Since, for these project types, the observed crash frequency for the existing design is not relevant to estimation of the future crash frequencies for the site, the EB Method is not needed and should not be applied. If the EB Method is applied to individual roadway segments and intersections, and some roadway segments and intersections within the project limits will not be affected by the major geometric improvement, it is acceptable to apply the EB Method to those unaffected segments and intersections.

If the EB Method is not applicable, do not proceed to the remaining steps. Instead, follow the procedure described in the Applications section of the applicable Part C chapter.

A.2.2. Determine whether Observed Crash Frequency Data are Available for the Project or Facility and, if so, Obtain those Data

If the EB Method is applicable, it should be determined whether observed crash frequency data are available for the project or facility of interest directly from the jurisdiction's crash record system or indirectly from another source. At least two years of observed crash frequency data are desirable to apply the EB Method. The best results in applying the EB Method will be obtained if observed crash frequency data are available for each individual roadway segment and intersection that makes up the project of interest. The EB Method applicable to this situation is presented in Appendix A.2.4. Criteria for assigning crashes to individual roadway segments and intersections are presented in Appendix A.2.3. If observed crash frequency data are not available for individual roadway segments and intersections, the EB Method can still be applied if observed crash frequency data are available for the project or facility as a whole. The EB Method applicable to this situation is presented in Appendix A.2.5.

If appropriate crash frequency data are not available, do not proceed to the remaining steps. Instead, follow the procedure described in the Applications section of the applicable Part C chapter.

A.2.3. Assign Crashes to Individual Roadway Segments and Intersections for Use in the EB Method

The Part C predictive method has been developed to estimate crash frequencies separately for intersections and roadways segments. In the site-specific EB Method presented in Appendix A.2.4, observed crashes are combined with the predictive model estimate of crash frequency to provide a more reliable estimate of the expected average crash frequency of a particular site. In Step 6 of the predictive method, if the site-specific EB Method is applicable, observed crashes are assigned to each individual site identified within the facility of interest. Because the predictive models estimate crashes separately for intersections and roadway segments, which may physically overall in some cases, observed crashes are differentiated and assigned as either intersection related crashes or roadway segment related crashes.

Intersection crashes include crashes that occur at an intersection (i.e., within the curb limits) and crashes that occur on the intersection legs and are intersection-related. All crashes that are not classified as intersection or intersection-related crashes are considered to be roadway segment crashes. Figure A-1 illustrates the method used to assign crashes to roadway segments or intersections. As shown:

- All crashes that occur within the curbline limits of an intersection (Region A in the figure) are assigned to that intersection.
- Crashes that occur outside the curbline limits of an intersection (Region B in the figure) are assigned to either the roadway segment on which they occur or an intersection, depending on their characteristics. Crashes that are classified on the crash report as intersection-related or have characteristics consistent with an intersection-related crash are assigned to the intersection to which they are related; such crashes would include rear-end collisions related to queues on an intersection approach. Crashes that occur between intersections and are not related to an intersection, such as collisions related to turning maneuvers at driveways, are assigned to the roadway segment on which they occur.

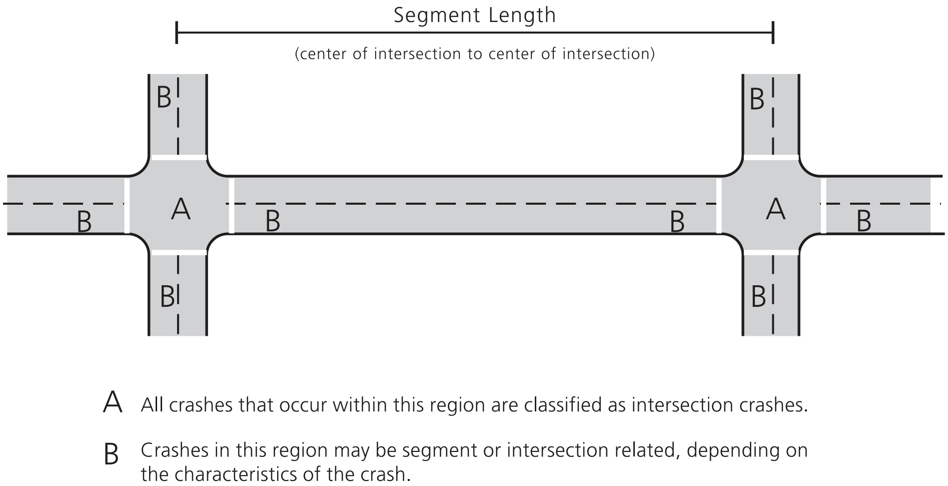

Figure A-1. Definition of Roadway Segments and Intersections

In some jurisdictions, crash reports include a field that allows the reporting officer to designate the crash as intersection-related. When this field is available on the crash reports, crashes should be assigned to the intersection or the segment based on the way the officer marked the field on the report. In jurisdictions where there is not a field on the crash report that allows the officer to designate crashes as intersection-related, the characteristics of the crash may be considered to make a judgment as to whether the crash should be assigned to the intersection or the segment. Other fields on the report, such as collision type, number of vehicles involved, contributing circumstances, weather condition, pavement condition, traffic control malfunction, and sequence of events can provide helpful information in making this determination.

If the officer's narrative and crash diagram are available to the user, they can also assist in making the determination. The following crash characteristics may indicate that the crash was related to the intersection:

- Rear-end collision in which both vehicles were going straight approaching an intersection or in which one vehicle was going straight and struck a stopped vehicle
- Collision in which the report indicates a signal malfunction or improper traffic control at the intersection

The following crash characteristics may indicate that the crash was not related to the intersection and should be assigned to the segment on which it occurred:

- Collision related to a driveway or involving a turning movement not at an intersection
- Single-vehicle run-off-the-road or fixed object collision in which pavement surface condition was marked as wet or icy and identified as a contributing factor

These examples are provided as guidance when an "intersection-related" field is not available on the crash report; they are not strict rules for assigning crashes. Information on the crash report should be considered to help make the determination, which will rely on judgment. The information needed for classifying crashes is whether each crash is, or is not, related to an intersection. The consideration of crash type data is presented here only as an example of one approach to making this determination.

Using these guidelines, the roadway segment predictive models estimate the average frequency of crashes that would occur on the roadway if no intersection were present. The intersection predictive models estimate the average frequency of additional crashes that occur because of the presence of an intersection.

A.2.4. Apply the Site-Specific EB Method

Equations A-4 and A-5 are used directly to estimate the expected crash frequency for a specific site by combining the predictive model estimate with observed crash frequency. The value of $N_{expected}$ from Equation A-4 represents the expected crash frequency for the same time period represented by the predicted and observed crash frequencies. $N_{predicted}$, $N_{observed,}$ and $N_{expected}$ all represent either total crashes or a specific severity level or collision type of interest. The expected average crash frequency considering both the predictive model estimate and observed crash frequencies for an individual roadway segment or intersection is computed as:

$$N_{expected} = w \times N_{predicted} + (1 - w) \times N_{observed} \tag{A-4}$$

$$w = \frac{1}{1 + k \times \left(\sum_{\substack{\text{all study} \\ \text{years}}} N_{predicted} \right)} \tag{A-5}$$

Where:

$N_{expected}$ = estimate of expected average crashes frequency for the study period;

$N_{predicted}$ = predictive model estimate of average crash frequency predicted for the study period under the given conditions;

$N_{observed}$ = observed crash frequency at the site over the study period;

w = weighted adjustment to be placed on the predictive model estimate; and

k = overdispersion parameter of the associated SPF used to estimate $N_{predicted}$.

When observed crash data by severity level is not available, the estimate of expected average crash frequency for fatal-and-injury and property-damage-only crashes is calculated by applying the proportion of predicted average crash frequency by severity level ($N_{predicted(FI)}/N_{predicted(total)}$ and $N_{predicted(PDO)}/N_{predicted(total)}$) to the total expected average crash frequency from Equation A-4.

Equation A-5 shows an inverse relationship between the overdispersion parameter, k, and the weight, w. This implies that when a model with little overdispersion is available; more reliance will be placed on the predictive model estimate, $N_{predicted}$, and less reliance on the observed crash frequency, $N_{observed}$. The opposite is also the case; when a model with substantial overdispersion is available, less reliance will be placed on the predictive model estimate, $N_{predicted}$, and more reliance on the observed crash frequency, $N_{observed}$.

It is important to note in Equation A-5 that, as $N_{predicted}$ increases, there is less weight placed on $N_{predicted}$ and more on $N_{observed}$. This might seem counterintuitive at first. However, this implies that for longer sites and for longer study periods, there are more opportunities for crashes to occur. Thus, the observed crash history is likely to be more meaningful and the model prediction less important. So, as $N_{predicted}$ increases, the EB Method places more weight on the number of crashes that actually occur, $N_{observed}$. When few crashes are predicted, the observed crash frequency, $N_{observed}$, is not likely to be meaningful, in statistical terms, so greater reliance is placed on the predicted crash frequency, $N_{predicted}$.

The values of the overdispersion parameters, k, for the safety performance functions used in the predictive models are presented with each SPF in Sections 10.6, 11.6, and 12.6.

Since application of the EB Method requires use of an overdispersion parameter, it cannot be applied to portions of the prediction method where no overdispersion parameter is available. For example, vehicle-pedestrian and vehicle-bicycle collisions are estimated in portions of Chapter 12 from adjustment factors rather than from models and should, therefore, be excluded from the computations with the EB Method. Chapter 12 uses multiple models with different overdispersion parameters in safety predictions for any specific roadway segment or intersection. Where observed crash data are aggregated so that the corresponding value of predicted crash frequency is determined as the sum of the results from multiple predictive models with differing overdispersion parameters, the project-level EB Method presented in Appendix A.2.5 should be applied rather than the site-specific method presented here.

Chapters 10, 11, and 12 each present worksheets that can be used to apply the site-specific EB Method as presented in this section.

Appendix A.2.6 explains how to update $N_{expected}$ to a future time period, such as the time period when a proposed future project will be implemented. This procedure is only applicable if the conditions of the proposed project will not be substantially different from the roadway conditions during which the observed crash data was collected.

A.2.5. Apply the Project-Level EB Method

HSM users may not always have location specific information for observed crash data for the individual roadway segments and intersections that make up a facility or project of interest. Alternative procedures are available where observed crash frequency data are aggregated across several sites (e.g., for an entire facility or project). This requires a more complex EB Method for two reasons. First, the overdispersion parameter, k, in the denominator of Equation A-5 is not uniquely defined, because estimate of crash frequency from two or more predictive models with different overdispersion parameters are combined. Second, it cannot be assumed, as is normally done, that the expected average crash frequency for different site types are statistically correlated with one another. Rather, an estimate of expected average crash frequency should be computed based on the assumption that the various roadway segments and intersections are statistically independent ($r = 0$) and on the alternative assumption that they are perfectly correlated ($r = 1$). The expected average crash frequency is then estimated as the average of the estimates for $r = 0$ and $r = 1$.

The following equations implement this approach, summing the first three terms, which represent the three roadway-segment-related crash types, over the five types of roadway segments considered in the (2U, 3T, 4U, 4D, 5T) and the last two terms, which represent the two intersection-related crash types, over the four types of intersections (3ST, 3SG, 4ST, 4SG):

$$N_{\text{predicted (total)}} = \sum_{j=1}^{5} N_{\text{predicted } rmj} + \sum_{j=1}^{5} N_{\text{predicted } rsj} + \sum_{j=1}^{5} N_{\text{predicted } rdj} + \sum_{j=1}^{4} N_{\text{predicted } imj} + \sum_{j=1}^{4} N_{\text{predicted } isj} \tag{A-6}$$

$$N_{\text{observed (total)}} = \sum_{j=1}^{5} N_{\text{observed } rmj} + \sum_{j=1}^{5} N_{\text{observed } rsj} + \sum_{j=1}^{5} N_{\text{observed } rdj} + \sum_{j=1}^{4} N_{\text{observed } imj} + \sum_{j=1}^{4} N_{\text{observed } isj} \tag{A-7}$$

$$N_{\text{predicted } w0} = \sum_{j=1}^{5} k_{rmj} N_{rmj}^2 + \sum_{j=1}^{5} k_{rsj} N_{rsj}^2 + \sum_{j=1}^{5} k_{rdj} N_{rdj}^2 + \sum_{j=1}^{4} k_{imj} N_{imj}^2 + \sum_{j=1}^{4} k_{isj} N_{isj}^2 \tag{A-8}$$

$$N_{\text{predicted } w1} = \sum_{j=1}^{5} \sqrt{k_{rmj} N_{rmj}} + \sum_{j=1}^{5} \sqrt{k_{rsj} N_{isj}} + \sum_{j=1}^{5} \sqrt{k_{rdj} N_{rdj}} + \sum_{j=1}^{4} \sqrt{k_{imj} N_{imj}} + \sum_{j=1}^{4} \sqrt{k_{isj} N_{isj}} \tag{A-9}$$

$$w_0 = \frac{1}{1 + \dfrac{N_{\text{predicted } w0}}{N_{\text{predicted (total)}}}} \quad \text{(A-10)}$$

$$N_0 = w_0 N_{\text{predicted (total)}} + (1 - w_0) N_{\text{observed (total)}} \quad \text{(A-11)}$$

$$w_1 = \frac{1}{1 + \dfrac{N_{\text{predicted } w1}}{N_{\text{predicted (total)}}}} \quad \text{(A-12)}$$

$$N_1 = w_1 N_{\text{predicted (total)}} + (1 - w_1) N_{\text{observed (total)}} \quad \text{(A-13)}$$

$$N_{\text{expected}/comb} = \frac{N_0 + N_1}{2} \quad \text{(A-14)}$$

Where:

$N_{\text{predicted (total)}}$ = predicted number of total crashes for the facility or project of interest during the same period for which crashes were observed;

$N_{\text{predicted } rmj}$ = Predicted number of multiple-vehicle nondriveway collisions for roadway segments of type j, j = 1..., 5, during the same period for which crashes were observed;

$N_{\text{predicted } rsj}$ = Predicted number of single-vehicle collisions for roadway segments of type j, during the same period for which crashes were observed;

$N_{\text{predicted } rdj}$ = Predicted number of multiple-vehicle driveway-related collisions for roadway segments of type j, during the same period for which crashes were observed;

$N_{\text{predicted } imj}$ = Predicted number of multiple-vehicle collisions for intersections of type j, j = 1..., 4, during the same period for which crashes were observed;

$N_{\text{predicted } isj}$ = Predicted number of single-vehicle collisions for intersections of type j, during the same period for which crashes were observed;

$N_{\text{observed (total)}}$ = Observed number of total crashes for the facility or project of interest;

$N_{\text{observed } rmj}$ = Observed number of multiple-vehicle nondriveway collisions for roadway segments of type j;

$N_{\text{observed } rsj}$ = Observed number of single-vehicle collisions for roadway segments of type j;

$N_{\text{observed } rdj}$ = Observed number of driveway-related collisions for roadway segments of type j;

$N_{\text{observed } imj}$ = Observed number of multiple-vehicle collisions for intersections of type j;

$N_{\text{observed } isj}$ = Observed number of single-vehicle collisions for intersections of type j;

$N_{\text{predicted } w_0}$ = Predicted number of total crashes during the same period for which crashes were observed under the assumption that crash frequencies for different roadway elements are statistically independent ($\rho = 0$);

k_{rmj} = Overdispersion parameter for multiple-vehicle nondriveway collisions for roadway segments of type j;

k_{rsj} = Overdispersion parameter for single-vehicle collisions for roadway segments of type j;

k_{rdj} = Overdispersion parameter for driveway-related collisions for roadway segments of type j;

k_{imj} = Overdispersion parameter for multiple-vehicle collisions for intersections of type j;

k_{isj} = Overdispersion parameter for single-vehicle collisions for intersections of type j;

$N_{\text{predicted } w1}$ = Predicted number of total crashes under the assumption that crash frequencies for different roadway elements are perfectly correlated (ρ = 1);

w_0 = weight placed on predicted crash frequency under the assumption that crash frequencies for different roadway elements are statistically independent ($r = 0$);

w_1 = weight placed on predicted crash frequency under the assumption that crash frequencies for different roadway elements are perfectly correlated ($r = 1$);

N_0 = expected crash frequency based on the assumption that different roadway elements are statistically independent ($r = 0$);

N_1 = expected crash frequency based on the assumption that different roadway elements are perfectly correlated ($r = 1$); and

$N_{\text{expected}/comb}$ = expected average crash frequency of combined sites including two or more roadway segments or intersections.

All of the crash terms for roadway segments and intersections presented in Equations A-6 through A-9 are used for analysis of urban and suburban arterials (Chapter 12). The predictive models for rural two-lane, two-way roads and multilane highways (Chapters 10 and 11) are based on the site type and not on the collision type. Therefore, only one of the predicted crash terms for roadway segments ($N_{\text{predicted } rmj}$, $N_{\text{predicted } rsj}$, $N_{\text{predicted } rdj}$), one of the predicted crash terms for intersections ($N_{\text{predicted } imj}$, $N_{\text{predicted } isj}$), one of the observed crash terms for roadway segments ($N_{\text{observed } rmj}$, $N_{\text{observed } rsj}$, $N_{\text{observed } rdj}$), and one of the observed crash terms for intersections ($N_{\text{observed } imj}$, $N_{\text{observed } isj}$) is used. For rural two-lane, two-way roads and multilane highways, it is recommended that the multiple-vehicle collision terms (with subscripts *rmj* and *imj*) be used to represent total crashes; the remaining unneeded terms can be set to zero.

Chapters 10, 11, and 12 each present worksheets that can be used to apply the project-level EB Method as presented in this section.

The value of $N_{\text{expected}/comb}$ from Equation A-14 represents the expected average crash frequency for the same time period represented by the predicted and observed crash frequencies. The estimate of expected average crash frequency of combined sites for fatal-and-injury and property-damage-only crashes is calculated by multiplying the proportion of predicted average crash frequency by severity level ($N_{\text{predicted}(FI)}/N_{\text{predicted(total)}}$ and $N_{\text{predicted}(PDO)}/N_{\text{predicted(total)}}$) to the total expected average crash frequency of combined sites from Equation A-14. Appendix A.2.6 explains how to update $N_{\text{expected}/comb}$ to a future time period, such as the time period when a proposed future project will be implemented.

A.2.6. Adjust the Estimated Value of Expected Average Crash frequency to a Future Time Period, If Appropriate

The value of the expected average crash frequency (N_{expected}) from Equation A-4 or $N_{\text{expected}/comb}$ from Equation A-14 represents the expected average crash frequency for a given roadway segment or intersection (or project, for $N_{\text{expected}/comb}$) during the before period. To obtain an estimate of expected average crash frequency in a future period (the after period), the estimate is corrected for (1) any difference in the duration of the before and after periods; (2) any growth or decline in AADTs between the before and after periods; and (3) any changes in geometric design or traffic control features between the before and after periods that affect the values of the CMFs for the roadway segment or intersection. The expected average crash frequency for a roadway segment or intersection in the after period can be estimated as:

$$N_f = N_p \left(\frac{N_{bf}}{N_{bp}}\right)\left(\frac{CMF_{1f}}{CMF_{1p}}\right)\left(\frac{CMF_{2f}}{CMF_{2p}}\right)\cdots\left(\frac{CMF_{nf}}{CMF_{np}}\right) \tag{A-15}$$

Where:

N_f = expected average crash frequency during the future time period for which crashes are being forecast for the segment or intersection in question (i.e., the after period);

N_p = expected average crash frequency for the past time period for which observed crash history data were available (i.e., the before period);

N_{bf} = number of crashes forecast by the SPF using the future AADT data, the specified nominal values for geometric parameters, and—in the case of a roadway segment—the actual length of the segment;

N_{bp} = number of crashes forecast by the SPF using the past AADT data, the specified nominal values for geometric parameters, and—in the case of a roadway segment—the actual length of the segment;

CMF_{nf} = value of the nth CMF for the geometric conditions planned for the future (i.e., proposed) design; and

CMF_{np} = value of the nth CMF for the geometric conditions for the past (i.e., existing) design.

Because of the form of the SPFs for roadway segments, if the length of the roadway segments are not changed, the ratio N_{bf}/N_{bp} is the same as the ratio of the traffic volumes, $AADT_f/AADT_p$. However, for intersections, the ratio N_{bf}/N_{bp} is evaluated explicitly with the SPFs because the intersection SPFs incorporate separate major- and minor-road AADT terms with differing coefficients. In applying Equation A-15, the values of N_{bp}, N_{bf}, CMF_{np}, and CMF_{nf} should be based on the average AADTs during the entire before or after period, respectively.

In projects that involve roadway realignment, if only a small portion of the roadway is realigned, the ratio N_{bf}/N_{bp} should be determined so that its value reflects the change in roadway length. In projects that involve extensive roadway realignment, the EB Method may not be applicable (see discussion in Appendix A.2.1).

Equation A-15 is applied to total average crash frequency. The expected future average crash frequencies by severity level should also be determined by multiplying the expected average crash frequency from the before period for each severity level by the ratio N_f/N_p.

In the case of minor changes in roadway alignment (i.e., flattening a horizontal curve), the length of an analysis segment may change from the past to the future time period, and this would be reflected in the values of N_{bp} and N_{bf}.

Equation A-15 can also be applied in cases for which only facility- or project-level data are available for observed crash frequencies. In this situation, $N_{\text{expected}/comb}$ should be used instead of N_{expected} in the equation.