Proceedings of 6th Transport Research Arena, April 18-21, 2016, Warsaw, Poland

Identification of hazardous locations in regional road network – comparison of reactive and proactive approaches

Jiří Ambros a,*, Pavel Havránek a, Veronika Valentová a, Zuzana Křivánková a, Radim Striegler a

aCDV – Transport Research Centre (Centrum dopravního výzkumu, v.v.i.), Lišeňská 33a, 636 00 Brno, Czech Republic

* Corresponding author. Tel.: +420 541 641 362; fax: +420 541 641 712.
E-mail address: jiri.ambros@cdv.cz

Abstract

The study objective is to compare three approaches to identification of hazardous road locations: (1) Traditional reactive accident-based approach, resulting in identification of accident black spots; (2) State-of-the-art empirical Bayes method using accident prediction model, which identifies critical locations, i.e. both real and potential black spots; (3) Proactive “preliminary” road safety inspection, identifying the risk factors, which may potentially increase accident occurrence and severity. Regional rural road network (approx. 1000 km) in South Moravia, Czech Republic was used. The methods were applied in identification and ranking of hazardous road locations in the studied network. It was found that black spot approach is not a suitable method, especially in low-volume road network with scattered accident occurrence. On the other hand risk index, based on road safety inspection, is a valid alternative, with ranking performance comparable to state-of-the-art empirical Bayes method. In addition both empirical Bayes and risk index methods are compared with respect to their data requirements. Their mutual application is recommended as a suitable replacement of traditional black spot management, as well as a step forward to a proactive road network safety management.

Keywords: Road safety; hazardous locations; black spot; empirical Bayes; risk index

1. Introduction

In recent years, European road safety orientation has been steered by European Directive 2008/96/EC on road infrastructure safety management (European Commision, 2008). It introduces four procedures, including road safety inspection and network safety ranking, which will be both further referred to in the text. Road safety inspections (RSIs) are defined as “an ordinary periodical verification of the characteristics and defects that require maintenance work for reasons of safety”. It is a preventive (proactive) tool, non-accident based, relying on subjective assessment of relevant
At the same time, it is recognized that most of safety issues are concentrated on regional roads – however due to lower traffic volumes, rural hazardous locations are less clustered, which complicates their identification (OECD, 1999; Gatti et al., 2007; Polidori et al., 2012). Nevertheless the Directive states that “such inspections shall be sufficiently frequent to safeguard adequate safety levels for the road infrastructure in question”, without providing any guidance how to efficiently conduct RSIs on large road networks.

Another procedure is network safety ranking (NSR) – a method for identifying, analysing and classifying parts of the existing road network (based on accident concentrations) according to their potential for safety development. Again no specific details are specified in the Directive, for example regarding the difference between NSR and traditional black spot management. The former approach is still mostly used – however its dependence on statistically rare historical accident data, with underlying random fluctuations, may lead to incorrect identifications. The identified and subsequently treated locations thus may not be the most hazardous ones, and true hazardous ones may further remain untreated. It is therefore desirable to attempt using approaches which are statistically more reliable and enable creating a priority list of road sections where an improvement of the infrastructure is expected to be highly effective.

In order to fulfill these requirements the empirical Bayes (EB) approach using accident prediction models has been recommended (Hauer et al., 2002; Elvik, 2008a; Montella, 2010).

The Directive requires that the road sections, identified during NSR, are evaluated by experts, using RSIs. This is where both procedures meet – a proactive (non-accident based) RSI is applied on sites, which were identified through reactive (accident-based) NSR. Alternative methods, using risk indices, present a combination of both approaches. Risk index reflects exposure (traffic volume and section length), as well as factors related to accident probability and potential severity consequences – they provide a proactive approach to network safety ranking. As surrogate safety measures, risk indices are especially attractive for low-volume regional rural roads with decreasing and scattered accident occurrence; for examples see de Leur and Sayed (2002), Montella (2005) or Cafiso et al. (2011).

However it is up to member states to implement their own procedures for practical conduct of all the Directive procedures. Applications of RSI in several European countries (Austria, Ireland, Norway) are reported in Cafiso et al. (2014). In the Czech Republic, RSI guidelines (CDV, 2013) have been developed according to recommendations of the Directive. In order to being able to cover large regional road networks, a “preliminary” RSI is firstly conducted (Pokorný and Striegler, 2014). The objective of this RSI is to review the given network and assess the risk of all sections for their prioritization. The top hazardous sections are to be subsequently subjected to on-site RSIs. Regarding NSR, traditional black spot management (Andres et al., 2001) is still commonly applied in the Czech Republic (Bartoš et al., 2015); applications of state-of-the-art EB methodology have only appeared recently (Valentová et al., 2014; Ambros et al., 2015; Ambros and Peltola, 2015; Ambros et al., 2016). EB estimates combine historical accident data with accident predictions, and are thus able to identify also potential hazardous road locations, where no accidents have yet occurred – the process thus provides both proactive and reactive perspective.

Given the amount of different approaches to identification of hazardous road locations, a comparative study in the Czech context is needed. The paper presents such study, conducted on regional rural road network in South Moravia, Czech Republic. The network consisted of two-lane undivided roads (including both sections and intersections), with approximate total length of 1000 km. For identification three approaches have been used:

1. Traditional approach based on accidents only, resulting in identification of accident black spots.
2. Empirical Bayes method using accident prediction model, which identifies critical locations, i.e. both real and potential black spots.
3. Proactive “preliminary” road safety inspection, identifying the risk factors, which may potentially increase accident occurrence and severity.

The study presents all the mentioned methods, applies them in identification and ranking of hazardous road locations in the studied network and compares (validates) their results. In addition the methods are compared with respect to their data requirements. Road administrators may use the results in order to decide on the most efficient approach, or potentially their combination within the life cycle of regional road network safety management.

2. Data and methods

The complete regional road network (2nd class roads) of South Moravia amounts to approx. 1500 km of length. After exclusion of urban sections, the length is approx. 1000 km. The network was segmented between the settlements,
so that each segment (including intersections) covers one link. In total there were 372 segments. The typical road section is paved, two-lane and undivided (single carriageway).

The segments were assigned accident frequency in 6 years (2009 – 2014), in all severity categories (property-damage only, slight/severe/fatal injury). Additional variables were selected based on previous modelling efforts on the same network (Valentová et al., 2014; Ambros et al., 2015; Ambros and Peltola, 2015; Ambros et al., 2016), see Table 1:

- Exposure variables: traffic volume (AADT) from National Traffic Census 2010 and segment length.
- Curvature change rate (CCR), as the alignment consistency indicator, computed as sum of angular changes divided by segment length.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min / max / mean / std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-year frequency of recorded accidents (R)</td>
<td>0 / 129 / 10.01 / 13.62</td>
</tr>
<tr>
<td>Traffic volume [veh/day] (AADT)</td>
<td>90 / 15000 / 2336.87 / 1993.42</td>
</tr>
<tr>
<td>Segment length [km] (L)</td>
<td>0.50 / 13.00 / 2.61 / 1.80</td>
</tr>
<tr>
<td>Curvature change rate [gon/km] (CCR)</td>
<td>0.00 / 833.21 / 118.00 / 103.92</td>
</tr>
</tbody>
</table>

Further data are mentioned in the following text. It describes the three methods (traditional black spot management, empirical Bayes approach, preliminary road safety inspection). Their results will be subjected to comparison, which will be also described.

2.1. Traditional black spot management (BSM)

According to the Czech guidelines (Andres et al., 2001, following FVS, 1990), black spot identification is based on a recorded accident frequency in a sections defined by sliding window of 250 m length, which meets a condition of at least three injury accidents in one year. Black spot analysis is processed each year, using accident data in overlapping 3-year periods, with results presented on web portal http://infobesi.dopravniinfo.cz/. Following variables were retrieved for all segments for time periods 2009 – 2011 and 2012 – 2014:

- number of black spots
- total accident frequency in these black spots (AF)

2.2. Empirical Bayes approach (EB)

Accident prediction model of frequency of all accidents (both property-damage-only and injury) was developed, using explanatory variables AADT, length and CCR. The model form was consistent with state-of-the-art:

\[ P_i = \beta_0 \cdot AADT_i^{\beta_1} \cdot L_i^{\beta_2} \cdot \exp(\beta_3 \cdot CCR_i) \]  

i.e. for each segment \( i \) expected (predicted) accident frequency \( P_i \) is estimated via multiplicative regression model, including explanatory variables AADT, length and CCR; \( \beta_i \) are regression coefficients to be estimated in modelling. Empirical Bayes estimates of expected accident frequency were then calculated:

\[ EB_i = w_i \cdot P_i + (1 - w_i) \cdot R_i \]  

\[ w_i = \frac{k_i}{k_i + P_i} \]
\[ k_i = \frac{k}{l_i} \]  \hspace{1cm} (4)

where \( EB_i \) are EB estimates, obtained as weighted average of predicted and reported accident frequencies \( (P_i \text{ and } R_i) \). Weights \( w_i \) are calculated using length-dependent overdispersion parameter \( (k_i) \) (Hauer et al., 2002).

### 2.3. Preliminary road safety inspection (RSI)

Preliminary RSI is based on data collected by instrumented vehicle. During driving through the studied network, road safety auditor records the presence of risk factors, which may potentially increase accident occurrence and severity. For each risk factor, following features were registered and evaluated:

- risk type (horizontal or vertical curve, road signing or marking, shoulder, access, vegetation, etc.)
- forgivingness (7-point scale, based on roadside hazard rating by Harwood et al., 2000)
- pavement quality (3-point scale)

Based on these assessments, risk severity grade on a 3-point scale is assigned to each risk factor (see Table 2). For each segment, risk severity factors are summed up and multiplied by AADT. In order to control for different segment lengths, the result is divided by length.

\[ r_{iri} = (\sum r_{os} r_{esi} \cdot \frac{A}{L}) \]  \hspace{1cm} (5)

Preliminary RSI was conducted in both driving directions, therefore the final risk index \( (RI) \) is a sum of two risk indices.

<table>
<thead>
<tr>
<th>Potential influence on:</th>
<th>Risk severity</th>
<th>- accident occurrence</th>
<th>- accident severity</th>
<th>RSI recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>low</td>
<td>may lead to traffic conflicts</td>
<td>minimal influence</td>
<td>none</td>
<td></td>
</tr>
<tr>
<td>medium</td>
<td>increases accident probability</td>
<td>may increase accident severity</td>
<td>should be eliminated</td>
<td></td>
</tr>
<tr>
<td>high</td>
<td>significantly influences accident probability</td>
<td>significantly increases accident severity</td>
<td>must be eliminated</td>
<td></td>
</tr>
</tbody>
</table>

### 2.4. Comparison

Each of the methods (BSM, EB, RSI) yields values, which will be used for comparison and ranking:

- BSM identification is dichotomous (segment either is or is not a black spot), which does not lend itself for comparison of continuous rankings. Therefore frequency of accidents on identified black spots \( (AF) \) was used as a continuous indicator. According to BSM principle, each black spot has equal length 250 m.
- EB method results in EB estimates. Given the non-uniform length of segments, the estimates were divided by segment length \( (EB/L) \). The same approach was applied by Montella (2005) and Cafiso et al. (2011).
- RSI produces risk indices \( (RI) \). The values of risk index are by definition (equation 5) divided by length, which makes them comparable across different segment lengths.

Cafiso et al. (2011) state that to effectively use RSIs (i.e. risk indices), the procedure must satisfy several objectives, including: (1) it must produce a safety evaluation correlated with accident history, (2) it must rank safety problems and rankings must be consistent. These two exercises will be further referred to as comparison of (1) level of safety, and (2) safety ranking. EB estimates \( (EB/L) \) will be used as a standard, against which BSM and RSI results \( (AF \text{ and } RI) \) will be validated. For level of safety comparison, Pearson correlation coefficients will be used.
The second comparison is conducted based on rankings of the lists of identification numbers of segments. These numbers are ranked according to decreasing values of criteria from the tested methods. The ranking based on $EB/L$ and the ranking based on $RI$ are then compared using the method consistency test (Cheng and Washington, 2008; Montella, 2010). Its premise is that a list of segments, which were identified as hazardous by the tested method, should be similar to the list of segments identified by the standard method. For a comparison the test criterion is following: the tested method, which identifies segments with the largest overlap with the list of segments by the standard method, is more consistent. In practice various samples of ranked segments may be selected (for example top 5% of distribution), according to available budget of a road administrator. For illustration, selections between top 20 and 100 segments were used (corresponding to approx. 5% to 25% upper tails of the whole studied network). However this method of ranking comparison is suitable for continuous variables ($EB/L$ and $RI$); this is not the case of $AF$, which is ordinal, with excessive zero values (more than 80% of segments). Alternative approach was therefore used: overlapping percentage was calculated from a number of black spots, which were identified within each of 5% to 50% upper tails of $EB/L$.

The third comparison uses epidemiological diagnostic criteria. The idea is that the tested method should identify as many of the truly hazardous segments as possible (sensitivity), and as few of truly non-hazardous segments as possible (specificity). Following Elvik (2008b), the segments, identified as hazardous by ranking based on 6-year accident frequency, were considered truly hazardous. In order to compare all three criteria, following sets were used:

- segments ranked by $AF$ from two time periods (2009 – 2011 and 2012 – 2014)
- segments ranked by $EB/L$ from two time periods (2009 – 2011 and 2012 – 2014)
- segments ranked by $RI$

The rankings were again tested for top 20 to 100 segments (5% to 25% upper tails). The results are presented in terms of numbers of positives and negatives. For explanation see Table 3.

<table>
<thead>
<tr>
<th>Is the segment identified using…</th>
<th>… 6-year accidents?</th>
<th>… 3-year data?</th>
<th>Then the segment is a…</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>yes</td>
<td></td>
<td>true positive (TP)</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
<td></td>
<td>false positive (FP)</td>
</tr>
<tr>
<td>no</td>
<td>no</td>
<td></td>
<td>true negative (TN)</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td></td>
<td>false negative (FN)</td>
</tr>
</tbody>
</table>

Diagnostic criteria of sensitivity and specificity are computed using following formulas:

\[
sensitivity = TP / (TP + FN) \tag{6}
\]

\[
specificity = TN / (FP + TN) \tag{7}
\]

According to Elvik (2008b), a good diagnostic test performs well in relation to both sensitivity and specificity; a trade-off has to be made. Therefore sum of both indicators was used to assess the performance of each ranking criteria.

3. Results and discussion

Data for comparison ($AF, EB, RI$) were collected according to the procedures described above (equations 2 and 5). Their descriptive characteristics are presented in Table 4.
Table 4. Descriptive characteristics of data used for comparison.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min / max / mean / std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-year accident frequency recorded in identified blackspots (AF)</td>
<td>0 / 65 / 1.94 / 5.91</td>
</tr>
<tr>
<td>6-year empirical Bayes estimate of expected accident frequency (EB)</td>
<td>0 / 129 / 10.01 / 13.58</td>
</tr>
<tr>
<td>Risk index (RI)</td>
<td>0 / 53 / 4.03 / 5.88</td>
</tr>
</tbody>
</table>

As previously mentioned, both indicators $AF$ and $RI$ were validated against $EB/L$. Pearson correlation coefficients were 0.46 for $AF$ and 0.52 for $RI$. Both values are close, while also relatively low. Nevertheless such low correlations are not unique: in previous studies even lower coefficients (approx. 0.3) were obtained (de Leur and Sayed, 2002; Montella, 2005). On the other hand, Cafiso et al. (2011) reached a correlation coefficient around 0.9. However their studied sample was considerably smaller than what was presented in this study (30 segments of total length 100 km); the authors also excluded intersections, which may improve homogeneity and correlation.

Consistency (in terms of percents of overlap with the ranked list based on $EB/L$) was further evaluated, using method consistency test (for $RI$) and percents of black spots (for $AF$), for top 20 to 100 segments (i.e. upper tails between approx. 5% to 25%) – see Fig. 1. The results of epidemiological diagnostic test are presented in Figure 2 (3-year time periods are labelled 3a and 3b).

![Fig. 1. Results of consistency tests for AF and RI (percents of overlap with the ranking based on EB/L).](image1)

![Fig. 2. Results of epidemiological tests for AF, EB/L and RI (sum of sensitivity and specificity).](image2)
In Fig. 1 consistency of identification using $RI$, with respect to $EB/L$, is between 70% and 80% for all variants of upper tails. This is comparable to what was achieved in other studies, such as Sacchi et al. (2015). On the other hand consistency using $AF$ is constantly decreasing to a level of 40% only.

In Fig. 2 diagnostic performance is almost similar for both variants of $EB/L$ and $RI$, while it is again lower for $AF$. The values are roughly comparable to Elvik (2008b), who reported approximate sums 1.5 and 1.7 for accident counts and EB estimates, respectively.

In both cases performance of $AF$ was inferior. It may be caused by several factors, which are inherent to black spot management approach:

- In principle the result of black spot investigation is dichotomous: segment either is or is not a black spot. This complicates the comparison with other criteria, which are in principle continuous.
- In order to circumvent this issue, ordinal criterion of total accident frequency in identified black spots was used. However majority of segments (more than 80%) had zero accident frequency.
- Total black spot accident frequency amounts to 18.5% of all accidents, and total length of black spots covers approx. 3% of the network. It is close to values of 20% accidents and 3% length, which were obtained also in other Czech studies (Bartoš et al., 2015). However it means that 97% of road network is not evaluated by this criterion, which again makes it incomparable to other continuous indicators.

The consistency test results confirm that identification of black spots is not suitable for low-volume roads where accidents are less clustered, as stated in the introduction. Low accident frequencies also limit possibilities of following analyses: according to Schermers et al. (2011), at least 5 accidents are needed for a meaningful analysis of common accident patterns and factors. However in the studied sample, on average only 2 accidents per segment were registered in a 6-year period (see Table 4).

Another dimension of consistency is temporal stability of true hazardous locations, which should not be influenced only due to random variations; therefore under unchanged conditions of exposure and infrastructure (i.e. no systematic variation), they should remain stable. This hypothesis was tested in previous study on the same network, using three overlapping 3-year periods and comparing traditional black spot criterion with EB identification (Valentová et al., 2014). The cases identified repeatedly in all 3 time periods amounted to 36% and 50% of all locations in the two methods, respectively. Therefore using EB approach one is more likely to identify true hazardous locations.

4. Conclusions

The objective was to compare three approaches to identification of hazardous locations in regional road network: traditional black spot criterion (AF), empirical Bayes estimates (EB), and risk index (RI) from preliminary road safety inspection. Two ranking comparisons were conducted:

- **Relative comparison**: Since EB approach is considered a state-of-the-art technique, it was used as a standard against which AF and RI rankings were validated. RI performed better than AF.
- **Absolute comparison**: Diagnostic performance of all three criteria was assessed using epidemiological test. EB and RI results were almost similar, while AF performance was again inferior.

The findings are consistent with previous studies: black spot approach, based on AF, is unsuitable safety ranking method, especially in low-volume road network with scattered accident occurrence. In these conditions risk indices, as proved in the study, provide an attractive alternative. Risk index enables a proactive assessment, uninfluenced by random accident variations. Its ranking performance was found comparable to EB method, which controls for random variations and thus provides adjusted estimates of expected accident frequency.

Both EB and RI may then be recommended for identification of hazardous locations in regional road networks. Regarding choice of accident prediction models for this application of EB method, it was recently indicated that simple variants (involving only exposure variables) are sufficient (Srinivasan et al., 2013) and previous studies on the same network (Ambros et al., 2015; Ambros and Peltola, 2015; Ambros et al., 2016) confirmed this fact. Based on equations 1 and 5, following variables are needed: AADT, length and CCR for EB estimation; AADT and length for risk index.
From the perspective of data requirements, a review of Candappa et al. (2011) recognizes three levels: low, intermediate, high. They see network screening (EB) as intermediate; in general, no new data are collected specifically for this purpose. For accident modelling new data may be collected (such as CCR in this study), but it is more often the case that data from road data banks are used. Road protection scoring (alternative of risk index) ranks intermediate or high in data requirements; it relies on observation, video or other instrumentation.

The requirements are thus relatively comparable. Nevertheless it is not necessary to choose one method or another; a combination may be useful as well. According to Sørensen and Elvik (2007), a gradual transition from black spot management to network safety management is “presumably the most relevant situation for most of the European countries now or within a short period of time”. In this perspective application of EB methodology and risk index for identification of hazardous locations in regional road networks will be a suitable replacement of traditional black spot management, and a step forward to a proactive road network safety management.

Acknowledgements

Appreciated is advice of Salvatore Cafiso on risk index validation, as well as help of Jan Novák and Robert Zůvala with consistency tests algorithms. The study was conducted with support of the Czech Ministry of Interior’s research project No. VG20112015013 “Identification and treatment of high risk road spots and sections” (IDEKO).

References


FVS (Forschungsgesellschaft für das Verkehrs- und Straßenwesen), 1990. Allgemeines / Verkehrssicherheit / Verkehrssicherheitsuntersuchung [General / Road safety / Road safety investigation]. Merkblatt RVS 1.21. FVS, Vienna, Austria. (in German)


