CALIBRATION OF SAFETY PREDICTION MODELS FOR PLANNING TRANSPORTATION NETWORKS

Andrew Tarko
Purdue University
School of Civil Engineering
550 Stadium Mall Drive
West Lafayette, IN 47907
Phone: (765) 494-5027
Fax: (765) 496-1105
Email: tarko@purdue.edu

ABSTRACT

The current planning practice addresses safety implicitly as a byproduct of adding capacity and operational efficiency to the transportation system. Safety conscious planning is a new proactive approach to the prevention of crashes by establishing inherently safe transportation networks through integrating safety consideration into the transportation planning process. One of the major concerns in predicting crashes in transportation networks is the applicability and accuracy of crash prediction models.

The paper presents two alternative formulations of the calibration problem consistent with the maximum likelihood approach. The proposed formulations can be viewed as a generalized version of the existing calibration procedure proposed in the past for individual crash prediction models. The proposed formulations are useful for road networks and for any transportation mode, provided that the needed prediction models are available.

The proposed calibration applied to individual elements of the test network yielded crash estimates that exhibited a considerable bias accumulation at the system level. The calibration task was redefined to focus on the prediction of the cumulative number of crashes in the user-defined sub-networks. The second method gave more acceptable results. The paper demonstrates the feasibility of the proposed approach, which may be helpful in developing a new class of tools for safety conscious planning.

Key words: safety conscious planning, safety prediction, calibration

Word count: 5,325 + 3 figures + 5 tables
INTRODUCTION

While traditional safety management identifies and remedies existing safety problems, preventing safety problems requires considering safety in the transportation planning and design phases. The prevailing planning practice addresses safety implicitly as a byproduct of adding capacity and operational efficiency to the transportation system. Safety impacts are sometimes assessed using crude methods supplemented with expert judgment.

To encourage more explicit consideration of safety in transportation planning, highway administration has been promoting safety conscious planning. Safety conscious planning is a proactive approach to the prevention of crashes by establishing inherently safe transportation networks. The objective is to integrate safety consideration into the transportation planning process at all levels starting with the Statewide Transportation Improvement Plans (STIP) and followed by consideration of safety objectives in long-range planning (FHWA, 2003).

Chatterjee et al. (2001) discussed the general issues of incorporating safety considerations into the planning process. They investigated common planning practices and concluded that the lack of data and suitable analysis tools led to inevitably relying on subjective assessment of the safety impacts. De Leur and Sayed (2002) proposed an approach to safety conscious planning that included safety-improvement strategies suitable in the planning stage and prediction tools for evaluating planning alternatives. They postulated that crash prediction models should include the risk exposure (traffic volume, mileage), the likelihood of involvement in a crash, and the crash severity. Dumbauch et al. (2004), on the other hand, investigated institutional barriers in involving highway safety agencies in safety conscious planning. In the conclusions, they pointed out the known barriers, but they also identified opportunities, such as better coordination, entrepreneurial approach, and flexibility in administering grant funds.

Although safety consideration can be incorporated into planning in many ways, safety prediction must be present in all alternatives of a planning process. Crash prediction models are available for traffic analysis zones (TAZ) and for transportation infrastructure
The TAZ crash models are suitable for long-term planning to evaluate the safety impacts of land use and economic growth (Tarko, et al., 1996; Hadayeghi, et al., 2003; de Geuvara, et al. 2004). Although some of these models include transportation characteristics (network density, VMT, etc.), the aggregate representation of the transportation infrastructure does not easily allow modeling infrastructure-specific solutions. TIF crash models are a better option as they include the road and traffic characteristics of individual links and nodes. The TIF crash models have been the subject of research for many years. These models reflect the gradual progress in modeling techniques, including OLS, Poisson, and Negative Binomial count models, and logit and probit models (Washington, et al., 2003).

The Indiana Department of Transportation (INDOT) assisted by Purdue research teams has been working on a Safety Management System for the last ten years. An advanced method of screening networks for roadway hazards, updated crash reduction factors, and crash predictive equations developed in the mentioned research have been developed and implemented through *Guidelines for Highway Safety Improvements in Indiana* (Tarko and Mayank, 2003) and a tool for a system-wide analysis of Indiana highways for targeted improvements (Lamptey, Labi, and Sinha, 2003).

Development of a GIS-based method of predicting safety in transportation networks for long and short-term planning is underway in Indiana. As emphasized by Washington et al. (2004), safety should be introduced in all steps of the planning process: starting with incorporating safety into the vision, goals, and objectives, through technical analysis, development of programs, and monitoring of the system. The current research project funded by INDOT through their Joint Transportation Research Program with Purdue University is aimed to implement the link and to further link safety prediction models to the long-term planning process of the state network through the GIS-based planning tools. Safety prediction with SPF is shown in Figure 1 as a process executed in parallel to the traffic demand prediction. The addition of the safety prediction component allows evaluating alternative planning solutions from the standpoint of traffic and safety.
One of the major concerns in predicting crashes in transportation networks is the applicability and accuracy of crash prediction models for network links and nodes (Lord and Persaud, 2003). Models developed in the past were frequently based on an insufficient number of observations or for regions other than the studied one. These models require calibration and validation, and this paper evaluates two methods of calibrating safety prediction models for user-defined set of network partitions. The paper presents a formulation of the calibration problems consistent with the maximum likelihood approach. The methods are tested and evaluated for the Indiana state road network.
CALIBRATION PROBLEM

A typical network representation includes various types of nodes and links supplemented with basic traffic, control, and roadway characteristics. Although many details of geometric and control design are not known in the planning stage of a transportation network, the main features of intersections and traffic control, classes of roads with their typical cross-sections, travel speeds, and daily and rush hour traffic volumes are known. These characteristics should be utilized to improve anticipation of future safety in the planned network.

The concept of Safety Performance Functions (SPF) seems to be the most reliable method of predicting future safety at individual road intersections and links. SPFs connect various roadway and traffic characteristics with crash frequency and various levels of severity. The most common structures of the models for links (segments) and nodes (intersections) are as follows:

\[ a = E \cdot \exp(\gamma_0 + \gamma_1 \cdot X_1 + ... + \gamma_n \cdot X_n) \]
\[ E = \begin{cases} 
L \cdot Q^\alpha & \text{for segments,} \\
Q^\alpha \cdot Q^\beta & \text{for intersections.} 
\end{cases} \]  

where:

- \( A \) = expected annual number of crashes of certain severity,
- \( E \) = exposure to risk function,
- \( Q \) = AADT along the segment or average AADT along one direction of the intersection,
- \( X_1...X_n \) = road, traffic, and other characteristics,
- \( \alpha, \beta, \gamma_1...\gamma_n \) = model parameters.

SPFs are expected to be used in the future Highway Safety Manual to predict and evaluate roadway safety. The initial tests of SPFs for predicting safety in road networks and traffic analysis zones indicate the usefulness of the functions (Lord and Persaud, 2003). For convenience, a safety performance function should consist of a base safety performance function (BSPF) and a set of Accident Modification Factors (AMFs). A model structure equivalent to Equation 1 is as follows:
\[ a = BSFP \cdot AMF_1 \cdot \ldots \cdot AMF_n \]  

where:

- \( a_o \) = crash frequency at a specific severity level;
- \( BSFP \) = basic safety performance function; \( BSFP = \exp(\gamma_0) \cdot E \)
- \( AMF_i \) = crash modification function \( i \); \( AMF_i = \exp[\gamma_i (X_i - M_i)] \),
- \( M_i \) = average or default value of characteristic \( X_i \).

The BSPF includes traffic volume and sometimes link length while the AMFs incorporate the impact of known roadway and control characteristics. In many cases, the AMFs are adapted from research done for regions other than the studied one. The differences between regions and the changes in safety over a period of years call for calibration of SPFs. Planners may also want to reflect sub-regional and local differences in large networks. For example, different parts of the region may experience different weather conditions (e.g., northern and southern California) or topographical conditions (e.g., northern and southern Indiana). Also, some parts of the region may be more developed than the rest and their safety may somewhat differ, such as the metropolitan city of Indianapolis versus smaller urban communities in Indiana. In such cases, the planner may want to use calibration factors to consider these differences when predicting future safety in road networks.

Harwood, et al. (2000) proposed a simple procedure of calibrating a valid SPF developed for a region other than the studied one. They propose a single calibration factor that was intended to remove an overall bias present in the original SPF. This procedure assumed that the functional form of the original model is correct and the model parameters associated with individual model variables reasonably reflect the relative safety impacts of these variables in the region to which the model is calibrated. Persaud, et al. (2003) checked the proposed procedure and found that in some cases the procedure worked satisfactorily, but in other cases, additional re-calibration of the parameters associated with the traffic volume was beneficial. For the purpose of this presentation, it is assumed that the re-calibration of the parameter associated with the volume is not needed.
Lord and Bonneson (2005) applied the calibration procedure proposed by Harwood at al. (2000) to interchange ramps. Instead of using a single calibration factor, they identified several cases of ramps and estimated a single separate calibration factor for each ramp case. In spite of the multiplicity of calibration factors, the authors could follow the approach proposed by Harwood, et al. because the identified calibration cases were separable and the factors could be estimated one by one, independently from each other.

This paper investigates a situation where calibration factors are allowed to depend on each other. This situation is expected to frequently occur in the calibration of SPFs for road networks as explained in the following part of the presentation. Two types of calibration are considered:

1. **Standard calibration** where one calibration factor applies to each SPF,
2. **User-defined calibration** where standard calibration factors are supplemented with factors for sub-networks defined by the user.

In a general case, each link or node (element) of a road network has a corresponding set of calibration factors. In the standard calibration, only one calibration factor applies to each network element and it corresponds to the link or node type. In the user-defined calibration, additional calibration factors may apply if the network element belongs to sub-networks defined by the user. A resulting set of calibration coefficients \( CF_j \) apply to crash predictions \( a_o \) made with original SPFs:

\[
a = a_o \cdot \prod_j CF_j .
\]  

The assignment of calibration factors to network elements is defined with a coincidence matrix \( \Delta \) with elements \( \delta_{ij} = 1 \) if the calibration factor applies to the network elements \( i \) and \( 0 \) otherwise. A calibration subset \( s \) is a set of network elements that have identical vector of coincidence elements \( \delta_{ij} \). In other words, the same set of calibration parameters are used within subset \( s \).

To better explain the concept, let us consider a simple network where all nodes are of one type (one SPF applies to all the nodes) and also all links are of one type (another single SPF applies to all the links). The standard calibration includes two calibration factors:
CF_{Node} and CF_{Link}. In addition, the user divides the region into two geographical areas (CF_{South}, and CF_{North}) and focuses the calibration on a specific corridor (CF_{Corridor}) in the southern part of the network. Figure 2 shows the breakdown of the network and Table 1 assigns calibration factors to network elements through coincidence elements $\delta$ and identifies calibration subsets $s$.

Figure 2 An example road network.
Table 1 Coincidence matrix $\Delta$ for example network elements and five calibration factors (resulted subsets $s$ are shown in the last column).

<table>
<thead>
<tr>
<th>Network element</th>
<th>Calibration factors</th>
<th>Subset $s$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$CF_{\text{Node}}$</td>
<td>$CF_{\text{Link}}$</td>
</tr>
<tr>
<td>01</td>
<td>1(^{a})</td>
<td>0(^{b})</td>
</tr>
<tr>
<td>02</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>07</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>08</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>09</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0102</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0105</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>0711</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0812</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0910</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0913</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>1620</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1920</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1314</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1317</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1415</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1418</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1519</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1718</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1819</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Notes:  
\(^a\) $\delta_{01,\text{Node}} = 1$;  
\(^b\) $\delta_{01,\text{Link}} = 0$. 

TRB 2006 Annual Meeting CD-ROM  
Paper revised from original submittal.
The calibration task is to find the optimal values of calibration factors $CF_j, j=1..n$, such that the adjusted crash predictions $a_0 \cdot \prod_j CF_j$ best fit the crash counts recorded in the network in the recent period.

The following sections present alternative formulations and select the most promising one from the point of view of cumulative prediction bias and computational robustness. The calibration is, in its entirety, a complex task and includes numerous issues such as omitted model variables, poor data, inadequate functional form of the model, predictive ability across time and space, over-fitting, etc. (Oh, et al., 2003). Additional issues related to the implementation of a calibration procedure include: missing data, underreporting of crashes, errors in network representation, lack of needed Safety Performance Functions, lack of suitable tools for assigning crashes to links and nodes, etc. All of these issues have been reported by numerous authors and remain the subject of research aimed at finding remedies or at least mitigations. This paper focuses on advancing the analytical component of estimating calibration factors in road networks to improve the accuracy of predicting safety in corridor studies and other system-level long-term planning analyses.

**ML METHOD FOR DISAGGREGATE CRASH COUNTS**

In the example presented in Figure 1 each subset $s$ includes at least one calibration factor which is included in at least one other subset $s$. It is not possible to calibrate individual factors separately from other factors. A single calibration problem defined for the entire road network is needed. For computational convenience, let calibration factor $CF_j$ be an exponential function $CF_j = \exp(b_j)$. Utilizing the coincidence matrix $\Delta$ and the original crash frequency predictions, the adjusted crash frequency for network element $i$ is:

$$a_i = a_{0i} \prod_{j=1,n} CF_j = a_{0i} \prod_{j=1,n} \exp(b_j \cdot \delta_{ij})$$

where:
- $a_i =$ adjusted crash frequency for network element $i$;
- $a_{0i} =$ original crash frequency calculated for network element $i$ with the original SPF;
\( \delta_{ij} = 1 \) if calibration factor \( j \) applies to element \( i \) and 0 otherwise; and
\( b_j = \) calibration parameter in calibration factor \( \text{CF}_j \),
n = number of calibration factors for the studied network.

A maximum likelihood (ML) method is our initial choice. Following the current practice
in modeling crashes, a Poisson model with random effects or Negative Binomial model,
is proposed (Washington, et al., 2003, pp. 248-250). The Negative Binomial distribution
\( \text{NB}(c_i, a_i, \alpha_s) \) estimates the likelihood of crash count \( c_i \) on link \( i \) or at node \( i \) when the
expected value is \( a_i \). An over-dispersion parameter \( \alpha_s \) represents unknown Gamma-
distributed random effects in the calibration group \( s \) to which the network element \( i \)
belongs. The corresponding optimization problem includes log-likelihood for the entire
transportation network. The formal optimization problem is:

\[
\max_{\{b\}, \{\alpha\}} LL = \max_{\{b\}, \{\alpha\}} \sum_{s \in S} \left( \sum_{i \in I_s} \ln \text{NB}(c_i, a_{0i} \prod_j \exp(\delta_{ij} \cdot b_j), \alpha_s) \right),
\]

s.t.
\( \alpha_s > 0 \) for all \( s \in S \).

where:
\( \{b\} = \) set of calibration parameters corresponding to factors \( \text{CF} \);
\( \{\alpha\} = \) set of over-dispersion parameters;
\( S = \) set of indices of subsets \( s \);
\( I_s = \) set of indices of the network elements that belong to subset \( s \);
\( \text{NB}(\cdot) = \) Negative Binomial probability distribution function;
\( c_i = \) number of crashes associated with network element \( i \);
\( a_{0i} = \) crash frequency calculated with non-calibrated SPF for network element \( i \);
\( \delta_{ij} = 1 \) if calibration factor \( j \) applies to element \( i \) and 0 otherwise;
\( b_j = \) calibration parameter in calibration factor function \( \text{CF}_j = \exp(b_j) \); and
\( \alpha_s = \) over-dispersion parameter for calibration subset \( s \).
For the robustness of the calculations, the Negative Binomial function in Equation 5 should include the Gamma function instead of factorials to allow searching for any positive value of $\alpha$:

$$NB(c,a,\alpha) = \frac{\Gamma(c+1/\alpha)}{\Gamma(1/\alpha) \cdot \Gamma(c+1)} \cdot \frac{(\alpha \cdot a)^c}{(1+\alpha \cdot a)^{c+1/\alpha}}.$$  \hspace{1cm} (6)

For evaluation purposes, the proposed calibration problem was solved for the Indiana state road network consisting of 19,151 links with counts of crashes recorded in 2004. Figure 3 depicts the size and density of the example state network. It has been coded in TransCAD. The safety performance functions suitable for state roads and calibrated in this test have been published in (Tarko and Kanodia, 2004) and they are shown in Table 2. Including local roads in the network representation would require additional SPF$s$ suitable for local roads. Another alternative is the use of state SPF$s$ and dedicated calibration factors, bringing the crash predictions in line with the crash statistics on local roads. We are currently assembling more elaborated SPF$s$ using results from the past research for Indiana and other regions. Since this work is not finished yet, the calibration procedures are tested with the base SPF$s$.

To test the method for user-defined calibration, the Indianapolis area has been specified as a user-defined sub-network. The Indianapolis, Indiana area was selected as considerably different from the rest of that state. Recent studies of driver behavior at signalized intersections indicated that Indianapolis drivers were more aggressive than in other Indiana areas (Tarko and Perez-Cartagena, 2004). It therefore would be interesting to check whether or not Indianapolis road safety is indeed different from the rest of Indiana. Adding Indianapolis to six types of links produced seven calibration factors:

1. Rural two-lane roads calibration factor,
2. Rural multilane roads calibration factor,
3. Rural interstates calibration factor,
4. Urban two-lane roads calibration factor,
5. Urban multilane roads calibration factor,
6. Urban interstates calibration factor, and
7. Indianapolis area calibration factor.

and corresponding nine calibration subsets:

1. Rural two-lane roads,
2. Rural multilane roads,
3. Rural interstates,
4. Non-Indianapolis urban two-lane roads,
5. Non-Indianapolis urban multilane roads,
6. Non-Indianapolis urban interstates,
7. Indianapolis urban two-lane roads,
8. Indianapolis urban multilane roads, and
9. Indianapolis urban interstates.

The breakdown of the Indiana state network between the calibration subsets by the number of links and basic link-specific data are presented in Table 3. The traffic volumes and link lengths have been obtained from the Indiana travel demand model implemented in TransCAD. The crash counts for 2004 have been obtained from the Indiana Crash Database and assigned to the link using GIS coordinates and the TransCAD standard features. The trends in average AADT and the link lengths are as expected. Urban links are shorter than rural but carry heavier traffic. Traffic increases with the number of lanes and the functional importance of the road.

When dividing the road network into calibration subsets, the total number of crashes per subset should be sufficiently large to allow confident calibration to avoid over-fitting the SPFs. Over-fitting takes place when the calibrated model follows the random fluctuation of the dependent variable in the sample, rather than the trend in the expected values. The percent standard deviation of the total annual number of crashes in the subset can be used as guidelines. The standard deviation can be estimated with the assumption of Poisson variability of crashes over years:
\[ \sigma = \frac{100}{Y \cdot \sqrt{C}}. \] (7)

Standard deviation exceeding the maximum value set by the user, let us say 5 \%, indicates that the subset with such a large standard deviation of total crashes is too small. A solution is to use more years of data or to abandon some of the network divisions. In our case, three calibration subsets seem to be rather small (marked in Table 3 with bold italicized font). One year of data does not allow for fine partitioning of the network. We keep these subsets as it does not defeat the purpose of this presentation. The issue of properly sizing the calibration subsets requires more research and it is outside of the scope of this paper.

Figure 3 Indiana state road network
Table 2 Basic safety performance functions for Indiana (Tarko and Kanodia, 2004)

<table>
<thead>
<tr>
<th>Rural two-lane segment</th>
<th>( a_{LF} = 0.208 \times L \times Q^{0.604} )</th>
<th>0.420</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( a_{FD} = 0.712 \times L \times Q^{0.592} )</td>
<td>0.430</td>
</tr>
<tr>
<td>Rural multilane segment</td>
<td>( a_{LF} = 0.107 \times L \times Q^{0.814} )</td>
<td>0.451</td>
</tr>
<tr>
<td></td>
<td>( a_{FD} = 0.634 \times L \times Q^{0.615} )</td>
<td>0.484</td>
</tr>
<tr>
<td>Urban two-lane segment</td>
<td>( a_{LF} = 0.105 \times L \times Q^{1.080} )</td>
<td>1.253</td>
</tr>
<tr>
<td></td>
<td>( a_{FD} = 0.603 \times L \times Q^{0.896} )</td>
<td>1.349</td>
</tr>
<tr>
<td>Urban multilane segment</td>
<td>( a_{LF} = 0.674 \times L \times Q^{0.435} )</td>
<td>1.588</td>
</tr>
<tr>
<td></td>
<td>( a_{FD} = 2.028 \times L \times Q^{0.460} )</td>
<td>1.946</td>
</tr>
<tr>
<td>Rural interstate</td>
<td>( a_{LF} = 0.044 \times L \times Q^{0.917} )</td>
<td>1.053</td>
</tr>
<tr>
<td></td>
<td>( a_{FD} = 0.169 \times L \times Q^{0.943} )</td>
<td>1.604</td>
</tr>
<tr>
<td>Urban interstate</td>
<td>( a_{LF} = 0.00048 \times L \times Q^{2.238} )</td>
<td>2.383</td>
</tr>
<tr>
<td></td>
<td>( a_{FD} = 0.0057 \times L \times Q^{1.954} )</td>
<td>2.704</td>
</tr>
</tbody>
</table>

Table 3 Calibration subsets

<table>
<thead>
<tr>
<th>Calibration subset</th>
<th>No. of Links</th>
<th>Av. Length (mi)</th>
<th>AADT (veh/day)</th>
<th>Fatal/Injury Total (Std Dev)</th>
<th>PDO Total (Std Dev)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural two-lane roads</td>
<td>11,112</td>
<td>0.685</td>
<td>4,128</td>
<td>3,525 (1.7)</td>
<td>11,386 (0.9)</td>
</tr>
<tr>
<td>Rural multilane roads</td>
<td>1,313</td>
<td>0.648</td>
<td>12,682</td>
<td>770 (3.6)</td>
<td>2,301 (2.1)</td>
</tr>
<tr>
<td>Rural interstates</td>
<td>307</td>
<td>2.507</td>
<td>26,769</td>
<td>534 (4.3)</td>
<td>2,546 (2.0)</td>
</tr>
<tr>
<td>Non-Indianapolis urban two-lane roads</td>
<td>3,057</td>
<td>0.225</td>
<td>9,822</td>
<td>1,491 (2.6)</td>
<td>4,081 (1.6)</td>
</tr>
<tr>
<td>Non-Indianapolis urban multilane roads</td>
<td>2,068</td>
<td>0.248</td>
<td>21,215</td>
<td>1,614 (2.5)</td>
<td>4,538 (1.5)</td>
</tr>
<tr>
<td>Non-Indianapolis urban interstates</td>
<td>235</td>
<td>0.872</td>
<td>39,830</td>
<td>294 (5.8)</td>
<td>1,225 (2.9)</td>
</tr>
<tr>
<td>Indianapolis urban two-lane roads</td>
<td>163</td>
<td>0.272</td>
<td>13,969</td>
<td>104 (9.8)</td>
<td>284 (5.9)</td>
</tr>
<tr>
<td>Indianapolis urban multilane roads</td>
<td>659</td>
<td>0.233</td>
<td>27,942</td>
<td>849 (3.4)</td>
<td>2,329 (2.1)</td>
</tr>
<tr>
<td>Indianapolis urban interstates</td>
<td>237</td>
<td>0.820</td>
<td>83,208</td>
<td>698 (3.8)</td>
<td>3,283 (1.7)</td>
</tr>
</tbody>
</table>

Calibration of the safety performance functions was done for fatal/injury and PDO crashes separately. This practice allows incorporating different costs of crashes at different levels of severity.
Columns titled ML for Links in Tables 4 and 5 show the calibration results. The first thought is that the values of some of the calibration factors considerably differ from one. The explanation may be in inaccurate estimation of the original SPRs based on limited samples. The limited size of the calibration subsets also contributes to this result.

As expected, the over-dispersion parameters estimated in the calibration (not shown) were typically higher than the ones reported in the original publication of the SPFs. The random effects of crash frequencies unexplained with a SPF are stronger when the model is applied to roads outside of the sample used to estimate the model.

The calibration factors for interstate rural and urban links are consistently lower than for those which indicate that the original SPFs for fatal/injury and PDO crashes overstate the crash frequencies. One possible explanation is that the SPFs were developed for links that included interchanges as part of the links. The TransCAD crash counts for interstate links exclude crashes that happened in the 250-foot vicinity of an interchange. It is symptomatic that the overestimation of crashes on urban interstate links is much higher than on rural interstate links. Denser traffic in urban areas exhibits much more interaction between vehicles (and resulted crashes) in the direct vicinity of ramp exit and entrance points than in rural areas.

There is a tendency toward calibration factors higher than one for links other than interstate roads. This tendency may reflect the improvements in crash reporting achieved in the last several years. Indiana modified its crash reports and started using GIS coordinates in 2001. These changes were accompanied by an intense training effort of police officers involved in reporting crashes. Years 2003 and 2004 witnessed a considerable increase in the number of crashes that can be precisely assigned to links and nodes.

Although many of the obtained results can be explained, the troubling observation is that the crash predictions aggregated in calibration subsets do not match the data well. It can be seen in the second parts of Tables 4 and 5. Some of the adjusted estimates differ by nearly 50% of the recorded number of crashes. Transportation planning utilizes
aggregated measures of performance when evaluating and comparing alternative solutions. Although it is important that the safety performance functions properly predict crashes at the level of network elements but it is even more important that the prediction errors do not accumulate beyond an acceptable level when aggregating the predictions. It seems that the maximum likelihood method of calibration applied to individual links and nodes generates results that may be questioned by planners. To address this issue, another approach to SPF calibration based on aggregated crash counts is proposed in the following section of the paper.

LS METHOD FOR COUNTS AGGREGATED IN CALIBRATION SUBSETS

The ML method applied to individual links (in the general case, to network elements including nodes) has not produced the most desirable results. To directly address the need of accurate safety estimation at the sub-network level, the maximum likelihood method will be applied to the total number of crashes in user-defined calibration subsets. Let \( C_s \) denote the total number of crashes recorded in calibration subset \( s \). If we accept the Poisson assumption for these counts, the variance of the total crash count is also \( C_s \). The Poisson distribution for variables with the mean sufficiently high can be approximated with the normal distribution with a standard deviation equal to \( \sqrt{C_s} \). A relevant ML estimation problem for total crash counts in calibration subsets is:

\[
\max_{\{b\}} LL = \max_{\{b\}} \sum_{s \in S} N(C_s, A_{0s}) \cdot \prod_{j=1}^{n} \exp(\delta_j \cdot b_j), \sqrt{C_s}.
\]

where:
- \( \{b\} \) = set of calibration parameters corresponding to factors CF;
- \( S \) = set of indices of subsets \( s \);
- \( N(\cdot) \) = Normal probability density function;
- \( C_s \) = total number of crashes in calibration subset \( s \);
- \( A_{0s} \) = total number of crashes predicted in the calibration subset \( s \) with original SPF;
- \( \delta_j = 1 \) if calibration factor \( j \) applies to subset \( s \) and 0 otherwise; and
- \( b_j \) = calibration parameter in calibration factor function \( CF_j = \exp(b_j) \).
Although the optimization problem in Equation 8 can be solved in its original form, an equivalent optimization problem of minimizing the sum of squares weighted with $1/variance$ is preferred (Washington, et al., 2003, pp.128-132) for its computational convenience:

$$
\min_{\{b\}} \sum_{s \in S} (C_s - A_{0s} \cdot \prod_{j=1}^{n} \exp(\delta_{sj} \cdot b_j))^2 / C_s
$$

where:

- $\{b\}$ = set of calibration parameters corresponding to factors $CF$;
- $S$ = set of indices of subsets $s$;
- $C_s$ = total number of crashes in calibration subset $s$;
- $A_{0s}$ = total number of crashes predicted in the subset $s$ with non-calibrated SPF;
- $\delta_{sj} = 1$ if calibration factor $j$ applies to subset $s$ and 0 otherwise; and
- $b_j$ = calibration parameter in calibration factor function $CF_j = \exp(b_j)$.

The ordinary least square technique favors large observations. The $1/C_s$ weight reduces this tendency by considering the increase in variability of large crash counts. The solutions obtained for the optimization problems in Equations 8 and 9 are identical. The calibration solutions are presented in Tables 4 and 5 in the columns titles SL for Subsets. It can be seen that the number of crashes predicted with the new method better match the sums of recorded crashes. The Sum of Squared Errors (SSE) was calculated for each original and calibrated SPF based on recorded and predicted crashes aggregated in calibration subsets. The link-based calibration reduced the original SSE for the Injury/Fatal SPF by nearly 70 percent while the subset-based calibration reduced it by nearly 90 percent. The reduction in the original SSE for the PDO SPF was even more appealing. The link-based calibration did not reduce the original SS while the subset-base calibration reduced the original SS by more than 90 percent.

The match between the recorded and predicted crashes is perfect for subsets with only a single calibration factor not shared with other subsets. This situation occurs for three subsets: rural two-lane roads, rural multilane roads, and rural interstates. The calibration parameters can be calculated independently from other factors as
\[ CF = \frac{C_s}{A_{0s}}, \]  

where:

\( C_s \) = total crash counts for subset \( s \), and

\( A_{0s} \) = total original crash prediction for subset \( s \).

Table 4 Calibration Results for Injury/Fatal Safety Performance Functions

<table>
<thead>
<tr>
<th>Calibration Factor</th>
<th>ML for Links</th>
<th>LS for Subsets</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial</td>
<td>Optimal</td>
</tr>
<tr>
<td>Rural two-lane roads</td>
<td>1</td>
<td>1.15</td>
</tr>
<tr>
<td>Rural multilane roads</td>
<td>1</td>
<td>1.29</td>
</tr>
<tr>
<td>Rural interstates</td>
<td>1</td>
<td>0.77</td>
</tr>
<tr>
<td>Urban two-lane roads</td>
<td>1</td>
<td>2.14</td>
</tr>
<tr>
<td>Urban multilane roads</td>
<td>1</td>
<td>1.93</td>
</tr>
<tr>
<td>Urban interstates</td>
<td>1</td>
<td>0.48</td>
</tr>
<tr>
<td>Indianapolis area</td>
<td>1</td>
<td>1.01</td>
</tr>
<tr>
<td>Log-Likelihood LL</td>
<td>-16,429</td>
<td>-16,035</td>
</tr>
<tr>
<td>Total Weighted Square Residuals</td>
<td>41,361</td>
<td>2,846</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calibration subsets</th>
<th>Recorded Crashes</th>
<th>Original Prediction</th>
<th>ML for Links</th>
<th>LS for Subsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural two-lane roads</td>
<td>3,525</td>
<td>3,271</td>
<td>3,755</td>
<td>3,525</td>
</tr>
<tr>
<td>Rural multilane roads</td>
<td>770</td>
<td>681</td>
<td>878</td>
<td>770</td>
</tr>
<tr>
<td>Rural interstates</td>
<td>534</td>
<td>685</td>
<td>525</td>
<td>534</td>
</tr>
<tr>
<td>Non-Indianapolis urban two-lane roads</td>
<td>1,491</td>
<td>798</td>
<td>1,706</td>
<td>1,411</td>
</tr>
<tr>
<td>Non-Indianapolis urban multilane roads</td>
<td>1,614</td>
<td>1,247</td>
<td>2,400</td>
<td>1,749</td>
</tr>
<tr>
<td>Non-Indianapolis urban interstates</td>
<td>294</td>
<td>379</td>
<td>183</td>
<td>139</td>
</tr>
<tr>
<td>Indianapolis urban two-lane roads</td>
<td>104</td>
<td>78</td>
<td>167</td>
<td>155</td>
</tr>
<tr>
<td>Indianapolis urban multilane roads</td>
<td>849</td>
<td>419</td>
<td>812</td>
<td>661</td>
</tr>
<tr>
<td>Indianapolis urban interstates</td>
<td>698</td>
<td>1,855</td>
<td>900</td>
<td>764</td>
</tr>
</tbody>
</table>
Table 5 Calibration Results for PDO Safety Performance Functions

<table>
<thead>
<tr>
<th>Calibration Factor</th>
<th>ML for Links Initial</th>
<th>ML for Links Optimal</th>
<th>LS for Subsets Initial</th>
<th>LS for Subsets Optimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural two-lane roads</td>
<td>1</td>
<td>1.15</td>
<td>1</td>
<td>1.03</td>
</tr>
<tr>
<td>Rural multilane roads</td>
<td>1</td>
<td>1.23</td>
<td>1</td>
<td>0.95</td>
</tr>
<tr>
<td>Rural interstates</td>
<td>1</td>
<td>0.87</td>
<td>1</td>
<td>0.89</td>
</tr>
<tr>
<td>Urban two-lane roads</td>
<td>1</td>
<td>1.75</td>
<td>1</td>
<td>1.32</td>
</tr>
<tr>
<td>Urban multilane roads</td>
<td>1</td>
<td>1.81</td>
<td>1</td>
<td>1.22</td>
</tr>
<tr>
<td>Urban interstates</td>
<td>1</td>
<td>0.75</td>
<td>1</td>
<td>0.55</td>
</tr>
<tr>
<td>Indianapolis area</td>
<td>1</td>
<td>1.08</td>
<td>1</td>
<td>1.07</td>
</tr>
<tr>
<td>Log-Likelihood LL</td>
<td>-29058</td>
<td>-28676</td>
<td>&lt; -700</td>
<td>-218</td>
</tr>
<tr>
<td>Total Weighted Square Residuals</td>
<td>280598</td>
<td>257411</td>
<td>3118</td>
<td>347</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calibration subsets</th>
<th>Recorded Crashes</th>
<th>Original Prediction</th>
<th>ML for Links</th>
<th>LS for Subsets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rural two-lane roads</td>
<td>11,386</td>
<td>11,013</td>
<td>12,660</td>
<td>11,386</td>
</tr>
<tr>
<td>Rural multilane roads</td>
<td>2,301</td>
<td>2,423</td>
<td>2,979</td>
<td>2,301</td>
</tr>
<tr>
<td>Rural interstates</td>
<td>2,546</td>
<td>2,871</td>
<td>2,506</td>
<td>2,546</td>
</tr>
<tr>
<td>Non-Indianapolis urban two-lane roads</td>
<td>4,081</td>
<td>2,969</td>
<td>5,184</td>
<td>3,930</td>
</tr>
<tr>
<td>Non-Indianapolis urban multilane roads</td>
<td>4,538</td>
<td>4,044</td>
<td>7,335</td>
<td>4,920</td>
</tr>
<tr>
<td>Non-Indianapolis urban interstates</td>
<td>1,225</td>
<td>1,512</td>
<td>1,129</td>
<td>829</td>
</tr>
<tr>
<td>Indianapolis urban two-lane roads</td>
<td>284</td>
<td>274</td>
<td>517</td>
<td>390</td>
</tr>
<tr>
<td>Indianapolis urban multilane roads</td>
<td>2,329</td>
<td>1,370</td>
<td>2,684</td>
<td>1,791</td>
</tr>
<tr>
<td>Indianapolis urban interstates</td>
<td>3,283</td>
<td>5,994</td>
<td>4,834</td>
<td>3,532</td>
</tr>
</tbody>
</table>

The simple method (Eq. 10) was proposed for calibrating the SPF for two-lane rural roads and intersections in (Harwood et al., 2000). All the calibration parameters can be calculated this way in the standard calibration case. Adding user-defined sub-networks
entangles some of the calibration factors together which requires joint calibration of the calibration factors. The ML or LS formulation of the calibration problem should be used.

The re-defined LS calibration problem is more robust. It does not use probability functions that may sometimes cause computational troubles if the original model poorly fits the data (very small likelihood values). Finding a solution to the re-defined problem is much faster because the computational burden is smaller. The calculations are performed on a limited number of calibration subsets and not on thousands of network elements.

CONCLUSIONS

This paper presented two methods of calibrating models predicting safety for the individual links and nodes of a transportation network. The calibration has been defined to address the specifics of network modeling in transportation planning. In both methods, a planner has the freedom to partition a road network in a way that addresses expected local and sub-regional safety differences. Furthermore, a planner may identify routes, corridors, and areas to focus calibration on these locations if the planning focuses on them.

The maximum likelihood method was first applied to individual network elements. The results, although plausible and possible to explain, indicate considerable differences between the recorded and predicted total number of crashes at the sub-system level. The calibration task was redefined to focus on the prediction of the cumulative number of crashes in the user-defined sub-networks. This time, the results are much more reasonable.

This paper has demonstrated the feasibility of the proposed approach. A generalization of a simple calibration method proposed for individual safety performance functions, it can be classified as a moment-based method although in the general version it involves maximization of the likelihood of moment estimates. A robust version based on weighted least squares was proposed.
The proposed approach can be incorporated to the GIS-based planning suits such as TransCAD that allow developing add-on tools for crash-network assignment, calculation of the crash frequencies, and estimation of the calibration factors with the method presented in this paper. These tools are being developed for Indiana and they can be easily modified to suit needs of other jurisdictions. The add-on tools will be used jointly with the Travel Demand Model for predicting future traffic and safety.

The presented calibration method is focused on TIF crash models developed for individual roadway facilities. These models typically involve neither socioeconomic variables nor area-wide infrastructure characteristics. TAZ crash models suitable to evaluate transportation policies should complement TIF crash models. The calibration approach proposed here are applicable to TAZ crash models and to other transportation modes such as public transit and commercial vehicles.

Time-related calibration of the safety prediction methods is partly addressed by calibrating the models to crash data more recent than the data used to develop the original models. Nevertheless, proper consideration of a time dimension in long-term predictions should include anticipation of the future temporary trends in safety. This consideration is out of the scope of this presentation but should be considered for future research.

ACKNOWLEDGEMENT

This work was supported by the Joint Transportation Research Program administered by the Indiana Department of Transportation and Purdue University. The contents of this paper reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein, and do not necessarily reflect the official views or policies of the Federal Highway Administration and the Indiana Department of Transportation, nor do the contents constitute a standard, specification, or regulation.

The author would like to thank Mike Inerowicz for importing crash data and assigning crashes to the Indiana network segments in the TransCAD, which was essential for the evaluation of the proposed method.
LITERATURE


