Comparing the Highway Safety Manual's Safety Performance Functions with Jurisdiction-Specific Functions for Intersections in Regina

Jason Young (Corresponding Author)
Graduate Research Assistant
Department of Civil and Geological Engineering
University of Saskatchewan
57 Campus Drive
Saskatoon, SK S7N 5A9
Phone: (306) 966-1314
Fax: (306) 966-5427
Email: jey606@mail.usask.ca

Peter Y. Park, Ph.D., P.Eng.
Assistant Professor
Department of Civil and Geological Engineering
University of Saskatchewan
57 Campus Drive
Saskatoon, SK S7N 5A9
Phone: (306) 966-1314
Fax: (306) 966-5427
Email: peter.park@usask.ca

Paper prepared for presentation at the
Implementation of Safety Prediction Methods in the Processes and Procedures of a Transportation Agency Session of the
2012 Annual Conference of the
Transportation Association of Canada
Fredericton, New Brunswick

ABSTRACT

The first edition of the Highway Safety Manual (HSM) includes a number of safety performance functions (SPF), which can be used to identify collision-prone locations on a roadway network. The HSM recommends that these SPFs be calibrated in order to more accurately reflect a specific jurisdiction's unique roadway characteristics, driver behavior, etc. Another alternative is the creation of jurisdiction-specific SPFs. For this study, negative binomial regression was used to develop a set of models using five years of collision data (2005-2009) from the city of Regina, Saskatchewan. Three intersection categories were investigated: 3-leg unsignalized, 4-leg unsignalized, and 3 and 4-leg signalized. The SPFs provided in the HSM were also calibrated using this data, and a set of calibration factors were produced. Statistical goodness of fit (GOF) tests were performed in order to determine the best-fitting SPFs for the study region. In addition to the statistical tests, CURE (cumulative residual) plots were utilized to perform two comparisons: between candidate jurisdiction-specific model forms, and between the jurisdiction specific SPFs and the HSM's SPFs (both calibrated and un-calibrated). It was found that the jurisdiction-specific SPFs provided the best fit to the data used in this study, and would be the best SPFs for predicting collisions at 3 and 4-leg intersections in the City of Regina.

INTRODUCTION

The primary goal of road safety engineering is reducing the frequency and severity of collisions on a roadway network. To achieve this goal, many safety authorities in North America currently use the six-step roadway safety management process (RSMP) presented in the Highway Safety Manual (HSM) (1). The RSMP consists of: 1) network screening, 2) diagnosis, 3) selecting countermeasures, 4) economic appraisal, 5) prioritizing projects, and 6) evaluating safety effectiveness. The first step of the RSMP, network screening, aims to identify locations that will most benefit from a program of safety improvements. These locations are also known as blackspots or hotspots. In recent decades, the question of the best way to identify hotspots has been the subject of great interest and discussion among safety engineers around the world. Numerous hotspot identification measures have been introduced, including 13 measures reported in the HSM. Examples include both conventional (e.g., collision frequency, collision rate) and advanced measures (e.g., EPDO average collision frequency with EB adjustments).

The advanced measures use mathematical formulas known as safety performance functions (SPFs). SPFs are designed to explain the statistical association between the location's attributes (e.g., traffic volume, roadway length) and collision frequencies (2,3) and are used to predict the number of collisions for target locations. We note that the HSM provides various SPFs for different classifications of roadways; for instance, the HSM's SPFs for rural two-lane highways are different from that of rural multilane highways – they are also different from that of urban/suburban arterials.

Over the past two decades, there have been intensive discussions among road safety engineers in regard to the development of better-fitting SPFs using a wide range of advanced statistical methods (4). Currently, the negative binomial model is regarded as the standard method in engineering practices to develop SPFs, and is indeed the method that has been used for the development of the HSM's SPFs. Although SPFs could use a variety of inputs (e.g., shoulder lane width, presence of lighting), the most common and important inputs are traffic volume – such as average annual daily traffic (AADT) – and road segment length. These two inputs are known as exposures. Equation 1 presents an example of an SPF provided in the HSM that is designed to predict the number of single-vehicle collisions on 4-leg signalized intersections on urban arterials.

$$N_{bisv} = exp^{(-10.21+0.68 \times \ln(AADT_{maj}) + \ln(AADT_{min}))} \qquad \text{[Equation 1]}$$

where:

$AADT_{maj}$ = average annual daily traffic volume for the major road; and
$AADT_{min}$ = average annual daily traffic volume for the minor road.

Equation 1 was developed using collision and traffic volume data collected from U.S. states such as Minnesota, North Carolina, and Michigan (3). Considering the empirical nature of SPFs, there is a high possibility that the HSM's given (or base) SPFs may not perform properly for roadway networks in other area. The HSM acknowledges this as an issue and recommends that the base SPFs in HSM be calibrated (i.e., updated or modified) using local collision data to predict collisions more accurately for target locations in a study network. (The HSM's guidance on calibration method is summarized in the "Methodology" section of this paper.)

Very recently, several studies have been undertaken to investigate the impact of calibration of HSM's SPFs on the performance of collision prediction for local roadway networks (5,6,7,8). While these studies state in general that calibrated HSM models show better performance (measured by model fit) than uncalibrated HSM models, a common challenge encountered by the researchers includes the large amount of data (e.g., roadway characteristics, traffic volumes, and multiple years of collision information) required for the HSM's calibration procedure.

An alternative approach to measuring the safety performance of a jurisdiction's roadway network is through the development of jurisdiction-specific SPFs, which are modeled using a specific jurisdiction's traffic volume, collision history, and roadway environment characteristics. By using jurisdiction-specific SPFs, road safety engineers may be able to better capture the unique collision characteristics of their study network in a particular region (3).

Many recent studies have described the development of jurisdiction-specific SPFs for rural roadway segments (9,10,11,12), urban roadway segments (11,13), rural intersections (14,15), and urban intersections (14,15,16). In a study conducted by Garber et al, jurisdiction-specific SPFs for intersections exhibited a better model fit than the HSM SPFs, however, this comparison was made to uncalibrated HSM SPFs (15). Lu et al. found that jurisdiction-specific SPFs fit their collision data better than calibrated HSM SPFs (13). Sacchi et al., in their investigation of the transferability of the HSM to Italy's roadways, used cumulative residual (CURE) plots to assess the validity of jurisdiction-specific models, but didn't perform a similar comparison to the HSM SPFs (9). These findings in previous studies beg the following question: will the given SPFs in HSM perform well in terms of predicting the number of collisions in roadway networks in Canadian municipalities (e.g., Regina, Saskatchewan)?

The objectives for this research are as follows:

1) Develop jurisdiction-specific SPFs using Regina's collision, traffic volume, and roadway environment data;
2) Calibrate the base SPFs provided in the HSM for intersections in Regina, Saskatchewan; and
3) Compare the performance of the two SPFs (i.e., calibrated HSM's SPFs vs. the jurisdiction-specific SPFs).

A detailed discussion of the methodology used to meet these objectives is given in the next section. The analysis results are then provided, followed by a discussion of these results. The study concludes with a summary of the important findings, and recommendations for future work.

METHODOLOGY

Description of Study Data

For this study, road network information, traffic counts, and collision databases were obtained for the city of Regina, Saskatchewan. The most recent five years of data (2005 to 2009) were used to develop and validate the jurisdiction-specific SPFs.

ArcGIS shapefiles for the city's road segments and intersections were provided by the City of Regina. These spatially-referenced files contained unique location identifiers (referred to as UGRIDS) for each segment and intersection, as well as street names, segment lengths, and functional classifications. A separate file containing information on intersection controls was also supplied; this file was joined to the existing database in order to allow further disaggregation.

Traffic volume information was also supplied. Since traffic counts are not performed every year, for every segment, the estimation procedure described in the HSM was followed:

- If data is missing for one or more years between known traffic counts, linear interpolation is used to estimate the missing years' data; and
- If data is missing before or after the first or last recorded counts, this value is assumed to remain constant for the missing portion of the study period (1).

Collision information was supplied by Saskatchewan Government Insurance (SGI). Collision records were separated into three databases – by collision, vehicle, and occupant – which could be linked to the street network and traffic volume databases through the UGRID code. Since the focus of this study was vehicle-to-vehicle collisions, SQL queries were performed on the collision databases to separate pedestrian-related collisions (using the occupant database) and bicycle-related collisions (using the vehicle database).

In recent years, a great deal of SPF research has been conducted on road segments – as such, the focus of this study was decided to be on intersections. The existing data set provided a wealth of intersection-related collision data; in the five-year period from 2005 to 2009, 62% of collisions in Regina occurred at intersections.

After the integrated databases were examined for errors (e.g. null traffic volumes) and inconsistencies, the intersections were divided into four categories: 3-leg signalized, 3-leg unsignalized, 4-leg signalized, and 4-leg unsignalized. As a low number of 3-leg signalized intersections (28) existed in the study region, these intersections were aggregated with the 4-leg signalized intersections for the purposes of this study. Table 1 shows the number of intersections for each category that were used for the development of jurisdiction-specific SPFs.

Each of these categories was divided into two subsections by random selection: estimation (70% of the data) and validation (30% of the data).

Calibration of the Highway Safety Manuals' Safety Performance Functions

To represent local roadway conditions, driver characteristics, and geometric design more accurately, the HSM recommends the calibration of the provided SPFs. The calibration procedure described in the HSM specifies two criteria in the selection of a dataset: 1) a minimum of 30 to 50 sites, and 2) a minimum of 100 collisions (1).

Calibration factors (e.g., C_i for intersections) are obtained by calculating the ratio of total number of observed collisions to the total number of predicted collisions obtained from the base SPFs. Therefore, the nominal value of the calibration factor (which would be the case if the observed collision frequency equaled the predicted crash frequency) is 1. When there are more collisions observed than are predicted by the default SPF, the calibration factor will be greater than 1. When there are fewer collisions observed than are predicted by the default SPF, the calibration factor will be less than 1. Once the required calibration factors are calculated, they can be applied to the default SPFs provided in the HSM. A table showing the calibration factors for the City of Regina is given in the "Results" section of this paper.

The following sections describe the estimation and validation methods that were used for the development of jurisdiction-specific SPFs.

Estimation of Local Safety Performance Functions

R-Language was used to develop jurisdiction-specific SPFs, using the negative binomial model, for each of three categories: total collisions, fatal/injury (FI) collisions, and property damage only (PDO) collisions. For each of these categories, four candidate model forms were developed, which were then compared using the methods described in the following section. Table 2 shows the initial forms that were created.

A number of methods were used to select the most appropriate form of SPFs for each data category, including investigation of each model's p-values, Akaike information criterion (AIC), Bayesian information criterion (BIC), cumulative residual (CURE) plots, and overdispersion parameters.

The AIC is a relative measure of a statistical model's goodness of fit (GOF), first proposed by Hirotsugu Akaike (17). It can be used to help determine the best-fitting model from several candidates – lower values are preferable – but provides no absolute information about a model's performance. The AIC is calculated using Equation 2 below.

$$AIC = 2 \times k - 2 \times \log(L) \qquad \text{[Equation 2]}$$

where:

k = number of parameters in the statistical model; and
L = maximized value of the likelihood function for the estimated model.

The BIC, first proposed by Gideon Schwarz, is similar to the AIC, but includes a term to quantify the number of data points in the model (18). The BIC is calculated using Equation 3 below.

$$BIC = k \times \log(n) - 2 \times \log(L) \qquad \text{[Equation 3]}$$

where:

n = number of data points in the study dataset.

A graphical method known as CURE plot is used to compare different forms of SPFs (19). In this method, each model's cumulative residuals (defined as the difference between the observed and predicted values for each site) are plotted horizontally. Residuals below zero indicate a model that over-estimates the predicted number of collision, while residuals above zero are indicative of a model that under-represents the predicted number of collisions. Additionally, good-fitting models can be identified by cumulative residuals that lie between the boundaries of two standard deviations, both above and below zero, which represent a 95% confidence limit. Figure 1 shows an example of a CURE plot, with the red line representing the cumulative residuals, and the blue and green lines showing the positive and negative two-standard deviation boundaries, respectively.

In addition, the overdispersion factor (K) – which is a parameter of negative binomial model – can be used to compare different form of SPFs. This parameter is a measure of heterogeneity (i.e., variance) in the collision data. The larger the overdispersion parameter is, the more the collision data vary; as such, a model with smaller overdispersion parameters (i.e., a smaller value of K) is preferred to a model with larger overdispersion parameters (20).

Validation of Local Safety Performance Functions

Washington et al. reported a series of statistical tests that can be used to validate models, and recommended that multiple GOF tests be assessed before making a decision in regards to a particular model's validity (20).

The mean square error (MSE) is applied to the estimation data. This test is a measure of the error associated with the model; smaller values are preferable to larger values. Equation 4 shows the formulation for this GOF test.

$$MSE = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n-p} \qquad \text{[Equation 4]}$$

where:

Y = number of predicted collisions;
$\hat{Y}$ = number of observed collisions;
n = data sample size; and
p = number of parameters in the statistical model.

The mean prediction bias (MPB) is applied to the validation data. This test provides a measure of the magnitude and direction of the average model bias – the smaller the value, the better the model is at predicting observed data. The form of the MPB is given in Equation 5.

$$MPB = \frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)}{n} \qquad \text{[Equation 5]}$$

The mean absolute deviation (MAD) is applied to the validation data. This test gives a measure of the average magnitude of variability of prediction, and unlike the MPB, values can only be positive. Smaller values are preferable to larger values. The form of the MAD is given in Equation 6.

$$MAD = \frac{\sum_{i=1}^{n} |\hat{Y}_i - Y_i|}{n} \qquad \text{[Equation 6]}$$

The mean squared prediction error (MSPE) is applied to the validation data. This test can be compared to the MSE to check for over-fitting of models to estimation data (MSPE > MSE) or under-fitting models (MSPE < MSE). Similar MSPE and MSE values are desired; the form is shown in Equation 7.

$$MSPE = \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{n} \qquad \text{[Equation 7]}$$

The Freeman Tukey R-Squared (R^2_{FT}) is applied to both estimation and validation data. This test, proposed by Fridstrom et al., measures GOF. Larger values indicate a better fit (21). Equations 8 through 11 show the equations required to compute this value (22).

$${R_{FT}}^2 = \frac{\sum_{i=1}^{n} (f_i - \bar{f})^2 - \sum_{i=1}^{n} {\hat{e}_i}^2}{\sum_{i=1}^{n} (f_i - \bar{f})^2} \qquad \text{[Equation 8]}$$

$$f_i = \sqrt{Y_i} + \sqrt{Y_i + 1} \qquad \text{[Equation 9]}$$

$$\bar{f} = \frac{\sum_{i=1}^{n} \left(\sqrt{Y_i} + \sqrt{Y_i + 1}\right)}{n} \qquad \text{[Equation 10]}$$

$$\hat{e}_i = \sqrt{Y_i} + \sqrt{Y_i + 1} - \sqrt{4 \times \hat{Y}_i + 1} \qquad \text{[Equation 11]}$$

The results of these statistical tests are shown in the "Analysis Results" section of this paper.

ANALYSIS RESULTS

Objective 1 - Develop jurisdiction-specific SPFs using Regina's collision, traffic volume, and roadway environment data.

The negative binomial model was used to develop SPFs for the three categories under investigation, using the estimation subset (70% of the data). A number of potential model forms were analyzed – using the comparison methods described in the "Methodology" section of this paper – in order to determine the best-fitting model.

Table 3 gives the selected model forms for each of the three intersection categories; Table 4 gives the coefficients for each of the SPFs. As Table 4 shows, all of the model parameters were found to be statistically significant at the 0.1% level, except for the minor AADT term in the FI 4-leg signalized intersection category, which had a p-value of 0.234. However, other comparison measures, such as the AIC, BIC and examination of CURE plots, showed that this model form fit the data better than the other alternatives.

CURE plots for the 4-leg signalized models, showing total, FI, and PDO collisions, as a function of major AADT, are shown in Figure 2. This figure shows that the residuals for the 4-leg unsignalized models fall within the 95% confidence interval (designated by the blue +2σ and green -2σ boundaries) for the entire range of major AADT. The models fluctuate around zero on the horizontal axis, which is indicative of a good-fitting model (19).

Statistical GOF tests were also performed on the validation subset (30% of the data). Table 5 shows the results of these tests. These results can be compared in order to assess the transferability of the developed models – which were developed using the estimation dataset – to the validation dataset. The MSE and MSPE values are similar in magnitude, particularly for the unsignalized models, indicating a high level of transferability. Additionally, since the MSPE values are lower in most cases, over-fitting of the regression models is unlikely. The Freeman Tukey R-Squared values are higher for the estimation datasets than they are for the validation datasets. This may be a result of the relatively low number of sites included in the validation data; for example, the 4-leg unsignalized validation dataset included only 36 sites (30% of the 121 4-leg sites investigated for this study).

Objective 2 - Calibrate the base SPFs provided in the HSM for intersections in Regina, Saskatchewan.

Using the observed collision numbers for the five most recent years of City of Regina data, the HSM SPFs were calibrated according to the procedure described in the "Methodology" section of this paper. Calibration factors were calculated for each of the five years (2005 to 2009), as well as the five-year average. The results of the calibration are shown in Table 6. As the table shows, the majority of the average calibration factors are greater than one, indicating that the HSM SPFs under-predict the number of collisions in these categories. For 3 and 4-leg signalized intersections, the average calibration factors for total collisions and PDO collisions are 2.25 and 2.79, respectively. This can be interpreted to say that the HSM SPFs predict 56% fewer total collisions and 64% fewer PDO collisions than actually occurred in Regina during the five-year study period.

FI collisions for both 3 and 4-leg unsignalized intersections have calibration factors less than one (0.65 and 0.66, respectively), indicating that the HSM SPFs over-predict collisions in these categories. For all three intersection categories, the FI calibration factor was consistently lower than the total and PDO calibration factors. For example, for 3-leg unsignalized intersections, the HSM SPFs predict 32% fewer total collisions and 45% fewer PDO collisions than actually occurred, but 54% greater FI collisions than actually occurred.

There are several possible reasons for these severity-based differences between models. For one, it's possible that collisions in the City of Regina are simply less severe due to local roadway conditions, climatic effects, and/or driver behavior. The quality of local emergency medical services (EMS) also has as effect on the number of fatalities and injuries that occur. As well, differences in reporting criteria between the City of Regina and the jurisdictions from which the HSM's urban/suburban intersection SPFs were developed (Minnesota and North Carolina) may play a role in the models' varying results (3). For example, the Minnesota Department of Public Safety defines a collision as "a collision that involves a motor vehicle in transport on a public traffic-way in Minnesota and results in injury, death, or at least $1,000 in property damage" (23). For the study period of 2005-2009, $1,000 was the PDO threshold for collisions records to be added to the province's collision database, but this amount has recently been increased to $5,000 (24). Changes to reporting thresholds in Saskatchewan in 1984, 1993, and 2010 have resulted in large decreases in PDO collisions in the province's collision database, and this highlights an important consideration. Since reporting thresholds, and other collision-related criteria, are embedded in the data, jurisdiction-specific SPFs (or calibration factors for HSM SPFs) are required in order to capture these regional collision-reporting characteristics.

A statistical comparison between the HSM's SPFs (both calibrated and uncalibrated) and the jurisdiction-specific SPFs is provided in the following section.

Objective 3 - Compare the performance of the two SPFs (i.e., calibrated HSM's SPFs vs. the jurisdiction-specific SPFs).

The jurisdiction-specific SPFs were compared to the HSM SPFs (both calibrated and uncalibrated) using a number of techniques, including visual plots, statistical GOF tests, and CURE plots. As the 3 and 4-leg signalized intersection category consisted primarily of 4-leg intersection data, this category was compared to the HSM SPFs for 4-leg signalized intersections.

Graphs showing the observed collisions (five-year averages), as well as jurisdiction-specific SPFS, HSM SPFs, and calibrated HSM SPFs, are shown in Figure 3. For 3-leg unsignalized intersections (shown in the left column) and 4-leg unsignalized intersections (middle column), the uncalibrated HSM SPFs (dark blue) under-predict total and PDO collisions, and over-predict FI collisions. For 3 and 4-leg signalized intersections (right column), the uncalibrated HSM SPFs consistently under-predict collisions in all categories. The calibrated HSM SPFs (light blue) tend to correspond more closely to the jurisdiction-specific SPFs (orange), up until approximately 15,000 major AADT (3-leg unsignalized), 8,000 major AADT (4-leg unsignalized) and 20,000 major AADT (3 and 4-leg signalized), at which point the jurisdiction-specific SPFs exhibit a decreasing trend. Though these graphs give a general idea of the models' fit to the observed collisions, it must be remembered that not all of the dots represent a single location (i.e., if two or more locations have similar major AADT and observed number of collisions, they will be represented by a single dot).

Statistical GOF tests were also utilized to analyze the models. It is important, when assessing a set of models using GOF tests, that several statistical tests be performed in order to identify the best-fitting model (20). Therefore, four of the statistical tests that were used to select the best-fitting functional form were also used to compare the jurisdiction-specific SPFs to the HSM SPFs. These tests include the MSPE, MPB, MAD, and Freeman Tukey R-Squared. Table 7 shows the results of this comparison. The highlighted boxes identify the models that performed best for each statistical test. In general, the jurisdiction-specific SPFs exhibited the lowest values for MSPE (other than 4-leg unsignalized FI collisions) and MAD (other than 3-leg unsignalized total and PDO collisions). The calibrated HSM SPFs, however, exhibited the lowest MPB values in most cases. The jurisdiction-specific SPFs exhibited the highest Freeman Tukey R-Squared results for all intersections and severities. The results from these GOF tests highlight the importance of performing multiple statistical tests, as outcome from one test may not necessarily reflect the majority of the results. Based on the results shown in Table 7, the jurisdiction-specific SPFs show the best-fitting results, based on their fit to Regina's collision data.

CURE plots were also developed for all of the intersection categories. Figure 4 shows CURE plots, as a function of major AADT, for 3-leg unsignalized, 4-leg unsignalized and 3 and 4-leg signalized intersections, based on total collisions. As the CURE plots show, the jurisdiction-specific SPFs' cumulative residuals fall within the 95% confidence interval (designated by the blue +2σ and green -2σ boundaries) for the entire range of major AADT, for both intersection categories. For these SPFs, the cumulative residuals fluctuate above and below the horizontal axis, which is representative of a good-fitting model (19). For all three intersection categories, the uncalibrated HSM SPFs' cumulative residuals show significant deviation from the 95% confidence interval; this deviation becomes more pronounced as the major AADT increases. The calibrated HSM SPFs' proximity to the 95% confidence intervals is similar to the jurisdiction-specific SPFs, which is to be expected, as these SPFs were calibrated to the same observed collision data. However, both of the calibrated HSM SPFs shown in Figure 4 do exhibit deviation from the 95% confidence intervals for major AADTs greater than 34,000 (3-leg unsignalized), 18,000 (4-leg unsignalized intersections), and 33,000 (3 and 4-leg signalized intersections).

In addition to the cumulative residuals, information about the variation inherent in each model can be determined by examining the size of the 95% confidence interval for each model. Figure 5 shows these intervals for total collisions in the 3-leg unsignalized category, as a function of major AADT. As the figure shows, the jurisdiction-specific SPF (shown in red) results in the narrowest 95% confidence interval over the entire range of major AADT values. The calibrated and uncalibrated HSM SPFs exhibit wider confidence intervals, which become pronounced at approximately 5,000 major AADT. The uncalibrated HSM SPF shows the widest confidence interval up to approximately 20,000 major AADT, at which point the calibrated HSM SPF's confidence interval expands past it. For the range between 20,000 and 35,000 major AADT, the calibrated HSM SPF's confidence interval is approximately double that of the jurisdiction-specific SPF, which indicates a higher degree of variance in the calibrated HSM SPF. As the 95% confidence interval is a function of the cumulative residuals for each model, an examination of these boundaries can also be used as a preliminary tool to determine which model best fits the observed data. Based on this comparison measure, the jurisdiction-specific SPFs displayed the best fit.

CONCLUSIONS & RECOMMENDATIONS

The research described in this paper investigated the performance of three collision prediction models for urban intersections: those given in the HSM (both uncalibrated and calibrated), as well as jurisdiction-specific SPFs that were developed using collision data from the City of Regina. A variety of methods were used to determine the best-fitting models, including statistical GOF tests and CURE plots. Key findings include the following:

- The SPFs provided in the HSM typically under-predicted collisions in the City of Regina, except for FI collisions at 3-leg and 4-leg unsignalized intersections, which exhibited over-prediction;
- CURE plots proved to be a useful tool for both choosing candidate model forms, as well as comparing the developed jurisdiction-specific SPFs to the SPFs given in the HSM;

- The 95% confidence intervals generated with the CURE plots provided an additional source of information in regards to a particular model's fit to the observed data; and
- Based on statistical GOF tests and CURE plots, the jurisdiction-specific SPFs and calibrated HSM SPFs were determined to exhibit a better fit than the uncalibrated HSM SPFs for all categories of intersections. Based on the comparison measures used in this study, the jurisdiction-specific SPFs showed a slightly better fit than the calibrated HSM SPFs.

The findings presented in this paper represent the first step in comparing the performance of jurisdiction-specific SPFs to the SPFs given in the HSM. Recommended future work includes the comparison of SPFs developed for roadway segments using data from the City of Regina. Additionally, a network screening using the empirical Bayes (EB) method (which combines predicted and observed collision information in order to identify collision-prone intersections or roadway segments) is recommended; a comparison of rankings would provide a real-world, quantitative means of comparing the developed SPFs to the HSM SPFs.

ACKNOWLEDGEMENTS

The authors wish to thank the Saskatchewan Centre of Excellence for Transportation and Infrastructure (SCETI), Saskatchewan Government Insurance (SGI), and the City of Regina for their financial and other support for this study.

REFERENCES

1. AASHTO. 2010. *Highway Safety Manual.* First Edition. American Association of State Highway and Transportation Officials.

2. Harwood, D., Council, F., Hauer, E., Hughes, W., Vogt, A. (2000). *Prediction of the Expected Safety Performance of Rural Two-Lane Highways.* Federal Highway Administration Report FHWA-RD-99-207. FHWA, U.S. Department of Transportation, McLean, Virginia.

3. Harwood, D., Bauer, K., Richard, K., Gilmore, D., Graham, J., Potts, I., Torbic, D., Hauer, E. (2007). *Methodology to Predict the Safety Performance of Urban and Suburban Arterials.* National Cooperative Highway Research Program Project 17-26. Available from http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_w129p1&2.pdf. [Accessed 2 March 2012].

4. Lord, D., Mannering, F., (2010). *The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives,* Transportation Research, Part A, Vol. 44 (5), pp. 291-305.

5. Xie, F., Gladhill, K., Dixon, K., & Monsere, C. (2011). *Calibrating the Highway Safety Manual Predictive Models for Oregon State Highways,* Proceedings from the 2011 TRB 90th Annual Meeting. Washington, D.C. January 2011.

6. Sun, X., Magri, D., Shirazi, H., & Gillella, S. (2011). *Application of the Highway Safety Manual: Louisiana Experience with Rural Multilane Highways*, Proceedings from the 2011 TRB 90th Annual Meeting. Washington, D.C. January 2011.

7. Banihashemi, M. (2012). Highway Safety Manual, *Calibration Dataset Sensitivity Analysis,* Proceedings from the 2012 TRB 91st Annual Meeting. Washington, D.C. January 2012.

8. Lubliner, H. & Schrock, S. (2012*). Calibration of the Highway Safety Manual Prediction Method for Rural Kansas Highways,* Proceedings from the 2012 TRB 91st Annual Meeting. Washington, D.C. January 2012.

9. Sacchi, E., Persaud, B. & Bassani, M. (2012). *Assessing International Transferability of the Highway Safety Manual Crash Prediction Algorithm and its Components,* Proceedings from the 2012 TRB 91st Annual Meeting. Washington, D.C. January 2012.

10. Brimley, B., Saito, M., & Schultz, G. (2012). *Calibration of the Highway Safety Manual Safety Performance Function and Development of New Models for Rural Two-Lane Two-Way Highways,* Proceedings from the 2012 TRB 91st Annual Meeting. Washington, D.C. January 2012.

11. Alluri, P. & Ogle, J. (2012). *Effects of state-specific SPFs, AADT estimations, and overdispersion parameters on crash predictions using SafetyAnalyst,* Proceedings from the 2012 TRB 91st Annual Meeting. Washington, D.C. January 2012.

12. Bornheimer, C., Schrock, S., Wang, M. & Lubliner, H. (2012). *Developing a Regional Safety Performance Function for Rural Two-Lane Highways,* Proceedings from the 2012 TRB 91st Annual Meeting. Washington, D.C. January 2012.

13. Lu, J., Gan, A., Haleem, K., Alluri, P. & Liu, K. (2012). *Comparing Locally-Calibrated and SafetyAnalyst-Default Safety Performance Functions for Florida's Urban Freeways,* Proceedings from the 2012 TRB 91st Annual Meeting. Washington, D.C. January 2012.

14. Tegge, R., Jo, J. & Ouyang, Y. (2010). *Development and Application of Safety Performance Functions for Illinois,* Research Report for Illinois Center for Transportation.

15. Garber, N., Rivera, G. & Lim, I. (2011). *Safety Performance Functions for Intersections in Virginia*, Proceedings from the 2012 TRB 90th Annual Meeting. Washington, D.C. January 2011.

16. Lyon, C., Haq, A., Persaud, B. & Kodama, S. (2005). *Safety Performance Functions for Signalized Intersections in Large Urban Areas: Development and Application to Evaluation of Left-Turn Priority Treatment,* Transportation Research Record: Journal of the Transportation Research Board, Volume 1908, pp. 165-171.

17. Burnham, Kenneth, P.; Anderson, David R.. (2002). *Model Selection and Multi-model Inference: A Practical Information-Theoretic Approach.* Springer.

18. Schwarz, G. (1978). *Estimating the Dimension of a Model,* Annals of Statistics. Volume 6, Number 2, pp. 461-464.

19. Hauer, E. & Bamfo, J. (1997). *Two tools for finding what function links the dependent variable to the explanatory variables,* Proceedings of ICTCT 97: International Cooperation on Theories and Concepts in Traffic Safety (pp. 1–7), Lund, Sweden.

20. Washington, S., Persaud, B., Lyon, C., & Oh, J. (2005). *Validation of Accident Models for Intersections.* FHWA Report, Publication No. FHWA-RD-03-037.

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

22. Hamidi, A., Fontaine, M., & Demetsky, M. (2010). *A Planning-Level Methodology for Identifying High-Crash Sections of Virginia's Primary System.* Virginia Transportation Research Council Research Report.

23. Minnesota Department of Public Safety. (2011). Minnesota Motor Vehicle Crash Facts. Available from https://dps.mn.gov/divisions/ots/educational-materials/Documents/CRASH-FACTS-2010.pdf [Accessed 13 March 2012].

24. Saskatchewan Government Insurance. (2012). *2010 Saskatchewan Traffic Accident Facts.* Available from http://www.sgi.sk.ca/pdf/tais/TAIS_2010_Annual_Report.pdf [Accessed 13 March 2012].

Table 1: Number of intersections for different categories used in the development of SPFs.

Category	No. of Intersections
3-Leg Unsignalized	123
4-Leg Unsignalized	121
3 & 4-Leg Signalized	143

Table 2: Initial candidate model forms.

Model No.	Model Form
1	$N = exp^{\beta_0} \cdot \left(\frac{AADT_{min}}{1000}\right)^{\beta_1} \cdot \left(\frac{AADT_{maj}}{1000}\right)^{\beta_2}$
2	$N = exp^{\beta_0} \cdot \left(\frac{AADT_{min}}{1000}\right)^{\beta_1} \cdot \left(\frac{AADT_{maj}}{1000}\right)^{\beta_2} \cdot exp^{\left(\beta_3 \cdot \frac{AADT_{min}}{1000}\right)}$
3	$N = exp^{\beta_0} \cdot \left(\frac{AADT_{min}}{1000}\right)^{\beta_1} \cdot \left(\frac{AADT_{maj}}{1000}\right)^{\beta_2} \cdot exp^{\left(\beta_3 \cdot \frac{AADT_{min}}{1000}\right)}$
4	$N = exp^{\beta_0} \cdot \left(\frac{AADT_{min}}{1000}\right)^{\beta_1} \cdot \left(\frac{AADT_{maj}}{1000}\right)^{\beta_2} \cdot exp^{\left(\beta_3 \cdot \frac{AADT_{min}}{1000}\right)} \cdot exp^{\left(\beta_4 \cdot \frac{AADT_{maj}}{1000}\right)}$

Table 3: Model forms for City of Regina intersections.

Category	Safety Performance Function
3-Leg Unsignalized	$N_{3U} = exp^{\beta_0} \cdot \left(\frac{AADT_{min}}{1000}\right)^{\beta_1} \cdot \left(\frac{AADT_{maj}}{1000}\right)^{\beta_2}$
4-Leg Unsignalized	$N_{4U} = exp^{\beta_0} \cdot \left(\frac{AADT_{min}}{1000}\right)^{\beta_1} \cdot \left(\frac{AADT_{maj}}{1000}\right)^{\beta_2}$
3 & 4-Leg Signalized	$N_{S} = exp^{\beta_0} \cdot \left(\frac{AADT_{min}}{1000}\right)^{\beta_1} \cdot \left(\frac{AADT_{maj}}{1000}\right)^{\beta_2}$

Table 4: Model coefficients for City of Regina intersections.

Category	Severity	Regression Coefficients						Overdispersion Parameter
		β_0	*p*-Value	β_1	*p*-Value	β_2	*p*-Value	
3-Leg Unsignalized	Total	-1.717	<0.001	0.826	<0.001	0.648	<0.001	0.396
	FI	-3.786	<0.001	0.858	<0.001	0.709	<0.001	0.485
	PDO	-1.832	<0.001	0.825	<0.001	0.626	<0.001	0.417
4-Leg Unsignalized	Total	-1.007	<0.001	0.732	<0.001	0.605	<0.001	0.281
	FI	-3.195	<0.001	0.241	0.234	0.959	<0.001	0.487
	PDO	-1.138	<0.001	0.799	<0.001	0.559	<0.001	0.257
3 & 4-Leg Signalized	Total	-0.929	<0.001	0.665	<0.001	0.684	<0.001	0.111
	FI	-3.011	<0.001	0.503	<0.001	0.945	<0.001	0.109
	PDO	-1.067	<0.001	0.704	<0.001	0.630	<0.001	0.119

Table 5: Goodness of fit tests for the developed models.

Category	Severity	Estimation Data (70%)		Validation Data (30%)			
		MSE	R^2_{FT}	MSPE	MPB	MAD	R^2_{FT}
3-Leg Unsignalized	Total	55.64	58%	55.09	-0.51	5.24	39%
	FI	3.10	33%	2.26	0.06	1.07	30%
	PDO	38.84	57%	42.94	-0.58	4.45	37%
4-Leg Unsignalized	Total	122.47	53%	120.05	-0.36	7.30	10%
	FI	4.12	30%	8.34	-0.65	1.84	4%
	PDO	92.94	53%	82.07	0.24	6.14	10%
3 & 4-Leg Signalized	Total	604.84	77%	502.48	0.56	16.76	73%
	FI	38.36	67%	32.72	0.00	4.65	64%
	PDO	409.85	76%	329.52	0.63	13.25	71%

Table 6: Calibration factors for the City of Regina.

Category	Severity	Calibration Factor					
		2005	2006	2007	2008	2009	Avg
3-Leg Unsignalized	Total	1.67	1.28	1.37	1.46	1.54	1.47
	FI	0.60	0.62	0.72	0.64	0.68	0.65
	PDO	2.14	1.57	1.65	1.80	1.87	1.81
4-Leg Unsignalized	Total	1.57	1.47	1.43	1.65	1.97	1.63
	FI	0.61	0.69	0.58	0.79	0.63	0.66
	PDO	2.12	1.92	1.91	2.14	2.73	2.17
3 & 4-Leg Signalized	Total	2.27	1.90	2.16	2.32	2.56	2.25
	FI	1.68	1.18	1.44	1.32	1.26	1.37
	PDO	2.67	2.34	2.63	2.94	3.33	2.79

Table 7: Statistical comparison between three sets of SPFs.

Category	Severity	Jurisdiction-Specific SPFs				HSM SPFs				Calibrated HSM SPFs			
		MSPE	**MPB**	**MAD**	**R^2_{FT}**	**MSPE**	**MPB**	**MAD**	**R^2_{FT}**	**MSPE**	**MPB**	**MAD**	**R^2_{FT}**
3-Leg Unsignalized	Total	54.12	-0.19	4.91	52%	75.84	-2.87	4.99	46%	60.62	0.00	4.90	51%
	FI	2.77	0.03	1.13	32%	3.89	0.71	1.40	8%	3.03	0.00	1.16	31%
	PDO	39.16	-0.23	4.20	51%	61.38	-3.44	4.61	36%	43.74	-0.69	4.03	50%
4-Leg Unsignalized	Total	118.71	-0.09	6.51	42%	161.57	-5.44	7.55	24%	119.79	0.00	6.59	40%
	FI	5.27	-0.21	1.57	22%	6.79	1.07	2.00	-8%	5.23	0.00	1.62	20%
	PDO	87.41	0.11	5.49	43%	148.25	-6.51	7.36	-4%	89.60	0.00	5.74	40%
3 & 4-Leg Signalized	Total	561.37	0.21	16.60	76%	2556.69	-36.32	36.83	-20%	753.57	0.00	20.50	66%
	FI	35.86	0.00	4.46	66%	57.78	-3.43	5.21	49%	39.71	0.00	4.69	61%
	PDO	377.09	0.28	13.49	75%	2060.92	-33.86	34.07	-61%	513.32	0.00	17.17	64%

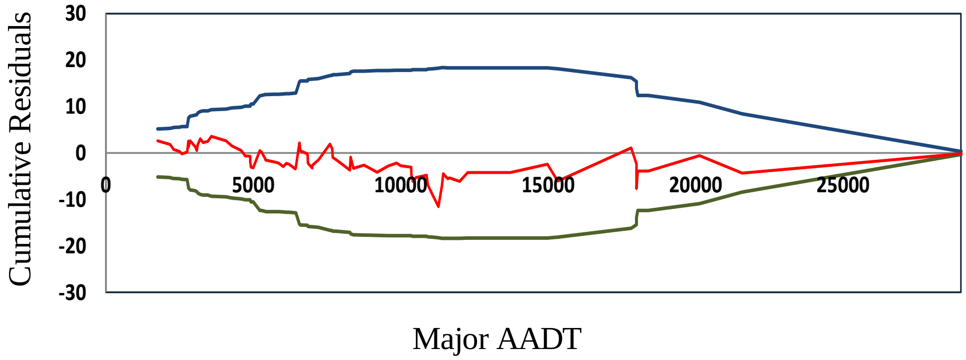

Figure 1: Sample CURE plot of a collision prediction model.

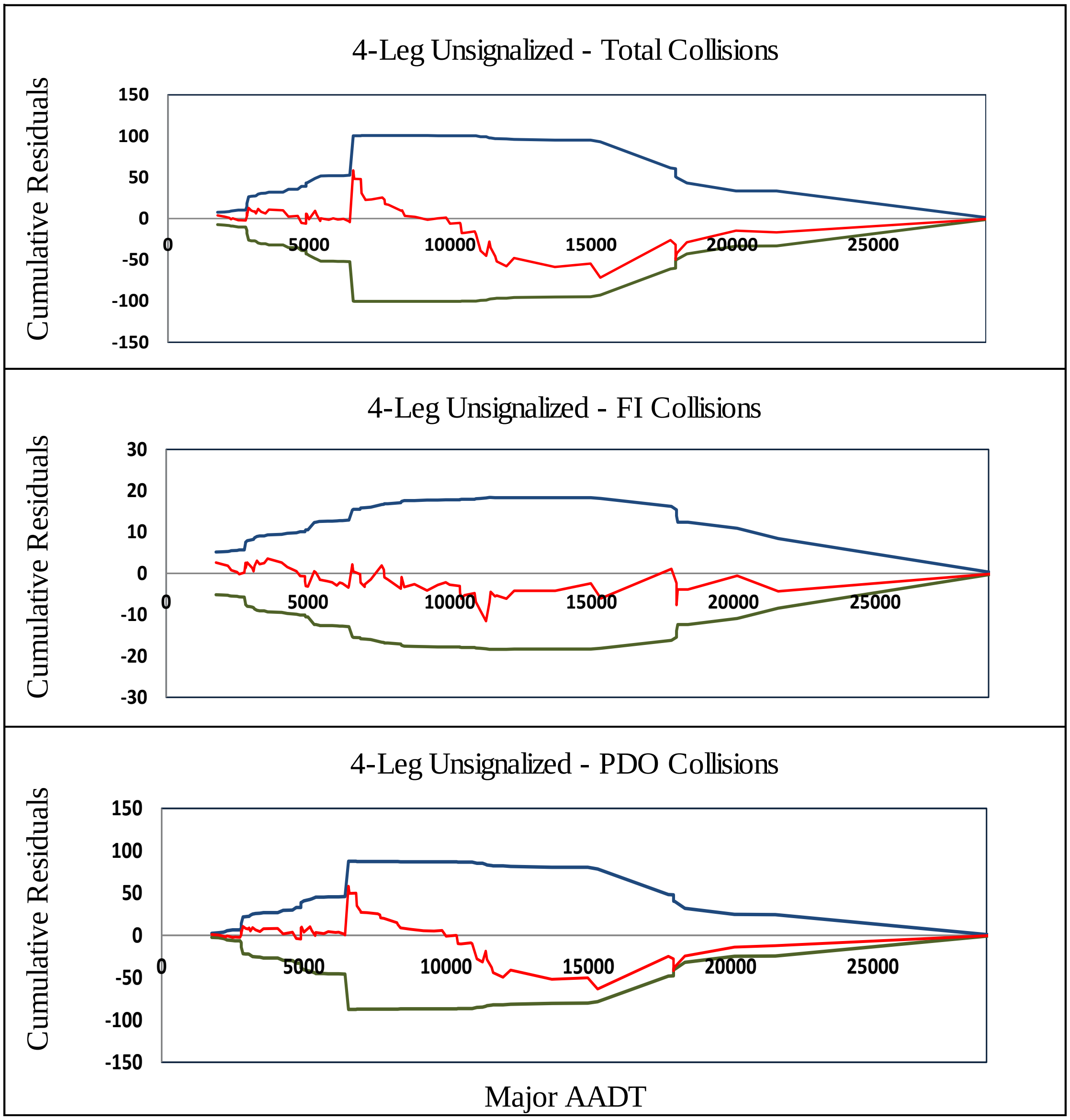

Figure 2: CURE plots for 4-leg unsignalized intersection models.

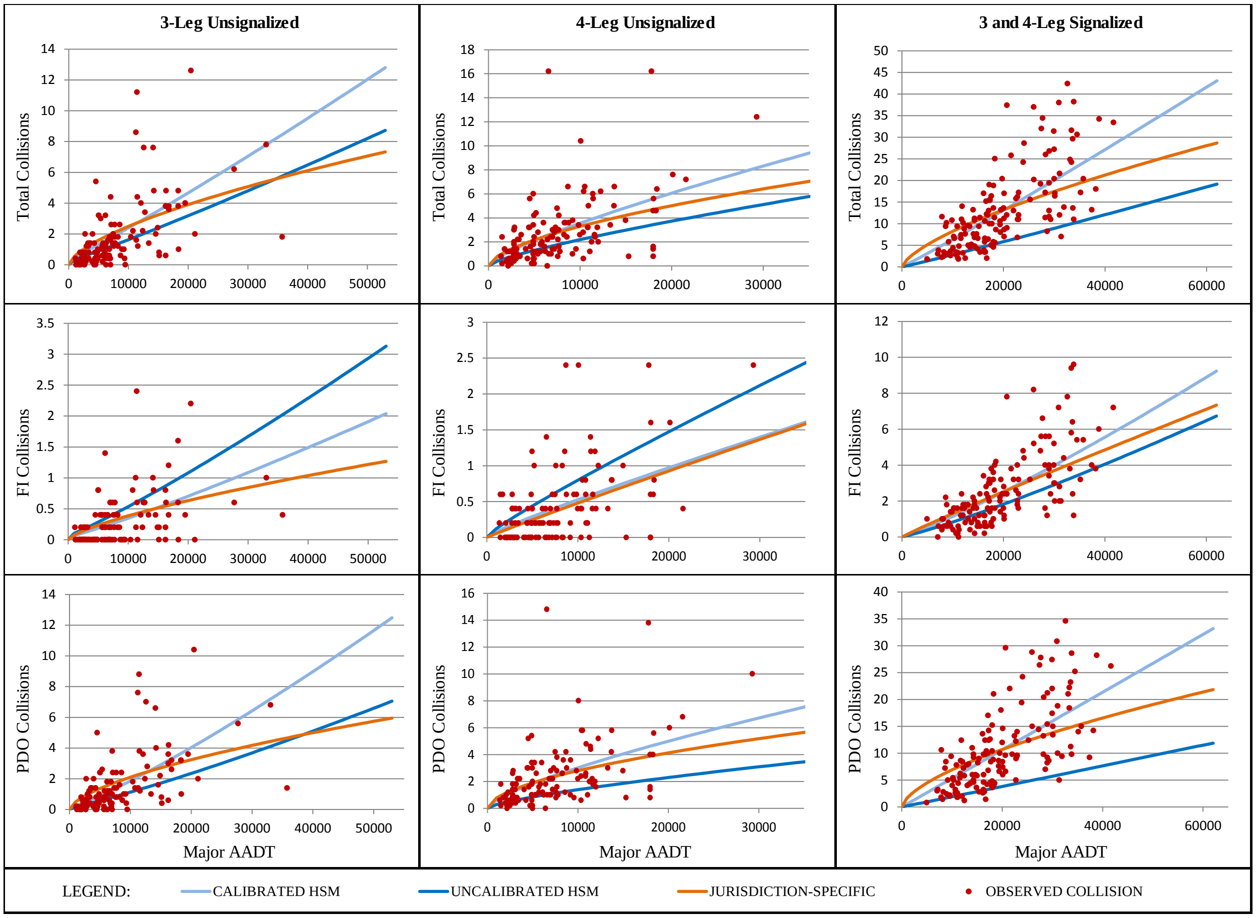

Figure 3: Total, FI, and PDO observed collisions and SPFs for
3-leg unsignalized, 4-leg unsignalized and 3 and 4-leg signalized intersections.

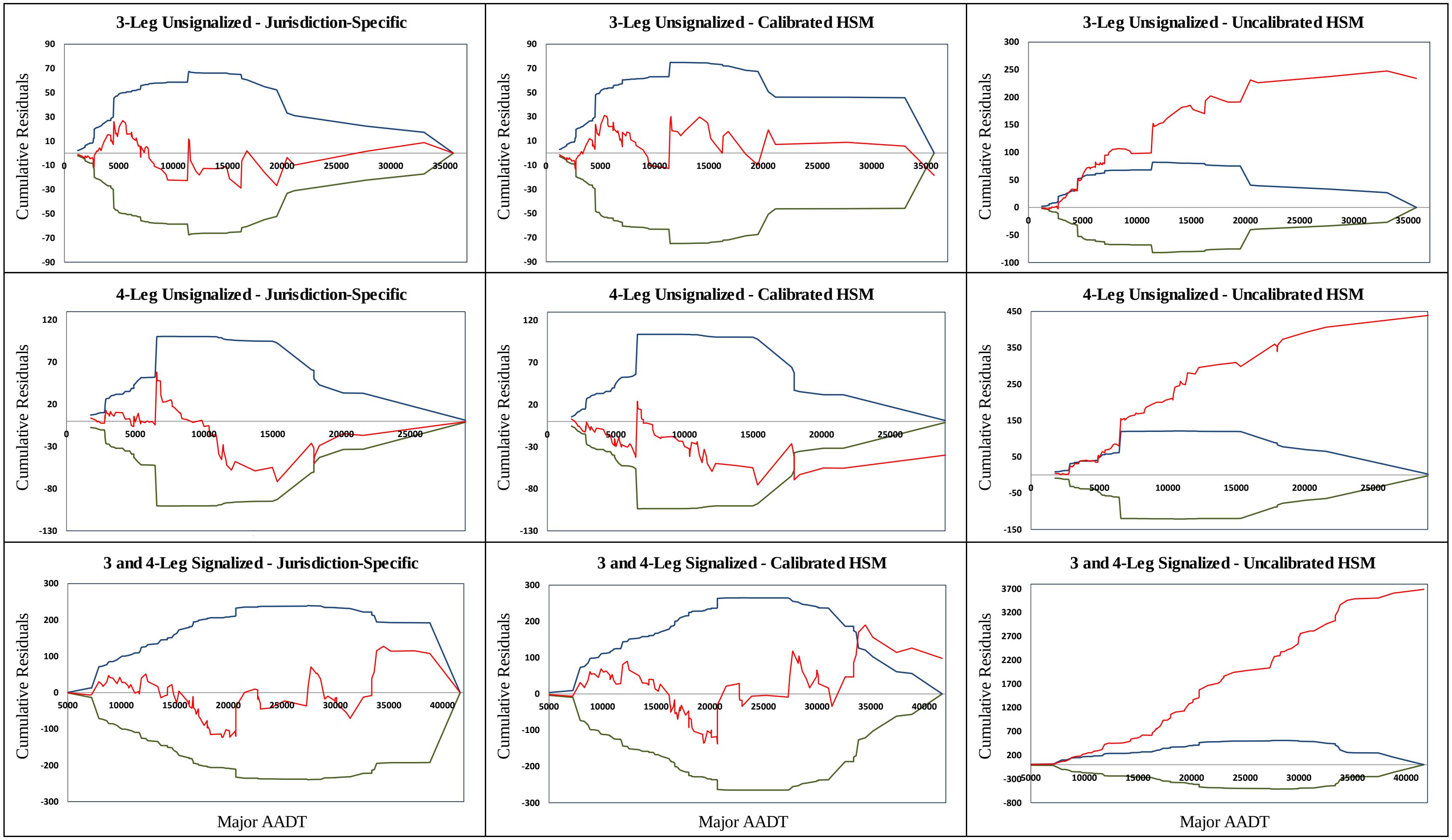

Figure 4: CURE plots as a function of major AADT for 3-leg unsignalized, 4-leg unsignalized and 3 and 4-leg signalized intersections.

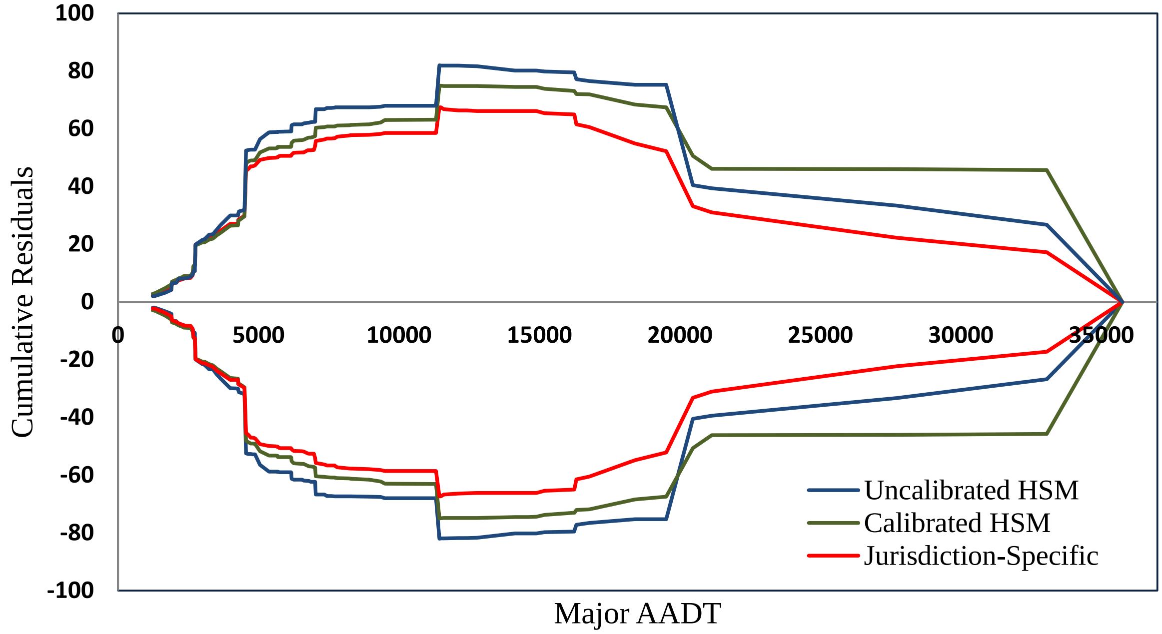

Figure 5: 95% confidence intervals for 3-leg unsignalized intersections.