SIIV - 5th International Congress - Sustainability of Road Infrastructures

Application of Highway Safety Manual to Italian divided multilane highways

Salvatore Cafiso*, Giacomo Di Silvestro, Giovanni Di Guardo

*Department of Civil and Environmental Engineering, School of Engineering, University of Catania, V. Andrea Doria 6 - 95125 Catania, Italy

Abstract

In the framework of the EU Directive on Road Infrastructure Safety Management, this paper will focus on ‘ranking of high accident concentration sections’, analyzing the opportunity to adopt the Highway Safety Manual (HSM) approach. With this goal, a pilot project of network screening for Italian motorways using HSM is presented. The more appropriate HSM procedures are selected and calibrated taking into account Italian national factors. Results are evaluated using as comparison Safety Performance Functions (SPF), developed directly with data for the specific Italian road infrastructure, which is assumed to provide more reliable estimates than HSM calibration.

© 2012 The Authors. Published by Elsevier Ltd. Selection and/or peer-review under responsibility of SIIIV2012 Scientific Committee

Keywords: Road safety, Network screening, HSM, Motorways.

1. Introduction

The new Directive 2008/96/CE of the European Parliament on Road Infrastructure Safety Management has been published on December 2008 and Member States have bring into force the laws, regulations and administrative provisions necessary to comply with this Directive by December 2011. The EU Directive aims to the establishment of procedures to ensure a consistently high level of road safety especially throughout the trans-European road network (TERN). TERN is the main road network connecting the European Countries composed above all by motorways and high quality roads which play an important role in long-distance traffic. In the framework of the actions considered in the EU Directive, this paper will focus on ‘ranking of high accident concentration sections’ considered as part of the safety management of the road network. In the EU Directive, ‘network safety ranking’ is defined as a method for “identifying, analysing and classifying parts of the existing road network, in operation for more than three years and upon which a large number of fatal crashes in proportion to the traffic flow, according to their potential for safety development and accident cost savings”.

* Corresponding author. Tel.: + 39 095 738 2213; fax: + 39 095 738 7913.
E-mail address: dcafiso@dica.unict.it
Even if the EU Directive defines a common approach to network safety management, its practical application is demanded to National guidelines to be adopted by Member States. In this context, the Highway Safety Manual (HSM) [1] approach for the evaluation of safety performance of Road network is very attractive. Among the fields of application of the HSM the “network screening” can be considered the equivalent of the directive task named ‘network safety ranking’. HSM describes network screening as composed by five major steps:
1. Establish Focus: Identify the purpose or intended outcome of the network screening analysis.
2. Identify Network and Establish Reference Populations.
4. Select Screening Method.
5. Screen and Evaluate Results.

The HSM assembles currently available information and methodologies on measuring, estimating and evaluating roadways in terms of crash frequency with special reference on north American state of the art and practice. To take into account, differences in factors, such as driver populations, roadway systems, traffic composition and traffic control measures, crash frequencies severity patterns and crash reporting systems, HSM contains calibration techniques to modify tools for local use. Also using the calibration process [2], applying the HSM outside the United States and Canada should be done with caution. In this case, the applicability and effectiveness of the calibration technique presented in the HSM should be properly evaluated.

In the paper a pilot project of network screening for Italian motorways will be presented. The more appropriate HSM procedures were selected and adapted to the EU Directive scope and were calibrated taking into account Italian national factors. Results are evaluated using as comparison Safety Performance Functions (SPF), developed directly with data for the specific Italian road infrastructure, which is assumed to provide more reliable estimates than HSM calibration. Also results obtained using naive approaches, such as simple crash count (observed crash), are analyzed and compared.

2. HSM calibration of rural four-lane divided model

The HSM safety predictive model consists of three parts, a Safety Performance Function (SPF) developed with respect to highway facility under given base conditions, the crash modification factors (CMF), and the calibration factor (Cr) computed in order to adapt the model to local conditions:

\[
N_{\text{predicted},i} = N_{\text{SPF},i} \times (\text{CMF}_{1i} \times \text{CMF}_{2i} \times \ldots \times \text{CMF}_{ij}) \times \text{Cr}
\]

where:
- \(N_{\text{predicted}}\) = predicted average crash frequency on the i-th site;
- \(N_{\text{SPF},i}\) = predicted average crash frequency for base conditions on the i-th site;
- \(\text{CMF}_{ij}\) = Crash Modification Factors related to the safety issue j at i-th site;
- \(\text{Cr}\) = calibration factor.

For each facility type, different SPFs and CMFs are taken into consideration. The paper deals with rural multilane highway segments consisting of divided four lane highways in rural area.

Different models are defined for total crashes, fatal/injury crashes and KAB crashes (KABCO scale where K=Killed; A=Incapacitating Injury; B=Non-Incapacitating Injury; C=Possible Injury; O=No Injury; and U=Injured, severity unknown). Among all, KAB classification better complies with the crash typology collected in the Italian National crash database. For KAB crashes, the HSM predictive model for rural divided roadway segments [1], is:

\[
N_{\text{SPF},ij} = e^{-8.505 + 0.874 \cdot \ln(AADT_{ij}) + \ln(L_i)}
\]

(2)
where:

- \( N_{spf, i} \): expected average crash frequency for the \( i \)-th roadway segment for the base conditions;
- \( \text{AADT}_{1,2} \): annual average daily traffic (vehicles per day) for the \( i \)-th roadway segment, both directions;
- \( L_i \): length of the \( i \)-th roadway segment roadway segment (miles);

The base conditions of the SPF for divided roadway segments on rural multilane highways are:

- Lane width (LW): 12 feet (3.65 m)
- Right shoulder width: 8 feet (2.45 m)
- Median width: 30 feet (9.15 m)
- Lighting: None
- Automated speed enforcement: None.

The effect of traffic volume (AADT) on crash frequency is taken into account through the SPF itself, while the effects of specific geometric design and traffic control features are accounted through the CMFs. Specifically, any feature associated with higher average crash frequency than the SPF base condition shows a CMF with a value greater than 1.00; any feature associated with lower average crash frequency than the SPF base condition shows CMF value less than 1.00.

The calibration factor allows to adjust the estimate to local state or geographic conditions, and is assessable through the following expression:

\[
C_r = \frac{\sum_{\text{sites}} \text{observed crashes}}{\sum_{\text{sites}} \text{predicted crashes}}
\]  

(3)

Moreover, the overdispersion parameter associated to the SPF is provided as a function of sites segment length as shown in the following equation:

\[
k_i = \frac{1}{e^{1.740 + \ln(L_i)}}
\]

(4)

where:

- \( k_i \): overdispersion parameter associated to the \( i \)-th roadway segment;
- \( L_i \): length of the \( i \)-th roadway segment (miles);

This parameter allows to integrate observed crashes data to the predictions of the model through the EB correction method.

### 2.1. Data and Sample segmentation

To apply the model performing the calibration procedure, the roadway must be divided into individual sites, consisting of homogenous roadway segments. The segmentation procedure produces a sample of roadway segments varying in length and homogeneous with respect to characteristics such as traffic volumes and key roadway design features (at least the same road features considered in the base model and CMFs: Segment length, Average annual daily traffic, Lane width, Shoulder width, Presence of lighting, Median width, Use of automated speed enforcement).

Motorway A18 Messina-Catania is a divided motorway part of TERN with two lanes for each direction (lane width = 3.75 m, right shoulder width = 3.00 m, median width = 3.00 m) and 6 intermediate interchanges. It has a total length of 76.85 km and the segmentation procedure has been performed considering all the features involved in the evaluation process. Segmentation process was carried out basing on the features proposed by the HSM method. Tunnels, interchanges and rest areas were excluded by the analysis. From the segmentation produced so far, 47 segments in common in the two carriageways and longer than 70 m were selected for a total length of 58.38 km. Their average length is 1.24 km, with a standard deviation of 1.77 km. Traffic volume in both directions resulted in the range 21,913 – 53,636 vehicles per day.
Crash data were obtained by reviewing police reports and National crash database reporting only fatal and injury crashes corresponding to KAB classification. Totally, 314 crashes occurred in the 4 years of analysis (2005-2008) were collected. The results of the calibration are listed in Table 1.

<table>
<thead>
<tr>
<th>Year</th>
<th>Total observed KAB crashes</th>
<th>Total Predicted KAB crashes</th>
<th>Calibration parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>75</td>
<td>61.68</td>
<td>1.21</td>
</tr>
<tr>
<td>2006</td>
<td>78</td>
<td>62.30</td>
<td>1.25</td>
</tr>
<tr>
<td>2007</td>
<td>90</td>
<td>62.95</td>
<td>1.43</td>
</tr>
<tr>
<td>2008</td>
<td>71</td>
<td>62.39</td>
<td>1.14</td>
</tr>
<tr>
<td>Total</td>
<td>314</td>
<td>249.45</td>
<td>1.26</td>
</tr>
</tbody>
</table>

The reader could note that the number of observed crashes per year is less than 100 as required by HSM. Anyway, it was assumed that the sample dimension, including all the A18 motorway, should not affect the results, but with reference for only A18 motorway.

3. SPFs development for an Italian motorway

With the bayesian approach [10], the measured parameter $x_i$ is considered a random variable. In a formulation of the problem based on the retrospective research of a change-point, while $x_1, \ldots, x_n$ are assumed to be independent random variables distributed according to the following function:

Application of the HSM model results will be compared to results carried out from two safety performance functions developed and calibrated purposely for the Italian motorway. Both of them were created using generalized linear modeling techniques and assuming a negative binomial distribution error structure [3,4,5,6].

A new segmentation was carried out to get the models. Differently from the previous one, this segmentation was produced keeping into consideration not only traffic data, but also geometrical features regarding horizontal and vertical alignment and cross section design, allowing further and deeper examinations. The road has slight different alignment layouts characterizing the two carriageways, therefore the analysis was performed considering independently each carriageway.

A total of 652 segments were defined. Their average length is 157.74 m, with a standard deviation of 89.73 m. The period of analysis was coincident to HSM’s (2005-2008); 308 severe crashes were connected to the sample sites.

The difference between the two new models is the variables accounted for. The former one is the so-called base model (B), which accounts only for the exposure variables (length $L$ and traffic AADT).

The SPF Base Italian model is the following:

$$N_{SPF,i} = e^{-18.52+1.17\cdot\ln(AADT_{i,1}) + \ln(L_i)}$$

(5)

where:

$N_{SPF,i}$: expected yearly crash frequency with Base model;
$L_i$: segment length (m);
$AADT_{i,1}$: average annual daily traffic one direction (vehicles per day).

The Base Model allows to compute what can be considered “normal” for a given site without considering differences in specific road or traffic features.

Considering that, for the road sample, all the CMFs in the HSM model are constant (CMFs=1), the model (B) represents the equivalent of the calibrated HSM model.

The latter model accounts for other explanatory variables changing in the sample, so that a more accurate and specific safety site analysis is possible.
As for multivariable modeling [7], many explanatory parameters were explored, including horizontal curvature, vertical alignment, cross section type and traffic volumes composition.

Several models were developed and evaluated depending on the statistical significance of the variables accounted for, on the model goodness of fit level and on the variables engineering congruity. Among the models evaluated, the best multivariable model selected is the one in the follow:

\[
N_{SPF,i} = e^{-19.67+1.26\cdot\ln(AADT_{i,1}) + \ln(L_i)+0.22\cdot\text{Curvi}+21.91\cdot\text{Gdi}}
\]  

(6)

where:

- \(N_{SPF,i}\) = predicted annual crash frequency with Multivariable model;
- \(L_i\) = segment length (m);
- \(AADT_{i,1}\) = average annual daily traffic one-direction (vehicles per day);
- \(\text{Curvi}\) = horizontal curvature \(1/R\) (km\(^{-1}\));
- \(\text{Gdi}\) = longitudinal downgrade slope (this parameter is equal to the absolute value of longitudinal slope in presence of a downgrade and equal to zero otherwise).

In order to verify the goodness-of-fit of the models both the Adjusted \(R^2_k\) and the Pearson \(\chi^2\) are compared with the value obtained from the \(\chi^2\) distribution at a p level of confidence. The goodness of fit measures for the two models are shown in Table 2.

<table>
<thead>
<tr>
<th>Variables significance (t-p)</th>
<th>(k)</th>
<th>Goodness of fit measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>COST</td>
<td>AADT</td>
<td>Curv</td>
</tr>
<tr>
<td>Base model</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Multivariable model</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

For comparison purposes, as HSM model (2) predicts crashes for both directions, it must be converted in one-directional SPF [8], as follows:

\[
N_{SPF,i} = 0.5\cdot e^{-8.505+0.874\cdot\ln(2\cdot AADT_{i,1}) + \ln(L_i)}
\]

(7)

where:

- \(N_{SPF,i}\) = expected average crash frequency one direction with the HSM model;
- \(AADT_{i,1}\) = annual average daily traffic one direction;
- \(L_i\) = length of the i-th roadway segment roadway segment (miles).

4. Performance measures selection and evaluation

Network screening is a process for reviewing a road network to identify and rank sites (usually called black spots) where the higher reduction in crash frequency, with implementation of appropriate countermeasures, is expected. Based on the results of the screening, only those sites that show potential for improvement are identified for additional analysis. Therefore, identification and ranking of such sites is the core of the network screening process. Any error or wrong identification and ranking will reflect on the whole process reducing the effectiveness of any safety policy. One of the major issues in network screening process is to select one or several performance measures. Generally, using multiple performance measures to evaluate each site may improve the results. To investigate how the HSM calibrated model works in the network screening for Italian motorways, and identify the potential application issues, results carried out using different SPFs and performance measures were compared. As first step the expected crash frequency on each segment in the last observation year was estimated using two different methods:

- Observed: the average crash frequency is estimated using only crash counts;
EB estimation: the observed average crash frequency and the predicted average crash frequency from a SPF are weighted together using the EB method. As SPF the calibrated HSM model(7), the Base model(5) and the Multivariable model(6) were separately used.

4.1. Data and Sample segmentation

For the site i, the estimation of the average crash frequency (OBS_i) in the last period of observation (year 2008) was carried out using crash counts from all periods assuming that the expected average crash frequency is not similar in all periods [1]:

$$OBS_i = \frac{\sum_{y=1}^{3} X_{iy}}{L_i \cdot \sum_{y=1}^{3} d_y}$$  \hspace{1cm} (8)

OBS_i = estimate of crash frequency in the last year of observation (2008);
X_{iy} = counts of accidents for site i at year y (2006, 2007, 2008)
L_i: segment length.
d_y = correction factor of the proportion of crashes in year y with respect to the last year

4.2. Expected Average Crash Frequency with Empirical Bayes Adjustment (CFEB)

The observed average crash frequency and the predicted average crash frequency from a SPF are weighted together using the EB method to calculate an expected average crash frequency that accounts for regression to the mean (RTM) bias. If 3 years of observations are available, the expected average crash frequency with empirical Bayes adjustment in the first year NEB,1 for site k is obtained from [1]:

$$N_{EB,1k} = w_i \cdot N_{SPF,1i} + (1 - w_i) \cdot \left( \frac{\sum_{y=1}^{3} X_{iy}}{\sum_{y=1}^{3} C_y} \right)$$  \hspace{1cm} (9)

with
N_{SPF,1} : crash frequency predicted by the SPF for the first year;
w: weight factor
C_y : adjustment factor for year y

$$w_i = \frac{1}{1 + k \cdot \sum_{y=1}^{3} N_{SPF,1y}}$$  \hspace{1cm} (10)
Finally, the expected average crash frequency with empirical Bayes adjustment for the last year of observation (EB) is:

$$\text{EB}_i = N_{\text{EB},1i} \cdot C_{3i} \ [\text{crash/year/km}]$$  (12)

$N_{\text{SPF},1}$, in the different simulations for comparison, was alternatively carried out from the calibrated HSM model one-direction (7), the SPF base model (5) or the SPF multivariable model (6).

The value of crash frequencies $(OBS_i, (11), \text{EB}_i(15))$ were used to rank the sites in the CF method.

Based on the estimation of the crash frequency, two more performance measures were added to identify and rank black spots, the Critical Crash Rate (CCR) and the Excess Expected Average Crash Frequency (ECF). These additional performance measures are able to take into account in the classification the traffic volume as required by the EU Directive.

5. Critical Crash Rate (CCR)

The crash rate $(\text{CR}_i)$ at each site $i$ is compared to a calculated critical crash rate $(\text{CR}_{ci})$ that is unique to each site.

$$\text{CR}_i = \frac{Y_i}{\text{MEV}_i} \ [\text{crash/10}^6 \text{ vehic. km}]$$  (13)

$Y_i$: expected crashes for site $i$.

Basing on the method used to estimate the crash frequency, $Y_i$ is computed as

- OBS) $Y_i = OBS_i \cdot L_i \ [\text{crash/year}]$
- EB) $Y_i = \text{EB}_i \cdot L_i \ [\text{crash/year}]$

$\text{MEV}_i$: exposure factor = $365 \cdot \text{AADT}_i \cdot L_i \cdot 10^{-6}$

$L_i$: segment length ; $\text{AADT}_i$: Annual Average Daily Traffic for segment $i$.

The critical crash rate is a threshold value that allows for a relative comparison among sites. The critical crash rate depends on the average crash rate at similar sites $(\text{CR}_{av})$, traffic exposition at site $i$ $(\text{MEV}_i \ [10^6 \text{ vehicle per km}])$, and a statistical constant $(P)$ that represents the standard normal variable at the $p^\text{th}$ percentile [10].

$$\text{CR}_{ci} = \text{CR}_{av} + P \sqrt{\frac{\text{CR}_{av}}{\text{MEV}_i}} + \frac{1}{2 \cdot \text{MEV}_i}$$  (14)

$$\text{CR}_{av} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} \text{MEV}_i}$$  (15)

$n$: number of sites composing the sample.

From equation (17) when posing $\text{CR}_{ci} = \text{CR}_i$ the $P_i$ value and the corresponding level of significance can be calculated for each site $i$. For example, if a value of $P=1.1$ is obtained, from the standard normal distribution a level of significance equal to 0.86 can be evaluated. The level of significance of each site was used to rank the sites in the CCR method. Higher the level of significance is, higher results the rank position.
6. Comparison of results and ranking

The sample of road segments used to calibrate the Base and the Multivariable SPFs, was adopted for the comparison study. The 152 sites composing the sample, were ranked in descending order according to each performance measure calculated using the average observed crash frequency (OBS) and the expected crash frequencies with EB adjustment (EB) obtained from the three different SPFs (HSM calibrated, Italian Base, Italian multivariable).

The top 23 sites, corresponding to 15% of the overall sample, resulting from the rank based on the expected crash frequencies EB obtained from the Italian multivariable SPF, were selected as reference group.

For each performance measure, to compare results the Normalized Root Mean Square Error (RMSE) of the performance parameter estimation and the squared differences between the expected rank and the observed one (R2) of the ranking were computed with reference to the Multivariate model results.

RMSE is measure of the differences between values predicted by two different models. Here, RMSE is a measure of precision in the performance parameter estimation. The better the rate fits the reference data of the Multivariable model, the closer the value of R2 is to one. As in the linear regression, R2 provides a measure of how well the rate outcomes are likely to be predicted by the model. Here, R2 is used as a simple measure of concordance in the rank estimation. Table 3 shows the rank comparison carried out using different performance measures and crash models.

<table>
<thead>
<tr>
<th>CF</th>
<th>CCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBS</td>
<td>SPF_B</td>
</tr>
<tr>
<td>False positive</td>
<td>R²</td>
</tr>
<tr>
<td>5</td>
<td>0.61</td>
</tr>
<tr>
<td>1</td>
<td>0.90</td>
</tr>
<tr>
<td>3</td>
<td>0.80</td>
</tr>
</tbody>
</table>

The differences in ranking are, also, represented in Figure 1 and Figure 2.

---

Fig. 1. CF Ranking comparison between SPF_M and other models
With reference to the use of CF as performance measure, Table 3 shows that using the observed crash frequency (OBS CF) a RMSE of 12% with respect to the average, can be expected in the estimation of the crash/year/km in the top rank sites. Looking at the ranking carried out using the observed CF (OBS), Figure 2 shows the highest dispersion of results with 5 sites that were not included in the top 23 (marks above the dashed line). Therefore, it can be concluded that 5 sites (about 22%) were classified as False Positive (a site erroneously classified as black spot) and an equal number of sites are missed as black spot (False Negative) in the ranking. Among the 23 sites ranked at the top by the two SPFs only one from SPF_Base (EB_Base) and three from HSM (EB_HSM) are not included in the top ranked by SPF_M. These sites are identified in Figure 2 by the marks positioned above the dashed line of the 23rd rank. The better ranking agreement of CF obtained from SPF_Base and HSM is represented in the figure by the gathering of the marks close to the diagonal line which represents the perfect concordance and by the value of R2 higher than 0.80 (Table 3).

Using CCR as performance measure gives similar results (Table 3, Figure 2) and trends.

7. Conclusions

It must be pointed out, that network can be prioritized using different performance measures, but one method to calculate the performance measures has to be selected. Using multiple performance measures to evaluate each site improves the level of confidence in the results, while the key considerations in selecting the method are data availability and regression-to-the-mean bias. Also knowledge and technical background of the transportation professionals in roadway safety management techniques and crash estimation methods can influence the choice of the method to apply.

Performing the network screening, both errors in ranking and in the estimation of the Performance measures produces failures. The error in ranking gives as consequence an incorrect selection of sites to be considered for safety improvements (i.e. some black spots are missed and erroneous sites are selected as black spots). Moreover, errors in quantifying the Safety Performance are reflected in bad evaluation of the expected benefits due to the crash reduction. Table 3 summarizes the results in terms of False positives (erroneous sites selected as black
spots), $R^2$ and RMSE for each Performance measure (CF, CCR) and method (OBS, SPF_B, SPF_HSM) compared in this study.

Considering both precision in the performance measure estimation and in the ranking positioning, this study confirms (Table 3) that developing SPFs directly with data for the specific road network produces the best results, but also the use, in Italy, of the calibration procedure allows to adopt the HSM model for the network screening of multilane divided road network. For network screening of motorway the development of multivariate SPF despite a notable increase in analysis effort appears doesn’t produce sensible improvements in the results when compared to the use of simpler SPF based on traffic volume and segment length only, or at least to the calibration of the HSM model.

As expected the use of observed crash frequency without EB correction, produces always bad results when compared with the ones derived by the other methods.

Results carried out in the present research could be generalized for motorway network in Italy, more caution and studies has to be adopted for the extension of the results to other road facility types included in the HSM, i.e. Rural two lane roads and Urban and Suburban arterials. The suggestion for greater prudence and experimental effort derives not only by the obvious differences in the road features but also by the difference in the exposure that characterizes motorways. The relatively higher traffic volumes of motorways makes the crash observations, for this kind of roads, particularly reliable if compared to other road networks. Therefore, the role of the SPF quality and EB adjustment in reducing RTM bias is less influential in the screening of motorway network than for other road networks characterized by fewer and more scattered crashes [11].

Acknowledgement

The authors express their gratitude to the east Sicily Highway Police Department and to the Management of Sicilian motorways (CAS) for their help in getting the data used for the analysis.

References