Develop Calibration Factors for Crash Prediction Models for Rural Two-Lane Roadways in Illinois

Michael Williamson\textsuperscript{a,*}, Huaguo Zhou\textsuperscript{a}

\textsuperscript{a} Department of Civil Engineering, Southern Illinois University Edwardsville Department of Civil Engineering Campus Box 1800
Edwardsville, IL 62026, USA

Abstract

This paper documents the development of calibration factors for crash prediction models in the new Highway Safety Manual (HSM) for rural two-lane roadways in Illinois. The crash prediction models (so called Safety Performance Functions (SPF)) in the HSM were developed using data from multiple states, therefore the models must be calibrated to account for local factors, such as weather, roadway conditions, and drivers’ characteristics. In this study, two calibration factors were developed for two different SPFs to give a better prediction of crash frequencies on rural two lane roadways in Illinois.

1. Introduction

In recent years, the safety of the traveling public on roadways in the United States has been an area of interest for transportation engineers. Current focus is on the reduction of fatal crashes by any possible means. The release of the National Highway Safety Manual (HSM) in 2010 improved the quantitative approach to highway safety management. The HSM method can be used to predict the number of expected crashes on a roadway, through consideration of the cross-sectional geometrics and traffic

\* Corresponding author. Tel.: 217-343-7512; Fax: 618-650-2555.
\textit{E-mail address:} micwill@siue.edu

© 2012 Published by Elsevier B.V. Selection and/or peer review under responsibility of Beijing Jiaotong University [BJU], Systems Engineering Society of China (SESC) Open access under CC BY-NC-ND license.
volumes. This paper will focus on rural two-lane two-way roadways, specifically the calibration of the HSM predictive method to this type of roadway in the State of Illinois.

The HSM predictive method includes the use of three separate variables to model a roadway’s safety: a safety performance function (SPF), crash modification factors (CMFs), and a calibration factor (C). The SPF was statistically derived from data obtained in several states and is intended to be used at the national level. Other SPFs developed by state or local agencies could be used in place of the national SPF and may better predict crashes in that region. Since SPFs are developed to predict crashes at a base condition, CMFs are needed when the cross-section of a roadway is altered from the base condition. Many CMFs are available and should be applied to accurately predict the crashes on roadways. The calibration factor (C) is used to adjust the national models to local conditions.

To date the HSM has not been calibrated for specific state or geographic locations including the State of Illinois. The SPF needs to be calibrated to adjust for driver population, roadway conditions, traffic volumes, weather conditions, and other regional specific factors. The objective of this study is to calibrate the SPF for rural two-lane roadways for the State of Illinois and compare the calibrated SPF with an SPF developed with Illinois crash data. The calibration of the HSM predictive model will assist transportation engineers in the selection of safety improvements that will reduce the number of fatalities experienced on rural roadways in Illinois. Use of the model is another step towards the national strategy on highway safety of moving toward zero fatalities.

Rural two lane roadways have the highest fatal rate of all road types accounting for 57% of national and 66% of Illinois fatalities, which can be linked to geometric designs and response times (Kalokata, 1994; FHWA 2010). This means that persons living or traveling in rural areas are more likely to be involved in a fatal crash than those living or traveling in urban areas (FHWA, 2010). The calibrated HSM models can be used to predict crashes at any roadway segments, identifying contributing factors that may cause high crash rates.

2. Literature review

An extensive literature review was conducted on prediction methods and factors contributing to crashes on rural roadways. It was determined that the geometric factors affecting the safety of a roadway including: lane width, shoulder width and type, clear zone distance, and road side hazards (Kalokata, 1994; FHWA 2010; FHWA 2009; Bahar 2009).

2.1. Crash prediction approaches

Traffic crashes are seen as rare events, so their prediction can be somewhat challenging. The factors causing crashes can be broken into the following groups: vehicle, human, highway, and context. In addition to the groups there are two ways to approach a safety analysis. The first is non-quantitative, which focuses on policy compliance, assessment using adjunct principles or guidelines (i.e. design consistency, driver work load, positive guidance, and other human factors), possibly within the context of a safety audit. The other approach is quantitative and uses crash reduction measures, statistical models, simulation surrogates, and driving simulators. Quantitative methods offer the most reliable approach to improve the safety of roadways (Pfefer, 2004). In Illinois the current method of practice is to develop an SPF for each road type and then determine the potential for safety improvement using a method of weighting (FHWA, 2010). The HSM is a quantitative method and more accurate than any current method of crash prediction.
2.2. Data collection methods

A traditional approach to highway safety improvements includes an extensive inspection of a roadway, review of crash reports, and meeting with stakeholders and citizens to gain local insight (Cafiso, 2004). The data collected using the traditional method is similar to the data collection process using the HSM predictive method.

Crash recording thresholds vary across the United States, for this reason the HSM must be adjusted to fit each State’s threshold. For example, Oregon, California, and Washington only report crashes at values above $1500, $750, and $700, respectively. The calibration factor (C) used in the predictive method is said to take care of this issue (Xie, 2011). Like most states Illinois has its own regulation for recording crashes.

3. Methodology

3.1. Calibration procedure

Before applying the HSM predictive method to state roadways, calibration factors needed to be developed to mitigate the impacts of weather conditions, driver populations, animal populations, and terrain. The following five steps had to be followed to develop an accurate calibration factor. The first step was to randomly select segments to be used in the study. The second step involved obtaining the site specific geometrics and cross-sectional data of each segment. The third step was to use the HSM predictive model to predict the crash frequency for each of the selected segments followed by the fourth step, comparing the predicted vs. observed crash frequencies. The final step adjusted the model to local conditions by calculating the calibration factor (C). The model could then be applied to regional specific segments and used to determine the number of expected crashes under various conditions. The model is useful when determining design improvements that will reduce the number of crashes along rural roadways.

The SFP for rural segments is a function of the AADT and segment length. It was used to predict the crash frequency for the base conditions. The base condition is defined as the condition that the SPF was developed for. When applying SPF to the real-world conditions, crash modification factors (CMF) and calibration factor need to use to account for the non-base conditions. CMFs including: lane width, shoulder width and type, grade, driveway density, roadside hazard rating (RHR), calibration factor, and countermeasures, are applied where applicable. The calibration factor is then applied to the predictive model to estimate a regions crash frequency more accurately. The HSM predictive model can be used to predict the expected crash frequency in future years, and can be seen in equation one (FHWA, 2010).

\[
N_{\text{Predicted}} = N_{\text{spf}} \times C \times (CMF_1 \times CMF_2 \times \ldots \times CMF_i)
\]

where,

- \(N_{\text{Predicted}}\) = Predicted average crash frequency for a specific site (Crashes/Year);
- \(N_{\text{spf}}\) = Expected crash frequency from SPF for base conditions (Crashes/Year);
- \(CMF_i\) = Crash modification factors consider effects of altering from base condition;
- \(C\) = Calibration factor.

3.2. SPF

The SPF used in the HSM method was developed at the national level, and is used to predict crashes at the selected segments. Formula two is the HSM SPF for rural two-lane highway with two independent
variables: AADT and segment length.

\[ N_{\text{spf(HSM)}} = AADT \times L \times 365 \times 10^{-6} \times e^{(-0.312)} \]  

where,

AAPT = Average Annual Daily Traffic  
L = Segment Length

3.3. Illinois SPF

Recently a SPF was developed specifically for Illinois and could also be used to predict crash frequency for rural two-lane highways. Formula three is the SPF developed for Illinois in 2010 (Robert A. Tegge, 2010). The format of this SPF is similar to the HSM SPF and uses AADT and segment length as the independent variables.

\[ N_{\text{spf(IL)}} = L \times e^a \times (AADT)^b \]  

where \( a = -4.435 \) \( b = 0.525 \)

3.4. Crash data collection

Traffic crashes are traditionally classified as either roadway segment or intersection related, with the recording officer using judgment to determine which. The HSM now classifies crashes as intersection related, if within 250 feet of an intersection. For segments the general rule is that all areas outside of the scope of an intersection are segment related. For the purpose of this study all crashes located at the distance of 250 feet or more from intersections are classified as segment crashes (AASHTO, 2010).

3.5. Calibration factor

The HSM (AASHTO, 2010) method compares the predicted crashes to the observed crashes at randomly selected segments and uses these values to determine a calibration factor (C) for an area or region. To determine the calibration factor for a specific area or region 30 to 50 segments must be selected, to adequately represent the physical and safety conditions of the region. The selected sites must include 100 crashes, to meet the minimum sample size requirement (AASHTO, 2010). The calibration factor is equal to the total number of observed crashes divided by the total number of predicted crashes. Formula four is used to determine the calibration factor for local conditions.

\[ C = \frac{\sum \text{Observed}}{\sum \text{Predicted}} \]  

3.6. Crash recording threshold

The crash recording threshold in Illinois prior to 2009 was $500. Effective January 1\textsuperscript{st} 2009 the crash recording threshold was increased to $1,500. This change has a significant impact on the number of recordable crashes. Fatal and injury crashes are still recorded regardless of the threshold involved, however Property Damage Only (PDO) crashes have seen a significant decrease from previous years. The recordable crashes for the year 2007 to 2009 were found to be 422,778; 408,258; and 292,106, respectively (IDOT, 2009). The drastic decrease in the number of crashes can be directly linked to the change in crash recording threshold.
When calibrating the HSM model multiple years of crash data will be used to ensure an even distribution and eliminate any bias caused by yearly winter weather. To use the available data for calibration in this study the recordable crashes needed to be adjusted to account for the change in the recording threshold to accurately predict crashes in future years.

An average number of crashes per year from 2005 to 2008 was taken and then compared to 2009. The average recordable PDO crashes for 2005 to 2008 were found to be 309,971, greater than the 2009 crashes of 202,105. This called for a reduction in recordable crashes of 34.8 percent for years prior to 2009. This resulted in a reduction of 8.8 percent for all crashes in years prior to 2009. Future thresholds must be checked when the data is available to see if the crash recording threshold remains at $1500. If the crash recording threshold is increase again, the model will need to be recalibrated to account for the reduction in recordable crashes.

4. Development of calibration factors

The purpose of this study was to calibrate the HSM model for rural two-lane two-way roadways segments in Illinois. In order to calibrate the model, several counties in Illinois were selected across the entire state. This random selection will ensure the model predicts the crashes on any rural segment accurately. Five random segments in six counties were selected for use in the calibration process. The following counties were used in this study: Edgar, Effingham, Kankakee, Madison, Sangamon, and Williamson.

To calibrate the HSM model, the predicted and the observed number of crashes at the selected segments were needed. The sum of the observed crashes had to be divided by the sum of the predicted number of crashes. This yielded the calibration factor for the study area. The three years of data analyzed from 2007 to 2009 included 165 crashes, exceeding 100 and meeting the requirement of the HSM. The three years of crashes were averaged and determined to be 34 per year, which was then compared to the one year of predicted crashes. Adjustment to the total number of observed crashes for 2007 and 2008 had to be made to account for the crash recording threshold difference. After adjustments were made to account for the recording difference, the total number of observed crashes was determined to be 31.

If the calibration factor is determined to be greater than one, the model has under predicted the crashes in the study area. Another way to think of this is that the study area experiences more crashes per year then the national average. In contrast if the calibration factor is less than one, the model has over predicted the crashes in the study area, and needs to be adjusted accordingly. Since two SPF were used in this study, two calibration factors had to be calculated. The HSM SPF predicted a total of 22.1 total crashes, while the SPF developed specifically for Illinois predicted 19.6 total crashes. The calibration factors computed for each SPF were found to be 1.40 and 1.58 for the HSM and Illinois models respectively. These factors are intended to be used to accurately predict the total crashes on a segment.

$$C_{\text{HSM}} = \frac{31.0}{22.1} = 1.40$$
$$C_{\text{Illinois}} = \frac{31.0}{19.6} = 1.58$$

The calibration factors calculated for each SPF were determined to be greater than one. This result provides evidence that the crashes on Illinois rural two-way two-lane roadway segments are higher than the national average.

5. Data analysis

5.1. Crash prediction

After determining the observed and predicted crashes an analysis was conducted to see which SPF
predicted crashes on Illinois roadways more accurately. By plotting the observed vs. predicted crashes using the un-calibrated HSM SPF, it can be seen that the model under predicts the crashes, with most of the plotted points fall below the 45 degree line. The correlation between the predicted and observed crashes does not appear to be normal. The results can be seen in Fig. 1.

\[
y = 0.581x + 0.077 \\
R^2 = 0.356
\]

Fig. 1. Comparison of the observed and predicted crashes using the HSM SPF

When the observed vs. predicted crashes are plotted using the SPF developed for Illinois, it is clear that crashes are more severely under-predicted than with the HSM SPF. The correlation between the predicted and observed crashes indicates a system error, shown in Fig. 2.

After comparing the two SPFs without calibration, it was observed that the HSM SPF better predicts the number of crashes that are likely to be seen at any given roadway segment in Illinois. After calibration the HSM predictive model was again tested to see which SPF would better predict the number of crashes on rural two-lane two-way roadways in Illinois. The observed and predicted crashes were again plotted for comparison. The HSM model was found to better predict crashes after the calibration process, which
can be seen on figure 3. Predicted crashes now fall evenly on each side of the 45 degree line in a realistic pattern, showing an even distribution of crashes. When using the calibrated Illinois SPF in the HSM model the prediction of crashes again improves, but with more points falling above the 45 degree line than with the HSM SPF. This shows the Illinois SPF better predicts the number of crashes, by providing a more even distribution. The distribution can be seen in figure 4. After calibration the SPF developed for Illinois better predicts the total number of crashes that will be experienced on rural two-lane two-way segments in Illinois.

Fig.3. Comparison of the observed and predicted crashes with the HSM SPF after calibration

Fig.4. Comparison of the observed and predicted crashes with the IL SPF after calibration

5.2. Model validation

To validate the calibration of the HSM predictive method, ten test segments were randomly selected in similar Illinois counties. The counties were in close proximity to those used in the calibration process. Three years of crash data was again used to determine the average observed yearly crashes on the selected roadway segments. The HSM method was then applied to the ten segments and the predicted number of crashes obtained using both the HSM and Illinois SPFs. The test results show a 53 and 59 percent
correlation between the observed and predicted crashes, suggesting that the calibrated models accurately predicted the observed crashes.

6. Conclusions

The calibrated HSM prediction method can now be used by transportation safety engineers to identify and make adjustments to problematic segments of rural two-lane two-way roadways in Illinois. The calibration factors developed could also be used at a regional level if the climate, driver populations, animal populations, crash reporting thresholds, and any other local factors are comparable. The factors having the greatest effect on the model were identified as AADT, driveways per mile, and shoulder width. However, the AADT remains as a governing factor within the predictive method. This paper could be used as a reference for further calibration of the HSM predictive method in other states or countries.

The largest obstacle overcome in the study was adjusting the crashes for the PDO recording threshold. Future use of the HSM in Illinois using crash data from 2009-present will not need to make this adjustment. While, the calibration constants (\( C_{\text{HSM}} = 1.40 \) and \( C_{\text{ILLINOIS}} = 1.58 \)) developed in this paper should be retained while the crash recording threshold remains at $1500, further adjustments will be needed if the threshold is again increased.

This study determined the SPF that best predicts the crashes was developed specifically for rural two-lane two-way roadways in Illinois. It is recommended that local SPFs be developed and compared to the HSM SPF when evaluating the safety of a roadway.

References

Administration, F. H. (2010). *Illinois 2010 Five Percent Report*. Washington, D.C.: Federal Highway Administration.

Administration, N. H. (2009). *Traffic Safety Facts 2009*. Washington, D.C.: U.S. Department of Transportation.

Geni Bahar, M. P. (2009). *Highway Safety Manual Knowledge Base*. Washington, D.C.: National Cooperative Highway Research Program.

IDOT, D. o. (2009). *Illinois Traffic Crash Report SR 1050*. Springfield, IL: Illinois Department of Transportation.

K. R. Kalokota, P. N. (1994). *Accident Prediction Models for Two-Lane Rural Highways*. Logan, UT: Utah State University.

Officials, A. A. (2005). *AASHTO Strategic Highway Safety Plan*. Washington, D.C.: American Association of State Highway and Transportation Officials.

Officials, A. A. (2010). *Highway Safety Manual*. AASHTO.

Pfefer, R. (2004). *Road Safety Analysis Methods and Procedures*. Washington, D. C.: TRB Task Force for the Development of a Highway Safety Manual.
Robert A and Tegge, J.-H. J. (2010). *Development and application of safety performance functions for Illinois*. Urbana, IL: Illinois Center for Transportation.

Salvatore Cafiso, G. L. (2004). *A Methodological Approach for the Safety Evaluation of Two-Lane Rural Roads with Low-Medium Traffic Volume*. Vol.1, No.11, pp. 24-26.

Xie, F. (2011). *Calibrating the Highway Safety Manual Predictive Methods for Oregon Rural State Highways*. Portland, Oregon: Oregon State University.