Examination of the Transferability of Safety Performance Functions for Developing Crash Modification Factors using Empirical Bayes Method

Jung-Han Wang
Mohamed Abdel-Aty
Jaeyoung Lee

Department of Civil Environmental Construction Engineering
University of Central Florida, Orlando
4000 Central Florida Blvd,
Orlando, FL 32816
Email: jwang@knights.ucf.edu

Nov 2015

Word Count: 4,978 words + 2 Figures + 6 Tables = 6,978 equivalent words
Submitted for possible presentation at the TRB 2016 Annual Meeting, and publication in the Transportation Research Record: Journal of the Transportation Research Board
ABSTRACT

In this study, crash modification factors (CMFs) for the effect of signalization at intersections in Florida are estimated. This paper applies the Empirical Bayes method to develop CMFs for KABCO, KABC, and rear-end crashes using several safety performance functions (SPFs) from different jurisdictions and adjusted by calibration factors. These SPFs were developed using Florida’s and Ohio’s data. Also, the SPFs suggested in the HSM were used to calculate CMFs. By developing and comparing the SPFs from different states, we concluded that it might not be suitable to apply SPFs from other states without thorough examinations. The CMF is 0.785 for KABCO using the SPF in Florida and it was significantly smaller than 1, which indicates that the signalization at intersections results in more total crashes. However, when applying the SPFs from Ohio and HSM, higher CMFs of 1.06 and 1.07 were obtained, respectively. Also they were both significantly larger than 1. It shows that the signalization brings about less total crashes. CMFs for KABC and rear-end crashes are discussed in this paper as well. The major finding of this paper is that the CMF values may be significantly different when applying SPFs developed from other states. Therefore, one would have biased CMFs if borrowing SPFs from other states without proper adjustments.

Keywords: Signalization, Crash Modification Factors, Empirical Bayes, SPF Transferability
BACKGROUND

The population of the US reached 313,914,040 people in 2012 (1). As population increase, so do the number of drivers on the road and the potential impact on safety and the operational characteristics of roadways. Roadway safety researchers treated intersections separately from roadway segments due to their unique characteristics and high exposure for crashes. According to the Fatality Analysis Reporting System (FARS) (2), 2.4 million intersection-related crashes occurred in 2007. This number consists of 40 percent of all US traffic crashes. Engineers seek solutions such as signalization to lower crashes. The Manual on Uniform Traffic Control Devices (MUTCD) chapter on traffic control signal warrant has nine warrants regarding signal installation (3). One of these warrants deal with crash experience, and thus intersection signalization could be considered as one of the major safety countermeasures. In Florida, traffic signal warrants are based on the MUTCD 2003 (4).

While both traffic researchers and engineers have concerted efforts to reduce roadway crashes. Nevertheless, in some occasions, gaps would occur between the research and actual practice. The Highway Safety Manual (HSM) (5) published by the American Association State Highway and Transportation Officials is designed to bridge these gaps. Conforming to the Highway Safety Manual Part D, most treatments have respective CMFs to predict the expected crashes (6). This prediction method is based on multiplying the specified CMF or CMFs by the previously predicted crashes based on a baseline Safety Performance Function (SPF). Quantitative values are specified in HSM so that engineers can easily measure the cost efficiency before treatments are implemented.

The purpose of this paper is to validate the transferability of SPFs using different states/sources (i.e., Ohio and HSM which developed based on data from Minnesota and North Carolina) and apply SPFs from these sources to compare the CMF values for signalization in Florida. We located the treated intersections which control type changed from two-way stop controlled to signal controlled. Using these target intersections, before-after study is conducted using Empirical Bayes (EB) method. In order to perform EB analysis, it is needed to develop SPFs and calculate the predicted crashes based on the SPFs to serve as priors. Since these treatments are located in the state of Florida, the SPFs in Florida are likely to have the highest accuracy. Under this assumption, this research compares the CMFs values among multiple SPFs from these 3 sources. If the CMFs calculated by the SPFs in HSM are close to the CMFs when using the SPFs in Florida, it would be a substantial benefit because it is not necessary to re-estimate SPFs based on local conditions for signalization.

The definition of an intersection-related area is a complicated task. Wang et al. (7) recommend that the intersection influence area varies according to the intersection type and configuration. From a practical point of view, each state in the US has its own regulation in defining intersection-related crashes. At 250 feet away from the center is a widely applied method in defining the influence area (8–13). According to HSM and Abdel-Aty et al. (6, 12), rear-end crashes are expected to increase after signalization. On the other hand, the angle and the left-turn crashes are significantly reduced after signalization (5). One previous research conducted by Cheng et al. (14) modeled the severity analyses in rear-end crashes. In fact, angle and left-turn crashes are usually
associated with higher severity levels. Therefore, examining fatality and injury crashes (KABC) is essential when estimating the safety effect of signalization.

Signalization is an important treatment from both operational and safety points of view. In the HSM (5), the CMFs of signalization are 0.56 for intersections in rural areas for total crashes and 0.95 for four-legged intersections in the urban area. On the other hand, the CMF for rear-end crashes is greater than 1 in HSM (5). Another signalization research was conducted by McGee and Persaud (15). This report states that the CMF for fatal and injury crashes is 0.86 for three-legged intersections and 0.77 for four-legged intersections. Abdel-Aty et al. (12) found that signalization effectively reduces angle, left turn, and KABC crashes, but increases rear-end crashes.

The issue of transferability of SPFs and calibrating SPFs is an important topic. Developing SPFs requires a tremendous effort of data collection and data analysis. If SPFs are transferable, researchers and engineers could skip the model development stage which is the most challenging part when developing new SPFs. Many states in the US have already developed their own calibration factors based on the SPFs provided in HSM (5). Nowadays, several studies investigate the impact of the calibration of the SPFs in HSM for local roadway networks (16–22). Besides, the calibration factors based on the SPFs in HSM are examined for places outside America such as Saudi Arabia (23) and Italy (24). These studies all pointed out that calibrated HSM models perform better (measured by model fit) than non-calibrated one.

**METHODOLOGY**

**Negative Binomial Model**

The SPF model is a negative binomial model for crash counts (25). In the model, \( y \) is the registered crash counts and the crash count is a target variable while AADT, the existence of red light running camera, operation class (rural or urban), etc. are covariates. The benefit of implementing negative binomial (NB) to model the distribution of crash frequencies is that the Poisson distribution limits the mean and the variance to be equal to \( \text{E}[y_i] = \text{VAR}[y_i] \) (13). When this equality does not hold (statistically), the data are said to be under dispersed \( \text{E}[y_i] > \text{VAR}[y_i] \) or over dispersed \( \text{E}[y_i] < \text{VAR}[y_i] \). The negative binomial model allows for over dispersion in that the mean of Poisson counts over sites \( i \) is itself gamma-distributed and has the following form as shown in equation 1:

\[
\lambda_i = \text{EXP} (\beta x_i + \varepsilon_i) = \text{EXP} (\beta x_i) \text{EXP} (\varepsilon_i) \tag{1}
\]

Where \( y \) is the registered crash counts, \( x \) is the covariate, \( \beta \) is the associated coefficient, \( \lambda \) is the expected crash count, and \( \text{EXP}(\varepsilon) \) is a gamma-distributed error term with mean 1 and variance \( \alpha^2 \).

The addition of this term allows the variance to differ from the mean as below:

\[
\text{VAR}[y_i] = \text{E}[y_i][1 + \alpha \cdot \text{E}[y_i]] = \text{E}[y_i] + \alpha \cdot \text{E}[y_i]^2 \tag{2}
\]

**Calibration Factor**

Based on the definition in HSM, “calibration factor is a factor to adjust crash frequency estimates produced from a safety prediction procedure to approximate local conditions. The factor is computed by comparing existing accident data at the state, regional, or local level to estimates
obtained from predictive models.” The calibration factor can be expressed in the equation shown in equation 3, which the calibration factor equals to the summation of the observed crash count divided by the summation of the predicted crash count using the SPF.

\[
\text{CalibrationFactor} = \frac{\sum \text{CrashCounts}_{\text{observed}}}{\sum \text{CrashCounts}_{\text{predicted}}} \quad (3)
\]

**Before and After Study using Empirical Bayes Adjustment**

The Empirical Bayes (EB) method combines the strengths of a before and after study that uses specific case-control techniques with regression methods for estimating safety. Unlike other methods, it increases the precision of estimation and it also corrects for the regression to the mean bias.

According to Hauer (26), the safety performance can be estimated through the following steps:

\[
\hat{\pi} = \gamma \cdot E\{k\} + (1 - \gamma) K 
\quad (4)
\]

Where \(\hat{\pi}\) = Expected crash count if there had been no treatment

\(E\{k\}\) = predicted crash counts based on SPF multiply by the calibration factor

\(K\) = observed crash counts before treatment

\(\gamma\) = Weight between observed crash counts and predicted crash counts

Afterward, we re-estimated \(E\{k\}\) based on SPFs and Calibration factors from different states and calculated the predicted crash counts accordingly. Then we can get the updated \(\hat{\pi}\) after substitution.

The method of the assigned weight is shown below as suggested by Hauer (26). The weight is inversely proportional to the variances of the corresponding random variables. When two estimates of unequal precision are joined, the weights \(\gamma\) and \(1 - \gamma\) that minimize the expected squared error of estimation are inversely proportional to the variance \(\varphi\) of the estimate.

\[
\gamma = \frac{1}{1 + \frac{\mu \cdot \varphi}{\varphi}} 
\quad (5)
\]

Where \(\mu\) = predicted crashes before treatment (per year)

\(Y\) = number of year(s)

\(\varphi\) = overdispersion parameter

In this term, the overdispersion parameter \(\gamma\) is different for each SPFs.

After \(\hat{\pi}\) is calculated, Gross et al. (27) suggest to use \(\hat{\pi}^*\) to adjust the value of \(\hat{\pi}\) which can be shown as:

\[
\hat{\pi}^* = \hat{\pi} \cdot \left( E\{l\} / E\{k\} \right) 
\quad (6)
\]

Where \(\hat{\pi}^*\) = Expected crash count if there had been no treatment after adjustment

\(\hat{\pi}\) = Expected crash count if there had been no treatment before adjustment

\(E\{l\}\) = predicted crashes after treatment
The CMF can be written in the form as follow:

$$\hat{\theta} = \frac{[\hat{\lambda} / \hat{\pi}]}{[1 + \text{Var}[\hat{\pi}] / \hat{\pi}^2]}$$  \hspace{1cm} (7)$$

Where $\hat{\theta} = \text{crash modification factor}$

$\lambda = \text{Observed after crash}$

When $\hat{\theta} < 1$, the treatment has a positive effect; when $\hat{\theta} > 1$ it is expected to have a negative effect on safety performance.

The variance of CMF is shown in the equation below.

$$\text{Var}(\hat{\theta}) = \frac{\hat{\theta}^2 \left[ \frac{\text{Var}(\hat{\lambda})}{\lambda^2} + \frac{\text{Var}(\hat{\pi})}{\pi^2} \right]}{[1 + \frac{\text{Var}(\hat{\pi})}{\pi^2}]^2}$$  \hspace{1cm} (8)$$

DATA DESCRIPTION

Data in Florida were collected and combined from the following five database sources: Roadway Characteristic Inventory (RCI), Crash Analysis Report (CAR), Florida Financial Management Search System, Transtat I-View, and Google Earth. The Financial Management Search System provides projects constructed for FDOT. The CAR system has all the reported crashes. Crash reports included in CAR have information such as severity, crash type, and other crash related characteristics. This system allows us to locate crashes from 2003 to date. Crashes are divided into 30 different crash types including rear end, head on, side swipe, angle, etc. Crashes are also divided into five crash severities: fatal, incapacitating injury, non-incapacitating, possible injury and property damage only (PDO).

In order to compare the transferability of applying SPFs from HSM, we specifically target four-legged intersection in urban/suburban areas. Since crash reports in Florida sometimes misclassify left-turn crashes as angle crashes, we could only estimate angle and left turn crashes together. In HSM, the expected crash count for different crash types such as rear-end crashes and angle crashes are calculated based on proportion to KABC crashes and PDO crashes. However, there is no proportion of left turn crashes provided by HSM. Therefore, it is not possible to calculate the predicted crash count for left-turn crashes using the ratio suggested in HSM. In this case, we cannot compare the CMFs value for angle crashes nor left turn crashes.

The treatment locations were chosen from two sources, one is from the Financial Project Search System and another part is from RCI, both maintained by Florida Department of Transportation (FDOT). We have selected the signal installation date from 2005 to 2010. After retrieving these data, we combined traffic volume on the major road and the minor road with the target intersections using GIS.
On the other hand, the Ohio data is collected from the Highway Safety Information System (HSIS). Crash data was combined from 2003 to 2011.

Summary of Data Collection
Twenty-nine intersections that were signalized were identified in Florida. The CMFs were estimated based on these 29 signalized intersections. For reference intersections in Florida, data for 126 intersections were located with major and minor AADT. On the other hand, we have more than 1000 reference locations in the state of Ohio. In order to compare the Florida SPFs with Ohio SPFs, we also control the sample size by randomly selecting 126 intersections from Ohio. The descriptive statistics for the treatment group, Florida reference group, and Ohio reference group are shown in Table 1. In the first part of the table, descriptive statistics for the treated sites are shown, if the variable is in before condition we inserted “Before”, if it is after condition we inserted “After”.

The roadway variables include:
1. Annual average daily traffic on major road (Maj_AADT)
2. Annual average daily traffic on minor road (Min_AADT)
3. Total Average annual daily traffic entering the intersection (Tot_AADT)

For each variable, simple statistics is provided with mean, standard deviation and minimum and maximum value also shown in Table 1. In addition, descriptive statistics of three crash types are also shown. On top of the crash count, the crash rate for each crash types is calculated. The unit of rate is “million vehicles entering the intersection per year per site”. The detailed explanations for these crash types are:

1. Total crashes (KABCO) - /per site per yr
2. Fatality and injury crashes (KABC) - /per site per yr
3. Rear-end crashes (Rear) - /per site per yr
4. Total crash Rate (KABCO_Rate) - / per mvmt per yr per site
5. Fatality and injury crash Rate (KABC_Rate) - / per mvmt per yr per site
6. Rear-end crash Rate (Rear_Rate) - per mvmt per yr per site

Overall, 29 treatment sites are located with 126 reference sites from Florida and 126 reference sites from Ohio. To insure the quality of the SPFs, we checked the reference group to verify that there is no major geometry change in the research period. For the data in Florida, we have checked street images from multiple years using Google Earth. On the other hand, for the data in Ohio, we made sure that the selected sites do not overlap with the treatment list in the research period.

The crash rate for KABCO is much higher in Ohio compared to Florida. Ohio has KABCO crash rate at 276 per million vehicles per year, and there are only 78,981 crashes per million vehicles per year in Florida. A similar situation was found for rear-end crashes. The rear-end crash rate in Ohio is more than 5 times than in Florida which is a significant different. On the other hand, crash rate for KABC in Ohio is 60 percent more than in Florida. Due to the differences in crash rate, it is expected that the predicted crash count for each state based on its own SPFs are different. Therefore, the calibration factor is needed to bridge this gap across regions.
1 **Table 1 Descriptive Statistics**

Descriptive Statistics for Treatment Sites  \( N=29 \)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maj_AADT_Before</td>
<td>35,954</td>
<td>24,820</td>
<td>6,800</td>
<td>110,500</td>
</tr>
<tr>
<td>Min_AADT_Before</td>
<td>10,513</td>
<td>13,024</td>
<td>1,416</td>
<td>63,500</td>
</tr>
<tr>
<td>Tot_AADT_Before</td>
<td>46,467</td>
<td>35,156</td>
<td>10,281</td>
<td>162,500</td>
</tr>
<tr>
<td>Maj_AADT_After</td>
<td>38,275</td>
<td>30,862</td>
<td>6,700</td>
<td>149,000</td>
</tr>
<tr>
<td>Min_AADT_After</td>
<td>7,728</td>
<td>8,684</td>
<td>700</td>
<td>37,000</td>
</tr>
<tr>
<td>Tot_AADT_After</td>
<td>46,002</td>
<td>33,906</td>
<td>8,100</td>
<td>169,000</td>
</tr>
<tr>
<td>KABCO_Before</td>
<td>11.655</td>
<td>9.893</td>
<td>0</td>
<td>40</td>
</tr>
<tr>
<td>KABC_Before</td>
<td>6.483</td>
<td>5.166</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Rear_Before</td>
<td>3.414</td>
<td>3.708</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Rear_After</td>
<td>6.172</td>
<td>9.312</td>
<td>0</td>
<td>42</td>
</tr>
</tbody>
</table>

Descriptive Statistics for Reference Sites in Florida  \( N=126 \)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maj_AADT</td>
<td>9,791</td>
<td>8,590</td>
<td>850</td>
<td>42,500</td>
</tr>
<tr>
<td>Min_AADT</td>
<td>1,864</td>
<td>1,902</td>
<td>100</td>
<td>15,000</td>
</tr>
<tr>
<td>Total_AADT</td>
<td>11,655</td>
<td>9,145</td>
<td>1,500</td>
<td>48,500</td>
</tr>
<tr>
<td>KABCO</td>
<td>1.198</td>
<td>2.186</td>
<td>0</td>
<td>13.4</td>
</tr>
<tr>
<td>KABC</td>
<td>0.613</td>
<td>1.128</td>
<td>0</td>
<td>5.2</td>
</tr>
<tr>
<td>Rear</td>
<td>0.291</td>
<td>0.737</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>KABCO_Rate (per mvmt/yr)</td>
<td>78.981</td>
<td>151.491</td>
<td>0</td>
<td>839.080</td>
</tr>
<tr>
<td>KABC_Rate (per mvmt/yr)</td>
<td>40.812</td>
<td>85.410</td>
<td>0</td>
<td>586.207</td>
</tr>
<tr>
<td>Rear_Rate (per mvmt/yr)</td>
<td>16.542</td>
<td>36.981</td>
<td>0</td>
<td>245.989</td>
</tr>
</tbody>
</table>

Descriptive Statistics for Reference Sites in Ohio  \( N=126 \)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maj_AADT</td>
<td>13,031</td>
<td>8,625</td>
<td>530</td>
<td>46,090</td>
</tr>
<tr>
<td>Min_AADT</td>
<td>4,401</td>
<td>3,593</td>
<td>259</td>
<td>19,400</td>
</tr>
<tr>
<td>Total_AADT</td>
<td>17,432</td>
<td>10,708</td>
<td>980</td>
<td>54,249</td>
</tr>
<tr>
<td>KABCO</td>
<td>5.6226</td>
<td>7.073</td>
<td>0</td>
<td>38.889</td>
</tr>
<tr>
<td>KABC</td>
<td>1.324</td>
<td>1.158</td>
<td>0</td>
<td>7.556</td>
</tr>
<tr>
<td>Rear</td>
<td>2.336</td>
<td>3.692</td>
<td>0</td>
<td>23.667</td>
</tr>
<tr>
<td>KABCO_Rate (per mvmt/yr)</td>
<td>276.309</td>
<td>240.000</td>
<td>0</td>
<td>1158.701</td>
</tr>
<tr>
<td>KABC_Rate (per mvmt/yr)</td>
<td>66.968</td>
<td>58.946</td>
<td>0</td>
<td>249.373</td>
</tr>
<tr>
<td>Rear_Rate (per mvmt/yr)</td>
<td>104.122</td>
<td>119.719</td>
<td>0</td>
<td>606.938</td>
</tr>
</tbody>
</table>
RESULTS

Safety performance functions were developed using the NB Model formulation. SPFs were developed based on total crashes, KABC (fatal and injury) crashes, and rear-end crashes respectively. In this study, we targeted on four-legged intersections as our locations. In this section, SPFs for Florida, Ohio and HSM will be developed. Calibration factors are also shown in the following paragraph. The predicted crash counts were calculated from the SPFs then adjusted by calibration factors. Using the different predicted crash counts from each source (Florida, OH, HSM), we estimated CMF accordingly.

Safety Performance Function

HSM and other research suggest that it is ideal to use the log of major AADT and the log of minor AADT to develop SPFs. However, after developing SPFs using major and minor AADT as separate variables, we found that the model fitness is worse than using the total AADT in Florida and Ohio. In addition, the log of minor AADT is not significant in Florida. Therefore, we estimated the models using the log of the total AADT. On the other hand, we applied the base condition model of the urban and suburban arterials in HSM. The model form in HSM is provided with the coefficient of the log of major AADT and the log of minor AADT. In this case, equation 9 represents the model form for developing SPFs for Florida and Ohio and equation 10 demonstrates the equation for HSM.

The variables included in the model form can be explained as follows:

i. Log AADT on Both Major and Minor Road (Log Total_AADT)
ii. Log AADT on Major Road (Log Major_AADT)
iii. Log AADT on Minor Road (Log Minor_AADT)

The equations can be written in this form.

\[ N = \exp(\beta_0 \times (\text{Total}_\text{AADT})^{\beta_1}) \] (9)

\[ N = \exp(\beta_0 \times (\text{Major}_\text{AADT})^{\beta_2} \times (\text{Minor}_\text{AADT})^{\beta_3}) \] (10)

where \( N = \) Crash Frequency \( \beta_0 = \) Intercept \( \beta_1 = \) Coefficient for log(Total_aadt) \( \beta_2 = \) Coefficient for log(Major_AADT) \( \beta_3 = \) Coefficient for log(Minor_AADT)

The relationship between total AADT and each crash types is shown in Figure 1. The y-axis is the predicted crash count per year and x-axis is AADT entering the intersections. Due to the limitation that we do not have the data for developing the SPFs in HSM, we cannot include the fitted line for HSM. In the figure, the dark gray lines represent fitted values for Florida and the lighter gray lines represent fitted values for Ohio. For crash types “KABCO” and “rear-end” crashes, Ohio’s predicted crashes is higher than Florida’s for our study AADT group. However, for KABC crashes, Ohio’s predicted crashes is higher only at low AADT intersections. Florida’s predicted crash count becomes higher when total AADT is higher than 15,955 (vehicles entering intersection).
The results in Table 2, show that three variables were selected to be included in the final SPF. According to the result, we could see that the coefficient is much different from each source. For the models developed for Florida and Ohio, all listed coefficient is significant at 99 percent level. When comparing the SPFs from Florida with Ohio, the coefficient of the log of total AADT is very different from the coefficient in Ohio. Besides, the SPFs from HSM consist of the major and minor AADT separately, which is different from Florida and Ohio condition as well. It is worth noting that there is no SPFs for rear-end nor left turn crashes in the SPFs from HSM. Instead, a proportion was suggested based on FI and PDO crashes respectively in HSM. Therefore, as shown in Table 2, the SPF for rear-end is stated as the pound sign.

It is worth noting that in HSM, the suggested way to calculate rear-end crashes is to estimate KABC and PDO crashes first, and then multiply the predicted crashes by a certain ratio to get the predicted count. In fact, the SPFs are developed based on the data from Minnesota and North
Wang, Abdel-Aty, and Lee

However, the proportion factors are developed based on the data collected in California. This inconsistency may cause potential bias when applying the SPFs to estimate rear-end crashes.

### Table 2 SPFs for each Crashes Types (The Urban 4-Leg Intersections)

#### Negative Binomial Model using Data in Florida  N=126

<table>
<thead>
<tr>
<th></th>
<th>KABCO</th>
<th>KABC</th>
<th>Rear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-18.601</td>
<td>-22.222</td>
<td>-21.754</td>
</tr>
<tr>
<td>log(Total_AADT)</td>
<td>1.985</td>
<td>2.304</td>
<td>2.147</td>
</tr>
<tr>
<td>Overdispersion Parameter</td>
<td>0.193</td>
<td>0.250</td>
<td>0.294</td>
</tr>
</tbody>
</table>

#### Negative Binomial Model using Data in Ohio  N=126

<table>
<thead>
<tr>
<th></th>
<th>KABCO</th>
<th>KABC</th>
<th>Rear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-12.212</td>
<td>-12.709</td>
<td>-16.864</td>
</tr>
<tr>
<td>log(Total_AADT)</td>
<td>1.416</td>
<td>1.321</td>
<td>1.791</td>
</tr>
<tr>
<td>Overdispersion Parameter</td>
<td>1.487</td>
<td>1.534</td>
<td>1.140</td>
</tr>
</tbody>
</table>

#### Negative Binomial Model using Data in HSM  N=96

<table>
<thead>
<tr>
<th></th>
<th>KABCO</th>
<th>KABC</th>
<th>Rear</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-8.97</td>
<td>-11.20</td>
<td>#</td>
</tr>
<tr>
<td>log(Major_AADT)</td>
<td>0.82</td>
<td>0.93</td>
<td>#</td>
</tr>
<tr>
<td>log(Minor_AADT)</td>
<td>0.25</td>
<td>0.28</td>
<td>#</td>
</tr>
<tr>
<td>Overdispersion Parameter</td>
<td>2.50</td>
<td>2.08</td>
<td>#</td>
</tr>
</tbody>
</table>

Note: #Use proportion in HSM
All coefficients are significant at 99% level
The CMFs were calculated as shown in Table 3. After calculating the predicted crash count from SPFs, the empirical Bayes method was utilized to estimate the crash modification factors (CMFs) for each category. The comprehensive CMFs result is shown in Table 3. Similar to previous research findings, signalization will result in more rear-end crashes. However, the CMFs for KABCO are different from HSM. According to HSM, the CMF is at 0.95 when signalizing an intersection in an urban area. But the CMF calculated in Florida is at 0.785 which is 0.165 lower than HSM. In fact, if we apply the SPFs and its corresponding calibration factors, the CMF values will become 1.062 and 1.072 which is significantly higher than using the SPF in Florida.

Table 3 Crash Modification Factors using SPFs of Different States w/o Calibration Factors

<table>
<thead>
<tr>
<th></th>
<th>FLORIDA</th>
<th>OHIO</th>
<th>HSM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard Error</td>
<td>Standard Error</td>
<td>Standard Error</td>
</tr>
<tr>
<td>KABCO</td>
<td>0.785</td>
<td>0.057</td>
<td>1.060</td>
</tr>
<tr>
<td>KABC_CMF</td>
<td>0.599</td>
<td>0.057</td>
<td>0.856</td>
</tr>
<tr>
<td>REAR_CMF</td>
<td>1.283</td>
<td>0.167</td>
<td>1.829</td>
</tr>
</tbody>
</table>

The calibration factors were estimated in this research as shown in Table 4. In Ohio, the calibration factors show the crash count is higher than that in Florida. Therefore, the calibration factors for all crash severities and types are below 1. For rear-end crashes, the calibration factor is as low as 0.23. This means that the rear-end crash count in Ohio is much higher than in Florida. On the other hand, the SPFs suggested in HSM were developed using data from Minnesota and North Carolina. It is worth noting that the number of KABC crashes depicted in HSM is much less than in Florida.

Table 4 Calibration Factors for OH and HSM SPFs Based on FL

<table>
<thead>
<tr>
<th></th>
<th>KABCO</th>
<th>KABC</th>
<th>REAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>OH</td>
<td>0.3835</td>
<td>0.8171</td>
<td>0.2342</td>
</tr>
<tr>
<td>HSM</td>
<td>0.785</td>
<td>2.067</td>
<td>0.5229</td>
</tr>
</tbody>
</table>

After applying calibration factors along with the predicted crash count based on the SPFs, the adjusted CMFs are shown in Table 5. Comparing the results without applying calibration factors as shown in Table 3 with the calibrated ones shown in Table 5, there is only minor different. This is due to the weight for predicted value $E\{k\}$ is small as shown in Equation 5. Therefore, after adjusted by Equation 6 which is the ratio suggested by (27), the differences are marginal.
Table 5 Crash Modification Factors using SPFs of Different States with Calibration Factors

<table>
<thead>
<tr>
<th></th>
<th>FLORIDA</th>
<th>OHIO</th>
<th>HSM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CMF</td>
<td>Error</td>
<td>CMF</td>
</tr>
<tr>
<td>KABCO</td>
<td>0.785</td>
<td>0.057</td>
<td>1.062</td>
</tr>
<tr>
<td>KABC_CMIF</td>
<td>0.599</td>
<td>0.057</td>
<td>0.857</td>
</tr>
<tr>
<td>REAR_CMIF</td>
<td>1.283</td>
<td>0.167</td>
<td>1.837</td>
</tr>
</tbody>
</table>

By plotting the CMF values in Table 5 to line chart as shown in Figure 2, we can observe the difference more closely. The CMFs using HSM and Ohio SPFs are significantly higher for KABCO and KABC crashes when using the locally developed Florida SPFs (after adjustment by the calibration factor). This is an important finding since the CMFs become insignificant in KABCO crashes when applying SPF from different sources. If we apply the SPF using Florida data, we would expect to get 21.5% crash reduction and it is statistically significant. However, substituting the Florida SPFs with the SPFs from HSM or Ohio, we would get CMF values slightly higher than 1 and not significant. In addition, the rear-end crashes have a similar pattern which CMFs from HSM and Ohio are also higher than Florida but not significant.
Figure 2 Comparison of CMF using SPFs from the Different States (90% Confidence Interval)

The estimated CMFs were compared from this research with others. According to the results shown in Table 6, the CMFs for KABCO and KABC crashes in this research is not significantly different from previous research (HSM (5) and NCHRP Vol. 491(15)). In fact, the results specified in this paper have a lower standard error which indicates that the CMFs are more precise. However, there is one issue worth noting which is that the CMFs were developed for all types of crashes in HSM and NCHRP Vol.491. However, the CMFs in this research are specifically for multi-vehicle crashes. The reason we did not use all types of crashes (including single vehicle crashes) is due to the SPFs in HSM separate the urban/suburban into multi-vehicle crashes and single vehicle crashes. In order to compare CMFs using SPFs provided in the HSM, we chose...
Wang, Abdel-Aty, and Lee

multi-vehicle crashes since the majority of crashes at intersections are multi-vehicle crashes. Therefore, the CMF values stated in Table 6 can still be compared with the sources from HSM and NCHRP Vol. 491.

**Table 6 Signalization Crash Modification Factors in HSM and NCHRP Report**

<table>
<thead>
<tr>
<th>Crash Severity</th>
<th>Crash Type</th>
<th>CMF</th>
<th>Standard Error</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>KABCO</td>
<td>All Crash</td>
<td>0.95</td>
<td>0.09</td>
<td>HSM (5)</td>
</tr>
<tr>
<td>KABCO</td>
<td>Multi</td>
<td>0.785</td>
<td>0.06</td>
<td>This paper (FL SPF)</td>
</tr>
<tr>
<td>KABCO</td>
<td>Multi</td>
<td>1.06</td>
<td>0.08</td>
<td>This paper (OH SPF)</td>
</tr>
<tr>
<td>KABCO</td>
<td>Multi</td>
<td>1.07</td>
<td>0.08</td>
<td>This paper (HSM SPF)</td>
</tr>
<tr>
<td>KABC</td>
<td>All Crash</td>
<td>0.77</td>
<td>0.27</td>
<td>NCHRP Volume 491 (15)</td>
</tr>
<tr>
<td>KABC</td>
<td>Multi</td>
<td>0.60</td>
<td>0.06</td>
<td>This paper (FL SPF)</td>
</tr>
<tr>
<td>KABC</td>
<td>Multi</td>
<td>0.86</td>
<td>0.09</td>
<td>This paper (OH SPF)</td>
</tr>
<tr>
<td>KABC</td>
<td>Multi</td>
<td>0.85</td>
<td>0.09</td>
<td>This paper (HSM SPF)</td>
</tr>
</tbody>
</table>

*Bold text* significant at 90% confidence interval

*Multi represents multi-vehicle crashes

## CONCLUSION AND RECOMMENDATIONS

In this research, we estimated CMFs for the safety effect of signalization implemented in Florida. This paper applied the Empirical Bayes method to develop CMFs for KABCO, KABC, and rear-end crashes using SPFs and calibration factors developed based on Florida, HSM (North Carolina and Minnesota), and Ohio. The results showed the significant difference between the CMFs using Florida, Ohio, and HSM SPFs. Therefore, we concluded that it is not suitable to apply data to SPFs developed from other jurisdictions without further investigation.

The CMFs were estimated for three crash types (i.e., KABCO, KABC, rear-end). For each crash type, the Florida, Ohio, and HSM SPFs were used to estimate CMFs. It was observed that the coefficients of SPFs from the three sources are not comparable. In addition, the calibration factors comparing the observed crash counts in Florida to the predicted crash counts from Ohio and HSM SPFs were also calculated. According to the results, the calibration factor for Ohio KABCO crashes is 0.38. On the other hand, the calibration factor for KABC crashes in HSM is 2.07. In fact, the different reporting threshold among states may be the key factor to explain why the calibration factor is needed. In Ohio, the reporting threshold is 1000+ of property damage. However, in Florida, the threshold is based on specific crash types. These crash types are, involve any injury, leaving the scene, commercial motor vehicles involved, require a wrecker to remove it from the scene, and driving under influence. The different reporting threshold has to be considered when using SPFs from a non-local jurisdiction. That is to say, the calibration factors not only adjusting the different crash behavior among states but also adjusting the different reporting threshold. It is challenging to separate the effect of different reporting threshold from behavioral differences. Therefore, in this paper, we used the calibration factors to consider the combination effect as suggested by HSM and found the calibration factors are really different for each crash type and state. Moreover,
it is desirable that the next edition of HSM address the issue of reporting threshold because its
SPFs are intended for nationwide use.

Comparing the CMFs calculated based on the calibrated SPFs with the non-calibrated SPFs, it was
found that differences are minimal. In detail, we found that when observed before and observed
after crashes at 150 (crash count for KABC) or more, there is no significant difference between
whether use the calibrated SPFs or not. Another important finding in this paper is that some CMF
values are statistically significantly different when using SPFs developed from other states. In
fact, the CMFs for KABCO and KABC crashes are 0.785 and 0.60 for Florida, 1.06 and 0.86 for
Ohio, and 1.07 and 0.85 for HSM, respectively. The CMFs using the data in Florida are
significantly lower than Ohio and HSM. This indicates that SPFs may not be transferrable and it
results in the biased estimation of CMFs. An erroneous judgment would be made if we borrow
SPFs from other states. The CMF is 0.785 for KABCO using the SPF in Florida which is a closer
estimate since the treatment targets are located in Florida. However, when applying the SPFs
from Ohio or HSM, we would get a higher CMF value at 1.06 and 1.07, respectively. In this case,
if Florida does not have its own SPFs and borrows the SPFs from HSM or Ohio, the judgment will
be different from using the Florida SPF. Therefore, according to these results, it is suggested to
apply SPFs from other jurisdictions after thorough examination and validation when developing
CMFs. It would be most desirable to use locally developed own SPFs if the sufficient traffic and

Acknowledgments: The authors acknowledge the financial support of the Florida Department of
Transportation (FDOT). The authors also wish to thank FDOT for providing the data used in this
research. Dr. Raghavan Srinivasan of UNC HSRC has provided the Ohio HSIS data that were used
in this study. All opinions and results are solely those of the authors.
REFERENCES


2. FHWA. Intersection Safety Issue Briefs - Safety | Federal Highway Administration.

3. FHWA. Chapter 4C - MUTCD 2009 Edition - FHWA.


18