Safety Performance Function Guidance: TPF-5(255)

Draft Work Plan

Raghavan Srinivasan & Daniel Carter

University of North Carolina Highway Safety Research Center

Karin Bauer

MRIGlobal

July 26, 2012

Objectives

Develop and provide information to assist state and local agencies calibrate the SPFs in the HSM and/or develop their own SPFs for their state or region.

Task 1. Develop Work Plan Outline and Conduct Initial Kick-Off Webinar with FHWA and the TAG

The outline of the draft work plan was submitted to FHWA on July 13, 2012. The kickoff meeting was held on Monday, July 23, 2012. At the kickoff meeting, the project team gave a presentation outlining the objectives and the tasks. The project team also presented a series of scoping questions for discussion. Following is the list of questions along with the response from the pooled fund states to these questions.

Scoping Questions and Response from Pooled Fund States

- Question: Who is the audience for each report, Tasks 2 and 3? Is it State DOT personnel, Statisticians, University Researchers, or all of the above?
- Response: Many States felt that the States are the primary audience, including safety engineers, planners, district engineers, statisticians with the States, and MPO/RPO. They felt that case studies and examples are important and will be helpful to the States. Some States felt that the development of SPFs (discussed in the Task 3 report) would likely not be done by State DOT personnel, but by Universities or consultants, hence they would be the primary audience for the Task 3 report. Even if the States do not develop SPFs on their own, they will need to know what to ask the Universities and consultants. Hence, the Task 3 report needs to discuss the requirements and the framework regarding these SPFs.

- Question: How standalone does each report need to be (Task 2 vs. 3)? Is FHWA planning to publish the outcomes of Task 2 and 3 separately, or as one report?
- Response: FHWA is planning to publish two separate reports. The two reports need to be complimentary. They should cite each other as appropriate.

- Question: Should the reports discuss EB evaluations and the ways in which SPFs are developed and used in those evaluations? If so, in what manner should they be addressed?
- Response: The discussion of the use of SPFs in EB is necessary and appropriate, as long as there is a discussion of other methods apart from EB.

- Question: SPF model form and error terms: Considering that the focus is on implementing the HSM, should the guidebook focus only on the traditional log-linear negative binomial SPFs?
- Response: The primary focus should be on the traditional log-linear negative binomial SPFs. However, there should be a discussion on other models forms and modeling issues, including the use of multivariate models, and logit/probit models that model proportion of crashes. There should also be some discussion of the pros and cons of different modeling approaches.

- Question: FHWA has suggested that examples of Statements of Work that States have used for calibrating/developing SPFs could be used as Appendices. Will these be useful?
- Discussion: The States agreed that these would be useful.

Deliverables:

- Annotated Work Plan Outline – Completed July 13, 2012
- Project Kickoff Agenda – Provided by FHWA on July 13, 2012
- Kickoff meeting – Completed July 23, 2012
- Meeting minutes – Completed July 26, 2012
- Draft work plan – Completed July 26, 2012
- Revised work plan

Task 2. Develop Report for States on How to Choose Between Calibrating SPFs from the HSM and Developing Jurisdiction-Specific SPFs

This task focuses on the first decisions that a state must make regarding the issue of SPFs. The report will present the decision of whether to calibrate SPFs from the HSM and/or develop jurisdiction-specific SPFs. Each option will be discussed in terms of its relative benefits and challenges. The report will list the factors that will affect a state's decision and indicate the importance of each factor.

We believe that the decision to calibrate the SPFs from the HSM and/or develop jurisdiction-specific SPFs may depend on the intended use of the SPFs. The main intent of the prediction methodology (and the SPFs) in Part C of the HSM is to predict crashes that can be used for project-level analysis. On the other hand, one of the intended uses of the methods (with SPFs) in Part B of the HSM is network screening.

The data requirements for the SPFs in these two parts are quite different. The Part C prediction methodology includes a base model and a series of crash modification factors to adjust the prediction for conditions different from the base condition. As mentioned in the HSM, if a jurisdiction wants to estimate jurisdiction-specific SPFs, it could be done in one of two ways:

1. Develop SPFs using only data that represent base conditions
2. Develop SPFs using a broader set of conditions than the base conditions, and then substitute values for the different variables in the model that correspond to the base conditions.

Either approach will require the jurisdiction to assemble a database with sufficient number of sites with information on site characteristics that can be used to determine whether a site corresponds to the base condition.

Unlike the SPFs in Part C of the HSM, SPFs that are used for network screening typically use a smaller set of variables. For example, in SafetyAnalyst, default SPFs are provided for different facility types, and the system automatically calibrates these default SPFs using the data from a particular jurisdiction. For roadway segments and ramps, the SPF predicts crashes per mile as a function of AADT. For intersections, the SPF predicted crashes per intersection as a function of major and minor road AADT. Hence, jurisdiction-specific SPFs for network screening could conceivably be developed for different facility types using just AADT and segment length.

In summary the decision to calibrate or develop jurisdiction-specific SPFs may depend on the following factors:

- Intended Use of the SPFs (i.e., network screening versus project-level analysis). It is possible that an agency may decide to develop jurisdiction-specific SPFs for network screening but calibrate the HSM SPFs for project-level analysis. As an example, Srinivasan and Carter (2011)

estimated North Carolina specific SPFs for 16 different roadway types for network screening purposes, and also calibrated the HSM SPFs using data from North Carolina for project-level analysis.

- Available expertise. The level of analytical or statistical expertise available to the agency may affect whether they choose to calibrate or develop SPFs. Calibration is a more straight-forward process that can be done by staff with limited to no statistical experience. However, developing jurisdiction-specific SPFs requires personnel with some background in statistical modeling. Agencies will need to consider whether this expertise is available in-house or whether they need to hire a consultant (e.g., university consultants).
- Facility type. It is possible that an agency may decide to develop jurisdiction-specific SPFs for certain facility types, but calibrate the HSM SPFs for other facility types. As mentioned earlier, developing jurisdiction-specific SPFs for project-level analysis will require assembling a database with sufficient number of sites with information on site characteristics that can be used to determine whether a site corresponds to the base condition. Depending on the variables available in the State's roadway inventory files, the process of assembling a database may be easier for certain facility types than for others.
- Available sample size. The HSM specifies that 30 to 50 sites with at least 100 crashes per year would be a reasonable sample for calibrating the HSM SPFs. However, developing jurisdiction-specific SPFs would typically require a larger sample. There is no straightforward way to determine the minimum sample size for developing SPFs. This decision is usually done based on experience and judgment, objectives, and available resources. The project team will review the literature to get an assessment on the sample sizes used by different agencies who have developed jurisdiction-specific SPFs (e.g., Srinivasan and Carter, 2011; Tegge et al., 2010).
- Available roadway data. Calibrating the HSM SPFs requires detailed roadway data for calculating the various CMFs that are part of the HSM predictive process. On the other hand, developing a basic jurisdiction-specific SPF for network screening may require only AADT and basic information about roadway types, but for a larger sample of sites.

As part of this Task, the project team will also seek to get examples of experiences from States that have already made a decision to either calibrate the HSM SPFs or develop their own jurisdiction-specific SPFs. Volpe will be conducting interviews of States as part of their SPF Resource Assessment project. The following is a list of questions that we would want to ask the States:

- What factors played into your decision to calibrate an existing SPF or develop your own?
- Were you concerned primarily with network screening or project-level SPFs?
- What data requirement concerns did you have?
- How did the availability of statistical expertise play into your decision? Did you use DOT staff or outside consultants?

If an agency decides that calibrating the HSM SPFs is the most appropriate option for them, then further information on the calibration procedure will be available in NCHRP Project 20-07 (Task 332) led by Dr. Geni Bahar. The project team will work closely with Dr. Bahar to ensure that both projects are coordinated.

Deliverables

- Draft outline of Task 2 Report
- Final outline of Task 2 Report
- Executive Summary of Task 2 report
- Draft Report of Task 2

- Final Report of Task 2

Task 3. Develop a "How-To" Guidebook for States Developing Jurisdiction-Specific SPFs

This task will focus on creating a guidebook to instruct agencies on how to develop SPFs for their state or region. It will comprehensively define SPF terminology and will provide examples where appropriate. It will present the states with the knowledge of what data, expertise, tools, and other resources are required to develop jurisdiction-specific SPFs.

This How to Guide will have two parts: the first part will describe the process to evaluate the applicability and quality of an SPF developed by another State/jurisdiction. The second part will describe the steps involved in developing jurisdiction-specific SPFs.

Evaluating SPFs developed by other States

Not all States may have the resources to develop their own jurisdiction SPFs, and hence there needs to be a way for them to evaluate SPFs developed by other States to determine if they would be appropriate for use in their State. Here are possible steps in this evaluation:

1. Assess the quality of the SPF. There are several steps to do this. First, since crash data are considered count data, the SPFs should be developed using statistical methods that are appropriate for count data. The second step is to examine the goodness-of-fit (GOF) of the SPF along with the standard errors of the parameter estimates. The third step is to determine if the signs of the parameter estimates are consistent with previous research or theory, e.g., one would expect total crashes to increase with increasing AADT.
2. Assess the appropriateness of the SPF for the state/jurisdiction. The first step in this process is to calibrate the SPF with data from the state/jurisdiction. However, instead of using all the available data for the calibration, an approach would be to select several random samples of minimum size, estimate the calibration factor for each random sample, and average the calibration factors so calculated. The variability of these calibration factors would provide an assessment of the appropriateness of the SPF of the state/jurisdiction.

Developing a "How to" Guide for Developing Jurisdiction-Specific SPFs

Clearly, the format and content of the "How to" guide will have to address the type of end users: State DOT personnel, Statisticians, or University Researchers. We are assuming that the primary user is State DOT personnel, but that other users will benefit from this document as well. Following are several issues that need to be considered in developing jurisdiction-specific SPFs, and will be discussed in the Guide:

- Structure of systematic variation and residual terms. As mentioned earlier, since crash data are count data, appropriate statistical methods are needed. In addition, crash data tend to be overdispersed, i.e., the variance is typically larger than the mean. To account for these conditions, the negative binomial model has become the most popular form for modeling crash data. In fact, all the SPFs in Part C of the HSM and in SafetyAnalyst are negative binomial models. Recently, other models including Poisson-lognormal models, zero-inflated Poisson or

Poisson-gamma models, and Conway-Maxwell models, have been used with some success (see Lord and Mannering, 2010). However, since the intent of this document is to facilitate the implementation of the HSM, the main focus will be on estimation using negative binomial regression models.

- Model form. Traditionally, negative binomial models are fit assuming a log-linear relationship between crash frequency and the independent variables. This is more due to convenience of estimation since this form allows the models to be estimated using a procedure called generalized linear models (McCullagh and Nelder, 1989). Most statistical packages assume this model form while estimating negative binomial regression models. In fact, all the SPFs in Part C of the HSM and in SafetyAnalyst are log-linear. However, Hauer (2004) and others (e.g., Kononov et al., 2011) have argued that this may not always be the most appropriate form. Hauer (2004) has argued for the use of a combination of additive and multiplicative terms and other functional forms that may be more appropriate to reflect the true relationship between crash frequency and the independent variables. However, again, since the intent of this document is to facilitate the implementation of the HSM, the main focus will be on estimation using negative binomial regression models with a log-linear form.
- Functional form of independent variables and interaction effects. Hauer (2004) argues that most traditional functional forms of independent variables constraints the relationships to be monotonic and does not allow them to have turning points (i.e., peaks and valleys). In addition, most models do not allow for interaction effects between independent variables. However, in some cases, interaction effects may be important, e.g., combination of sudden change in vertical alignment and cross section may have a larger effect than the sum of these effects (Elvik, 2011).
- Goodness of fit. A number of goodness-of-fit parameters have been used by previous researchers (e.g., Pseudo R-square, Freeman-Tukey R-square, Deviance). These GOFs typically provide information about the overall goodness of fit. Hauer and Bamfo (1997) have proposed the use of Cumulative Residual (CURE) plots that provide information about how well the model fits in different ranges of the independent variables, and hence, further insight into the functional form of the relationship between the independent variables and crash frequency.
- Statistical tools. Researchers use a variety of statistical tools (statistical software) to estimate SPFs. The common ones include SAS, STATA, and GENSTAT. These software packages might be cost-prohibitive for an agency. Hence, it is necessary to talk about other tools such as R, an open source programming language, and Microsoft Excel. The intent is to show examples on how SPFs can be estimated using some of the more available tools.

In addition to discussing these issues with examples, the guidebook will also briefly discuss other modeling issues that are outside the scope of the HSM, but still relevant to SPF development in the future, for example:

- Multivariate modeling and analysis to deal with injury severity and crash type correlation
- Logit/probit models that are used to predict proportion of different crash types and severities
- Temporal and spatial correlation
- Bayesian versus maximum likelihood estimation

When these topics are covered, there will also be some discussion of the pros and cons of different modeling approaches. Addressing these issues in detail will require a trained statistician with experience in the analysis of crash data. Nevertheless, these will be discussed briefly with guidance on where users can find further information.

Deliverables:

- Draft outline
- Final outline
- Executive Summary
- Draft Report
- Final Report

Task 4. Project Management

This task will focus on the daily and monthly management of the process and coordination with FHWA and contractors from other related efforts. Activities will include:

- Facilitation of monthly conference calls
- Submittal of monthly progress reports
- Coordination with Volpe team (Lee Biernbaum) and 20-7 calibration efforts (Geni Bahar)

Deliverables:

- Monthly Progress Reports
- Quarterly Newsletter for HSM Implementation Pooled Fund Study

References

Elvik, R. (2011), Assessing Causality in Multivariate Accident Models, *Accident Analysis and Prevention*, Vol. 43, pp. 253-264.

Hauer, E. (2004), Statistical Road Safety Modeling, *Transportation Research Record: Journal of the Transportation Research Board*, No. 1897, Transportation Research Board of the National Academies, Washington, DC.

Hauer, E., and J. Bamfo (1997), Two tools for finding what function links the dependent variable to the explanatory variables, *Paper presented 5-7 November 1997 at the ICTCT Conference*, Lund.

Kononov, J., Lyon, C., and Allery, B.K. (2011), Relationship of Flow, Speed, and Density of Urban Freeways to Functional Form of a Safety Performance Function, *Transportation Research Record: Journal of the Transportation Research Board*, No. 2236 (pp. 11-19), Transportation Research Board of the National Academies, Washington, DC.

Lord, D. and Mannering, F. (2010), The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives, *Transportation Research Part A*, Vol. 44, pp. 291-305.

McCullagh, P. and Nelder, J.A. (1989), *Generalized Linear Models*, CRC Press.

Srinivasan, R. and D. Carter (2011), *Development of Safety Performance Functions for North Carolina*, Report FHWA/NC/2010-09, North Carolina Department of Transportation.

Tegge, R.A., J.-H. Jo, and Y. Ouyang (2010), *Development and Application of Safety Performance Functions for Illinois*, Report FHWA-ICT-10-066, Illinois Department of Transportation.

Schedule of Deliverables

The delivery schedule below outlines the estimated time to complete each task. All draft deliverables shall assume a 4-week FHWA review and comment period

Task	Deliverable	Due Date
Task 1. Initial Kick-off webinar with designated lead agency	Annotated work plan outline	July 13, 2012
	Project Kickoff Agenda	July 13, 2012
	Kickoff meeting	July 23, 2012
	Meeting minutes	July 26, 2012
	Draft work plan	July 26, 2012
	Revised work plan	August 10, 2012
Task 2. Develop Report for States Deciding Whether to proceed with HSM Calibration or Safety Performance Function (SPF) Development.	Draft outline	August 24, 2012
	Final outline	September 7, 2012
	Executive Summary	September 21, 2012
	Draft Report	December 28, 2012
	Final Report	January 31, 2013
Task 3. Develop a SPF Guidebook	Draft outline	August 24, 2012
	Final outline	September 7, 2012
	Executive Summary	October 5, 2012
	Draft Report	March 29, 2013
	Final Report	April 26, 2013
Task 4. Project Management	Monthly Progress Report	by 15th of following month
	Quarterly Newsletter for HSM Implementation Pooled Fund Study	by 15th of first month of next quarter