Netw Spat Econ (2006) 6: 97–110
DOI 10.1007/s11067-006-7695-2

# A Comparative Study of Alternative Model Structures and Criteria for Ranking Locations for Safety Improvements

**Luis F. Miranda-Moreno · Liping Fu**



**Abstract** In transportation safety literature, many statistical models and criteria have been proposed for quantifying risk at transportation facilities such as roadway intersections and highway-rail grade crossings, and identifying candidate locations, or blackspots, for engineering improvements. There are, however, few systematic studies on the comparative performance and practical implications of these models and criteria. The primary goal of this study is to investigate the relative impacts that the use of various alternative models and ranking criteria can have on identifying blackspots. Three alternative models are considered in this investigation, including the popular negative binomial model (NB), the heterogeneous negative binomial model (HNB), and the zero inflated negative binomial model (ZINB). The expected accident frequency based on both marginal distribution and posterior distribution is considered as a ranking criterion. A sample of highway–railway grade crossings located in the Canadian railway network is used in this investigation as an application environment.

**Keywords** Blackspot identification · Empirical Bayes approach · Zero-inflated models

## 1. Introduction

One of the first tasks in developing a safety improvement program for a set of transportation facilities (e.g., road sections, roadway intersections, highway-rail grade crossings, etc.), is the identification of a list of locations that show evidences

L. F. Miranda-Moreno (✉) · L. Fu
Department of Civil Engineering,
University of Waterloo, Waterloo, ON, Canada
e-mail: {lfmirand; lfu}@uwaterloo.ca

L. F. Miranda-Moreno
Instituto Mexicano del Transporte, Querétaro, México

of high accident risk and warrant for further engineering inspections. These locations are commonly referred to as hazardous locations or blackspots and are considered as the candidates for cost-effective remedial actions such as installation of new control devices and improvement of location geometry (Schluter et al., 1997; Heydecker and Wu, 2001).

Hazardous locations could be identified by ranking all the locations according to their accident history or raw accident rates (e.g., number of crashes per vehicle entries in intersections). However, this approach can be very sensitive to temporal variations because of the randomness and rarity of accident events. For instance, a location that experienced a high number of accidents in one period of time may not register accidents in the following periods. Furthermore, accident frequency at specific locations may tend to the general accident mean. This phenomenon is known as *regression to the mean* (Hauer, 1997; Schluter et al., 1997; Miaou and Song, 2004).

Instead of using a simplistic approach based on the raw accident rates, a more commonly accepted method for blackspot identification is the model-based approach. This method consists in ranking locations according to one or more ranking criteria that can be computed using a regression model. Among the most popular models applied in road safety analysis are the standard Poisson and negative binomial (NB) models (e.g., Miaou, 1994; Austin and Carson, 2002; Saccomanno et al., 2004). Extensions of these two models are the zero inflated Poisson model (ZIP) and zero inflated negative binomial model (ZINB) which have also been utilized for modeling accident data (Miaou, 1994; Shankar et al., 2003). More recently, other random effect or Bayesian models have been proposed to deal with long-term trends and/or spatial correlation (e.g., Lord and Persaud et al., 2000; Miaou and Song, 2004).

Among the ranking criteria, the expected number of accidents based on the marginal distribution of a given model (e.g., NB model) has been utilized for blackspot identification (Saccomanno et al., 2004). Using Bayesian analysis, other ranking criteria derived from the posterior distribution have been also proposed in the literature for this task. Example criteria are the posterior mean of accident frequency, the potential of accident reduction, the posterior probability that the accident frequency exceeds a specific value and the posterior expectation of ranks (e.g., Persaud et al., 1999; Heydecker and Wu, 2001; Schluter et al., 1997).

As we can see, there are a number of alternative models and ranking criteria available for the identification of hazardous locations. There are, however, few systematic studies on the comparative performance of alternative models and the practical implications of their use. Many important issues still remain. For instance, the application of alternative models can result in different lists of blackspots; but can these differences be relevant? What is the impact of the use of alternative criteria from a decision-making point of view (e.g., posterior mean versus marginal mean of accident frequency)?

The primary objective of this study is to illustrate the relative impacts for using alternative models or ranking criteria in the context of blackspot identification and provide some guidelines on how to evaluate and select the appropriate model. A sample of highway–railway grade crossings located in the Canadian railway network is used in this investigation. The paper is organized as follows. First, a brief description of three alternative models for accident data analysis is provided. Next, for each model the posterior distribution and the posterior mean of accident

frequency are defined. This is followed by a short description of the accident dataset in which these models are applied. Finally, the calibrated models along with two ranking criteria are compared in terms of goodness-of-fit and blackspot identification.

## 2. Models for Accident Data Analysis

For the analysis of accident data, the Poisson regression model has traditionally been the starting point (Miaou, 1994). This model assumes that the number of accidents occurring over a period of time at a given location $i$, is independently Poisson distributed with a mean of $\mu_i$, that is:

$$Y_i|\mu_i \sim Poisson(\mu_i), \tag{1}$$

where $\mu_i$ is commonly assumed to be an exponential function of a vector of covariates, that is, $\mu_i = \exp\left(\mathbf{x_i'}\boldsymbol{\beta}\right)$, where $\mathbf{x_i} = (1, x_{i1}, \ldots, x_{ik})'$ is a vector of covariates and $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_k)'$ is a vector of regression parameter to be estimated from the data.

The Poisson model is based on the equal-dispersion assumption that the mean is equal to the variance ($E[Y_i|\mu_i] = Var[Y_i|\mu_i]$). This assumption, however, does not hold in many cases, in which the variance can either be larger (over-dispersed) or smaller than the mean (under-dispersed). This is especially common in accident data where over-dispersion is a norm since it is impossible to capture all effects associated with $\mu_i$ (e.g., Maher and Summersgill, 1996). Assuming a Poisson distribution for accident data with problems of over-dispersion would result in underestimation of the standard error of the regression coefficients, which can lead to a biased selection of covariates (Cameron and Trivedi, 1998). A common approach to addressing the over-dispersion problem for unobserved heterogeneities is to consider random effect or mixed Poisson models such as the popular negative binomial, also called Poisson-gamma model. The negative binomial and other two alternative models are discussed in the following subsections.

### 2.1. Negative Binomial (NB) Model: Fixed-$\phi$

Instead of assuming that the mean of accident frequency to be fixed as in the standard Poisson model, in the NB model it is assumed to be random, denoted by $\widetilde{\mu_i}$. With this assumption, the NB model can be written as follows (Lawless, 1987):

$$\begin{aligned} Y_i|\widetilde{\mu}_i &\sim Poisson(\widetilde{\mu}_i), \\ \widetilde{\mu}_i &= \mu_i \exp(\varepsilon_i), \\ \exp(\varepsilon_i) &\sim Gamma(\phi, \phi), \end{aligned} \tag{2}$$

where, as previously defined $\mu_i = \exp(\mathbf{x_i'}\boldsymbol{\beta})$ and $\exp(\varepsilon_i)$ is assumed to follow a gamma distribution with $E[\exp(\varepsilon_i)] = 1$ and $Var[\exp(\varepsilon_i)] = 1/\phi$. This is obtained by

considering that the two parameters of the gamma distribution are equal. In order to obtain the marginal distribution of the NB model, the random effect $\exp(\varepsilon_i)$, is integrated. From this, the NB marginal distribution can be obtained as:

$$f(y_i|\mu_i,\phi) = \frac{\Gamma(y_i+\phi)}{y_i!\Gamma(\phi)}\left(\frac{\phi}{\mu_i+\phi}\right)^{\phi}\left(\frac{\mu_i}{\mu_i+\phi}\right)^{y_i}, y_i = 0,1,\ldots n, \tag{3}$$

where, the dispersion parameter $\phi$ is assumed fixed for all the sites. For computational convenience it is usually written as the reciprocal of α, that is, $\phi = 1/\alpha$. In addition, the conditional mean and variance of the NB marginal distribution are $E[y_i|\mu_i, \phi] = \mu_i$ and $Var[y_i|\mu_i, \phi] = \mu_i(1+\mu_i/\phi)$, respectively. In order to estimate the model parameters, the marginal likelihood of this model can be maximized numerically using the Newton–Raphson algorithm (Cameron et al., 1998).

### 2.2. Heterogeneous Negative Binomial (HNB) Model: Varying-$\phi_i$

The heterogonous negative binomial (HNB) model is an extension of the NB model. The only difference is that, in the HNB model, $\phi_i$ is expressed as a function of some location attributes, such as traffic-flow conditions (Greene, 2002). It follows that, the magnitude of $\phi_i$ varies among locations with the idea of structuring the unobserved heterogeneities (Miaou and Lord, 2003). Modeling $\phi_i$ can increase model flexibility and precision of accident estimates. In the HNB model, $\phi_i$ is computed using the following link function:

$$\phi_i = {}^{1}\!/_{\alpha} \times \exp\left(\mathbf{z_i'}\gamma\right), \tag{4}$$

where $\mathbf{z_i} = (z_{i1},\ldots,z_{ik})'$ is a vector of covariates representing traffic conditions or other site characteristics (not necessarily the same as $\mathbf{x_i}$) and $\boldsymbol{\gamma} = (\gamma_1,\ldots,\gamma_k)'$ is a vector of parameters. The mean and variance of the HNB marginal distribution are the same as the NB model, with the particularity that the parameters in the HNB model are estimated allowing variability in $\phi_i$ and $\mu_i$.

### 2.3. Zero Inflated Negative Binomial (ZINB) Model

As unobserved heterogeneity, another source of over-dispersion can be the high frequency of zeros in the counts. In this situation, the standard Poisson and NB models may be inappropriate. To deal with the problem of excess of zeros, Lambert (1992) introduced the zero-inflated Poisson (ZIP) model where counts are assumed to be generated from two sources. One source represents a perfect state that reflects zero occurrences while the other represents a normal counting process that follows a Poisson distribution. The ZIP model has been applied in road safety analysis to model datasets with high frequency of zero accidents (Miaou, 1994). While the ZIP distribution can handle the problem of excess zeros,

it does not consider the possibility that the problem of over-dispersion is caused by both, unobserved heterogeneity and excess of zeros. To deal with these two possible sources of over-dispersion, a more flexible model can be the zero-inflated negative binomial (ZINB), which can be expressed as follows (Washington et al., 2003):

$$\begin{aligned} &Y_i \sim 0, \text{with probability } \psi_i, \\ &Y_i|\widetilde{\mu}_i \sim Poisson(\widetilde{\mu}_i), \text{with probability } (1-\psi_i), \end{aligned} \tag{5}$$

where $\psi_i$ is a parameter that represents the proportion of zeros added to the NB distribution, $0 < \psi_i < 1$. The ZINB marginal distribution may be written as follows:

$$\begin{aligned} f(y_i|\mu_i,\psi_i,\phi) &= \psi_i + (1-\psi_i)\left(\frac{\phi}{\phi+\mu_i}\right) \text{ for } y_i = 0, \\ f(y_i|\mu_i,\psi_i,\phi) &= (1-\psi_i)\frac{\Gamma(y_i+\phi)}{y_i!\Gamma(\phi)}\left(\frac{\phi}{\mu_i+\phi}\right)^{\phi}\left(\frac{\mu_i}{\mu_i+\phi}\right)^{y_i} \text{ for } y_i > 0. \end{aligned} \tag{6}$$

To allow for covariates in $\psi_i$, one can specify a logistic link function: $\psi_i = \exp(\mathbf{v_i'\omega})/[1+\exp(\mathbf{v_i'\omega})]$, where $\mathbf{v_i} = (1, v_{i1}, \ldots, v_{ik})'$ represents a vector of site attributes, usually different from $\mathbf{x_i}$, and $\boldsymbol{\omega} = (\omega_0, \ldots, \omega_k)'$ is a parameter vector. The marginal mean and variance of the ZINB are given by $E[y_i|\mu_i,\psi_i,\phi] = \mu_i(1-\psi_i)$ and $Var[y_i|\mu_i,\psi_i,\phi] = \mu_i(1-\psi_i)[1+\mu_i(\psi_i+1/\phi)]$, respectively. Note that when $\psi_i$ is close to 1, the expectation of the number of accidents is close to 0. This implies that locations with a $\psi_i$ value of near 1 are sites most likely in a safe state. For parameter estimation, the ZINB marginal likelihood can also be maximized and the resulting problem can be solved using a numerical method such as the Newton–Raphson algorithm.

## 3. Empirican Bayes Approach and Ranking Criteria

The use of posterior distributions through the Bayesian approach has been widely recommended for identifying hazardous locations (e.g., Persaud et al., 1999; Heydecker and Wu, 2001). Its main advantage is that the Bayesian approach combines the information on the safety status of a facility brought by the accident history, with the prior knowledge we have about the safety of the facility into a posterior distribution. Within the class of Bayesian methods, we can distinguish two main approaches—the full Bayes and empirical Bayes (EB) methods (Andrew et al., 2004). An essential difference between these two approaches is in the manner the parameters of the prior distribution are computed. In this work, we utilize the EB approach in which the parameters of interest are estimated by maximizing the marginal likelihood of each model. In the following section, we present the posterior mean of accident frequency for the three previous models.

### 3.1. Posterior Mean of Accident Frequency under the NB and HNB Models

For the NB and HNB models, the prior of $\widetilde{\mu}_i$ is a conjugated gamma distribution[1] and the posterior distribution, denoted by $\pi(\widetilde{\mu}_i|y_i, \mu_i, \phi_i)$, is also gamma distributed and can be written as:

$$\pi(\widetilde{\mu}_i|y_i, \mu_i, \phi_i) \sim Gamma(y_i + \phi_i, 1 + \phi_i/\mu_i), \tag{7}$$

where $\phi_i$ is fixed under the NB model or varies according to Eq. (4) under the HNB model. The posterior expectation of this gamma probability density function is one of the most popular ranking criteria and is often expressed as follows (Hauer and Persaud, 1987):

$$E[\widetilde{\mu}_i|y_i, \mu_i, \phi_i] = \frac{y_i + \phi_i}{1 + \phi_i/\mu_i} = w_i\mu_i + (1 - w_i)y_i, \tag{8}$$

where, $w_i = 1/(1 + \mu_i/\phi_i)$ and the parameters $\mu_i$ and $\phi_i$ are the maximum likelihood (ML) estimates. As previously mentioned, $y_i$ is the observed number of accidents at location $i$ for a given period of time. Note that for a specific value of $\mu_i$, when $\phi_i$ increases, $w_i$ also increases and the weight on $y_i$ decreases. Thus, the parameter $\mu_i$ can have an important weight in locations with high uncertainty.

Alternatively to the posterior mean, other ranking criteria can be easily computed utilizing the EB or full Bayesian approach. Among those criteria, we can mention the posterior probability that $\widetilde{\mu}_i$ excess a standard value, $\pi(\widetilde{\mu}_i \geq c|y_i)$, where $c$ denotes a standard or upper limit of the "acceptable" mean of accident frequency, specified by practitioners depending of the application under consideration. Furthermore, it is also possible to make inference based on the posterior distribution of ranks denoted by $\pi(R_i|y_i)$, where $R_i$ is the rank at location $i$ and is defined as $R_i = \sum_{i \neq j} I(\widetilde{\mu}_i \leq \widetilde{\mu}_j)$, where $j$ denotes for all sites except $i$ (Rao, 2003). In this paper, we concentrate only in the posterior mean of accident frequency.

### 3.2. Posterior Mean of Accident Frequency under the ZINB Model

Applying again Bayes theorem, we obtain the posterior distribution of the ZINB model, which can be written as follows:

$$\begin{aligned} &\pi(\widetilde{\mu}_i|y_i, \psi_i) \propto L(y_i|\widetilde{\mu}_i\psi_i)\pi(\widetilde{\mu}_i), \\ &\pi(\widetilde{\mu}_i|y_i, \psi_i) \propto [\psi_i(1 - d_i) + (1 - \psi_i)Poisson(\widetilde{\mu}_i)]Gamma(\phi, \phi/\mu_i) \end{aligned} \tag{9}$$

where $d_i = \min\{y_i, 1\}$, $L(y_i|\widetilde{\mu}_i, \psi_i)$ is the ZINB likelihood [see Eq. (5)], $\pi(\widetilde{\mu}_i)$ denotes the gamma prior distribution and $Poisson(\widetilde{\mu}_i)$ refers to the Poisson probability density function. Then, the posterior mean after observing $y_i$ accidents during a given unit of time may be denoted by:

$$E[\widetilde{\mu}_i|y_i, \psi_i, \mu_i, \phi] = \psi_i\mu_i(1 - d_i) + (1 - \psi_i)\left(\frac{y_i + \phi}{1 + \phi/\mu_i}\right), \tag{10}$$

[1] Note that based on the properties of the gamma density function, we can write that $\widetilde{\mu}_i \sim$ Gamma$(\phi_i, \phi_i/\mu_i)$, which is equivalent to say that $\widetilde{\mu}_i = \mu_i \exp(\varepsilon_i)$ with exp($\varepsilon_i$)~Gamma $(\phi_i, \phi_i)$.

The use of zero inflated models into the EB framework has not been applied in the context of blackspot identification, being one of the contributions of this paper.

## 4. Relative Performance of Ranking Criteria—A Case Study

The aim of this section is to compare the three alternative models and the two ranking criteria (e.g., marginal mean versus posterior mean of accident frequency) defined previously. For that, a sample of highway–rail grade crossings is utilized as an application environment. Some details of this sample as well as the main results of the model calibration and evaluation are also discussed in this section.

### 4.1. Data Description

The dataset used in this analysis includes two databases provided by Transport Canada and the Transportation Safety Board. One database consists of an inventory, containing information of approximately 29,500 grade crossings (public and private) located in the Canadian railway network. The other is the accident occurrence database which includes information of collisions for several years. In the inventory database, four groups of attributes are included for each crossing: spatial location, type of warning device, geometric characteristics and traffic conditions (e.g., number of road vehicles and trains daily). In this application, we consider a sample of public crossings with passive warning devices (i.e., crossbucks, pavement markings and parallel track signs), which includes approximately 13,241

**Table 1** Summary of the crossing attributes and collision history

| Category | Variable | Description | Average or percentage | Minimum | Maximum |
|---|---|---|---|---|---|
| Road and railway attributes | Posted road speed | Km/hr | 54.4 | 5.0 | 100.0 |
| | Road type | Dummy variable: arterial or collector = 1, 0 otherwise | 3.5(%) | 0.0 | 1.0 |
| | Main track | Dummy variable: main track = 1, 0 otherwise | 86.8 (%) | 0.0 | 1.0 |
| | Track number | Number | 1.1 | 1.0 | 9.0 |
| | Track angle | Degrees | 69.1 | 0.0 | 90.0 |
| | Train speed | Mile/hour | 34.3 | 5.0 | 95.0 |
| Traffic volumes | AADT | Number | 321.1 | 1.0 | 29000.0 |
| | Daily trains | Number | 4.9 | 1.0 | 59.0 |
| | Exposure | ln [AADT×daily trains] | 4.6 | 0.0 | 13.2 |
| Dependent variable | Collision history in a five-year period | Accident average per crossing | 0.03 | 0.0 | 3.0 |
| | | Proportion of zeros in percentage | 97.4 (%) | – | – |

*AADT* Average annual daily traffic.

crossings, as well as the history of accidents of the period 1997–2001 (5 years of accident information). A brief description of the crossing attributes is presented in Table 1. In this sample, 99.6% of the crossings report AADT's between 1 and 10,000 vehicles, (only 50 crossings register traffic volumes greater than 10,000 vehicles). The number of daily trains varies from 1 to 59 trains, however 99.6% of the crossings report between 1 and 35 daily trains (only 55 crossings report more than 35 daily trains).

In order to identify high linear correlation between crossing attributes, a correlation matrix was estimated, the results of this matrix are presented in Table 2. Small or moderate linear association was found between the crossing attributes involved in this application, with correlation coefficients less than 0.5. The highest correlation was identified between daily trains and train speed, with a correlation coefficient of 0.45. Linear correlation between the number of daily trains and AADT is almost zero.

### 4.2. Functional Form Definition and Model Calibration

For estimating the model parameters, the functional form of $\mu_i$ was first be specified. In this case, we adopted the popular exponential form (e.g., Miaou and Lord, 2003):

$$\mu_i = \exp(\beta_0 + \beta_1 \ln x_{i1} + .. + \beta_k x_{ik}) = x_{i1}^{\beta_1} \exp(\beta_0 + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}), \qquad (11)$$

where $[x_{i1}]$ is a measure of traffic exposure defined as $x_{i1} = \ln[\text{AADT} \times \text{trains daily at site } i]$ and $x_{i2}, \ldots, x_{ik}$ are other crossing attributes. This functional form attempts to capture the nonlinear relationship between $\mu_i$ and $x_{i1}$, which has been found in previous work (Hauer, 1997). Since there is no information on daily variations of rail and highway traffic or other traffic characteristics, it is not possible to explore more precise measures of exposure. For estimating the parameter $\phi_i$ in the HNB model and $\psi_i$ in the ZINB model, we consider only the traffic conditions, AADT and trains daily, as potential covariates. In the case of the ZINB model, we suppose that intersections with very low traffic volumes can be considered as those almost in a safe state or with a low probability of accidents occurrence.

Before the calibration, the grade crossing sample (including a total of 13,241 crossings) was split into two random sub-samples: one consisting of 75% of the crossings was used for calibration and the other 25% was employed for validation. This section discuss the parameter estimation results for the three models: NB, HNB and ZINB. The statistical software package LIMDEP 8.0 was utilized for model calibration. *t*-tests with a confidence level of 95% were used to identify statistically significant crossing attributes. A summary of the model parameter estimates and the corresponding *t* values is shown in Table 3.

The crossing attributes that were found to be statistically significant for the three models are, posted road speed, maximum train speed and traffic exposure, being the last the most important. In addition, the coefficients of these three attributes are positives as logically expected. Regarding the dispersion parameter $\phi_i$ for the HNB, we found that the number of daily trains, is statistically significant with a negative coefficient ($\gamma_1$). Therefore, as the traffic of trains increases the certainty in the accident estimates increases as well. The daily train traffic is also linked to the parameter $\psi_i$ of the ZINB model. As supposed, crossings with lower

**Table 2** Correlation matrix

| Variable | Road speed | Road type | Main track | Track number | Track angle | Train speed | AADT | Daily trains | Exposure | Collisions |
|---|---|---|---|---|---|---|---|---|---|---|
| Road speed | 1.000 | 0.088 | 0.044 | −0.052 | −0.071 | −0.030 | −0.011 | −0.099 | 0.122 | 0.027 |
| Road type | 0.088 | 1.000 | −0.009 | 0.022 | 0.000 | −0.007 | 0.095 | 0.041 | 0.169 | 0.055 |
| Main track | 0.044 | −0.009 | 1.000 | −0.132 | −0.017 | 0.477 | −0.366 | 0.108 | −0.405 | −0.018 |
| Track number | −0.052 | 0.022 | −0.132 | 1.000 | 0.101 | −0.002 | 0.072 | 0.201 | 0.251 | 0.035 |
| Track angle | −0.071 | 0.000 | −0.017 | 0.101 | 1.000 | −0.003 | 0.029 | 0.59 | 0.087 | 0.031 |
| Train speed | −0.030 | −0.007 | 0.477 | −0.002 | −0.003 | 1.000 | −0.238 | 0.453 | −0.016 | 0.057 |
| AADT | −0.011 | 0.095 | −0.366 | 0.072 | 0.029 | −0.238 | 1.000 | −0.048 | 0.437 | 0.058 |
| Daily trains | −0.099 | 0.041 | 0.108 | 0.201 | 0.059 | 0.453 | −0.048 | 1.000 | 0.328 | 0.106 |
| Exposure | 0.122 | 0.169 | −0.405 | 0.251 | 0.087 | −0.016 | 0.437 | 0.328 | 1.000 | 0.146 |

**Table 3** Model calibration results

| Model | Variable | Parameter | Standard error | *t*-ratio | *p* value | Log-likelihood |
|---|---|---|---|---|---|---|
| Poisson | Constant ($\beta_0$) | −7.047 | 0.280 | −25.187 | 0.000 | $\ell = -1200.2$ |
| | Exposure ($\beta_1$) | 0.463 | 0.028 | 16.283 | 0.000 | $n = 9931$ |
| | Train speed ($\beta_2$) | 0.017 | 0.003 | 6.668 | 0.000 | |
| | Road speed ($\beta_3$) | 0.006 | 0.003 | 1.976 | 0.048 | |
| NB | Constant ($\beta_0$) | −7.178 | 0.311 | −23.113 | 0.000 | $\ell = -1188.0$ |
| | Exposure ($\beta_1$) | 0.480 | 0.033 | 14.758 | 0.000 | $n = 99.31$ |
| | Train speed ($\beta_2$) | 0.017 | 0.003 | 6.144 | 0.000 | |
| | Road speed ($\beta_3$) | 0.006 | 0.003 | 2.015 | 0.044 | |
| | Alpha (α) | 2.141 | 0.650 | 3.294 | 0.001 | |
| HNB | Constant ($\beta_0$) | −7.114 | 0.345 | −20.622 | 0.000 | $\ell = -1186.0$ |
| | Exposure ($\beta_1$) | 0.478 | 0.040 | 12.046 | 0.000 | $n = 99.31$ |
| | Train speed ($\beta_2$) | 0.016 | 0.003 | 5.433 | 0.000 | |
| | Road speed ($\beta_3$) | 0.006 | 0.003 | 1.974 | 0.048 | |
| | Alpha (α) | 3.257 | 1.268 | 2.568 | 0.10 | |
| | Daily trains ($\gamma_1$) | −0.020 | 0.009 | −2.147 | 0.32 | |
| ZINB | Constant ($\beta_0$) | −6.307 | 0.425 | −14.839 | 0.000 | $\ell = -1180.8$ |
| | Exposure ($\beta_1$) | 0.414 | 0.044 | 9.440 | 0.000 | $n = 9931$ |
| | Train speed ($\beta_2$) | 0.009 | 0.004 | 2.393 | 0.017 | |
| | Road speed ($\beta_3$) | 0.008 | 0.003 | 2.863 | 0.004 | |
| | Alpha (α) | 1.673 | 0.845 | 1.980 | 0.048 | |
| | Logit constant ($\omega_0$) | 1.109 | 0.540 | 2.054 | 0.040 | |
| | Daily trains ($\omega_1$) | −0.832 | 0.380 | −2.189 | 0.029 | |

train traffic have lower expected number of accidents than those with higher train volumes. We found that the range of $\psi_i$ goes from 0 to 0.6.

### 4.3. Overdispersion, Excess of Zeros and Goodness of Fit

Once the parameters have been estimated, the next step is to detect over-dispersion in the data. This can be done by testing the null hypothesis of the dispersion parameter, $H_0 : \alpha = 0$, that is, by contrasting the variance-mean equality assumption of the Poisson model against an alternative model in which the variance exceeds the mean. The comparison can be made using the likelihood ratio statistic, defined as $T_{LR} = -2(\ell_P - \ell_A)$ (Cameron and Trivedi, 1998), where, $\ell_P$ is the log-likelihood value of the restricted (Poisson) model and $\ell_A$ is the log-likelihood estimate of the unrestricted model that considers over-dispersion (e.g., NB model). Knowing that $T_{LR}$ approximately follows a chi-square distribution ($\chi^2_{df}$) (*df*—degrees of freedom), the null hypothesis is rejected if it exceeds a critical value. From Table 3, the log-likelihood values of the Poisson and NB models are −1200.2 and −1188.0, respectively. Thus, $T_{LR}$ is equal to 24.4, exceeding the 1% critical value of $\chi^2_1 = 5.41$. Note that in this case, the number of restrictions or degrees of freedom is equal to 1. Using log-likelihood estimates of the HNB and ZINB models, the values of the $T_{LR}$ statistic are even greater than the critical value. From this, the presence of over-dispersion was detected with the three alternative models.

Given the high proportions of zeros in the dataset, we consider the use of the ZINB model. To evaluate the appropriateness of this model as compared to the NB model, we use the Vuong statistic denoted by $V$, which is useful in comparing non-nested models (Washington et al., 2003). In order to obtain the $V$ statistic, we first compute for each site $L_i = \ln[P_1(Y_i)/P_2(Y_i)]$, where $P_1(Y_i)$ and $P_2(Y_i)$ are the marginal probability distributions of the NB and ZINB models, respectively. Then, the $V$ statistic is calculated as, $V = \sqrt{n}(\overline{L})/S_L$, where $\overline{L} = (1/n)\sum_{i=1}^{n} L_i$, $S_L$ is the standard deviation of $L_i$, and $n$ is the sample size. Since the $V$ statistic follows asymptotically the standard normal distribution, it can be compared with $z$-values. Hence, if $|V|$ is greater than $V_c = 1.96$ (critical value assuming a 95% confidence level) the test favors the selection of the ZINB model. In other words, large positive values of $V$ support the ZINB model whilst large negative values support the alternative. For the dataset studied here, $V$ is approximately 2.0, which is very close to the critical value. From this we can see that despite the high frequency of zeros in the data, the superiority of the ZINB model over the NB model cannot be statistically confirmed.

To evaluate the goodness-of-fit of the three models, we utilize the hold-out dataset for validation purpose. With this dataset, we compute and compare the frequency distributions produced by classifying the crossings according to the observed and estimated number of accidents. The observed frequencies $O_j$, are the number of locations with 0, 1, 2...$J$ accidents, during the time period of analysis. The estimated frequencies denoted by $E_j$, are computed as $E_j = \sum_i^n P(Y_i = j)$, where $P(Y_i = j)$ is the probability of having $j$ ($j$ = 0,1,2...$J$) accidents at location $i$ and $n$ is the sample size. The classification of crossings according to their observed and estimated accident frequency is presented for each model in Table 4. From this, we can observe that the three models produce similar results. Although the HNB and ZINB models fit slightly better the data in the two main categories with 0 and 1 accident. In addition, we can compare the goodness-of-fit of non-nested models calibrated with the same data using the Akaike information criterion—AIC (Cameron and Trivedi, 1998). This criterion is computed as AIC = −2 log-likelihood + 2$k$, where $k$ = number of estimated parameters included in the model. The model with lowest AIC is preferred. In the present application, $k$ is equal to 5, 6 and 7 parameters, and then the AIC values are equal to 2386 2,384 and 2376 for the NB, HNB and ZINB models, respectively. In summary, we can notice that although the HNB and ZINB models fit slightly better the data than the NB model, the use of alternative model structures did not have an important improvement in terms of goodness-of-fit for this particular dataset.

**Table 4** Goodness-of-fit: observed versus estimated frequencies

| Collisions per crossing | Number of crossings | | | |
|---|---|---|---|---|
| | Observed | NB | HNB | ZINB |
| 0 | 3236 | 3219 | 3230 | 3229 |
| 1 | 65 | 83 | 71 | 73 |
| 2 | 7 | 7 | 8 | 7 |
| 3 | 2 | 1 | 1 | 1 |
| Total | 3310 | 3310 | 3310 | 3310 |

### 4.4. Decision Implications for Using Alternative Models and/or Ranking Criteria

The application of alternative models or ranking criteria may lead to different ranks and lists of blackspots. To investigate this hypothesis, we made a comparative analysis between the marginal and posterior expectation of each model, which are denoted as follows:

- *Expected accident frequency* under the marginal distribution of each model, i.e., $E[y_i|\mu_i]$. We denote as $\widehat{\mu}_i^{NB}$, $\widehat{\mu}_i^{HNB}$ and $\widehat{\mu}_i^{ZINB}$ the marginal mean under the NB, HNB and ZINB models, respectively.
- *Posterior mean of accident frequency* under the posterior distribution, i.e., $E[\widetilde{\mu}_i|y_i]$, denoting as $\widehat{\mu}_i^{B-NB}$ the NB posterior mean [Eq. (8)], $\widehat{\mu}_i^{B-HNB}$ the HNB posterior mean Eq. (8) and $\widehat{\mu}_i^{B-ZINB}$ the ZINB posterior mean [Eq. (10)].

As a fist step, we ranked the same hold-out sample according to each criterion, resulting in six different lists. Then, the crossings at the top of each list were selected as blackspots, e.g., the top 100 crossings of each list. Finally, in order to measure the differences between two lists of blackspots, the percentage deviation was computed as follows:

$$\%\ \ deviation = 100 \times \left(1 - {}^{\kappa}/_{r}\right), \tag{12}$$

where $\kappa$ is the number of locations that are common in two lists and $r$ is the list size. Thus, in order to measure the relative impact of using alternative model structures, we computed the % deviations between two blackspot lists identified by using the same ranking criterion but two different models, for instance, $\widehat{\mu}_i^{B-NB}$ versus $\widehat{\mu}_i^{B-HNB}$. In addition, in order to observe the relative impact of applying an alternative criterion, we computed the % deviation using the same model but two different criteria. Figure 1 shows the % deviation among different models and ranking criteria under different list sizes ($r$), from which the following observations can be made:

- The blackspot lists identified using two alternative models and the same ranking criterion are quite similar. For example, approximately 90% of blackspots identified with $\widehat{\mu}_i^{B-NB}$ are the same as the ones identified by $\widehat{\mu}_i^{B-HNB}$. The discrepancies between $\widehat{\mu}_i^{B-NB}$ and $\widehat{\mu}_i^{B-ZINB}$ are less than 20% in all the cases. Small differences were also found among the marginal expectations (these results were not included in the paper). This result suggests that the use of alternative models would have small repercussions on the ranks.
- Conversely, the differences between using two different ranking criteria are fairly significant (e.g., $\widehat{\mu}_i^{NB}$ versus $\widehat{\mu}_i^{B-NB}$). This is especially true when the number of blackspot to be identified is small (e.g., $r \leq 100$). Here, we see clearly that the inference based on the marginal distributions may produce very different results than using the posterior distribution.
- In the present application, the choice of ranking criteria has more implications than the choice of model structure. Note that the ranks based on the marginal distributions are basically determined as a function of location attributes, whilst accidents history plays an important role when making inference according to the posterior distribution.

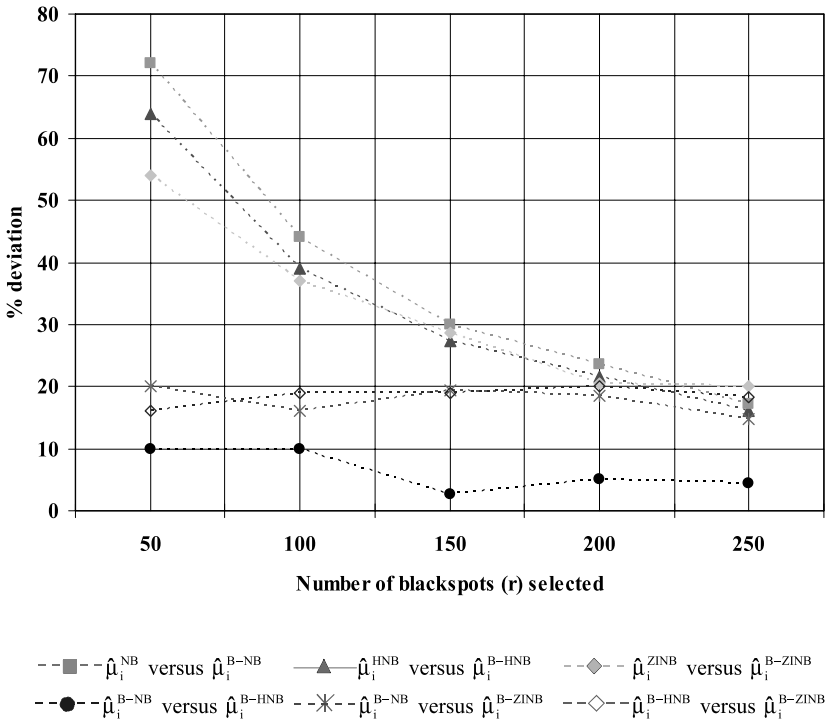


**Fig. 1** Percentage deviations between alternative model structures and criteria

## 5. Conclusions

This paper investigates the implications of applying alternative models and ranking criteria in the identification of hazardous locations. This work was motivated by the vast set of models available in the literature and the lack of formal guidance for model selection in the transportation safety context. Three alternative models were calibrated and evaluated using a sample of highway–rail grade crossings located in the Canadian railway network. From these models, accident risk estimators were computed and compared utilizing two alternative criteria, the marginal and posterior expectation of accident frequency. The analysis indicated that small differences in the ranks were obtained when applying alternative models and the same criterion. That is, the lists of blackspots were pretty similar when utilizing the accidents estimators derived from the posterior mean of each model. In contrast, substantially different lists of blackspots were identified when ranking the sites based on the marginal mean and posterior expectation of accident frequency. In general, the blackspots identified with a marginal distribution may be extremely different from those selected based on a posterior distribution.

In addition, the HNB model was presented as a more flexible option than the traditional NB model for analysis of accident data. The advantage of the HNB model is that observed variability is allowed in the dispersion parameter. This may

improve goodness-of-fit and accuracy of the accident estimators. Furthermore, the ZINB model was considered given the extremely high frequency of zeros in the accident dataset studied. However, this model fitted only slightly better the accident data than the NB model. This shows that the NB model can still be a good candidate for modeling accident data with high proportion of zeros. Here, the use of the ZINB model has been extended in the context of blackspot identification into an empirical Bayesian framework.

As for future research, a simulation study will be carried out to validate the conclusions reached in this paper. Measures other than point estimates, such as the posterior probability of excess based on $\theta_i$ or $R_i$, will be investigated. Furthermore, some general decision rules for the identification of blackspots, such as thresholds or cutoffs values, will be developed to help not only in the ranking but also in the selection of hazardous sites for cost-effective safety improvements.

## References

Andrew G, Carlin JB, Stern HS, Rubin DB (2004) Bayesian data analysis. Chapman and Hall/CRC, Washington, District of Columbia

Austin RD, Carson JL (2002). An alternative accident prediction model for highway–rail interfaces. Accident Anal Prev 34:31–42

Cameron AC, Trivedi PK (1998) Regression analysis of count data. Cambridge University Press, Cambridge, UK

Greene W (2002) Econometrics Modeling Guide, LIMDEP, Version 8.0. Econometric Software, Inc

Hauer E (1997) Observational before–after studies in road safety. Pergamon

Hauer E, Persaud BN (1987) How to estimate the safety of rail–highway grade crossing and the effects of warning devices. Transp Res Rec 1114:131–140

Heydecker BG, Wu J (2001) Identification of sites for accident remedial work by bayesian statistical methods: an example of uncertain inference. Adv Eng Softw 32:859–869

Lambert D (1992) Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics 34-1:1–14

Lawless JF (1987) Negative binomial and mixed Poisson regression. Can J Stat 15:209–225

Lord D, Persaud BN (2000) Accident prediction models with and without trend: application of the generalized estimation equations procedure. Transp Res Rec 1717:102–108

Maher MJ, Summersgill (1996) A comprehensive methodology for the fitting of predictive accident models. Accident Anal Prev 28(3):281–296

Miaou SP (1994) The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Anal Prev 26-4:471–482

Miaou SP, Lord D (2003) Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and bayes versus empirical bayes. Transp Res Rec 1840:31–40

Miaou SP, Song JJ (2004) Bayesian ranking of sites for engineering safety improvement. 83rd annual meeting of the transportation research board. Washington, District of Columbia

Persaud B, Lyon C, Nguyen T (1999) Empirical bayes procedure for ranking sites for safety investigation by potential for safety improvement. Transp Res Rec 1665:7–12

Rao J (2003) Small area estimation. Wiley, New York

Saccomanno F, Fu L, Miranda-Moreno LF (2004) Risk-based model for identifying highway–rail grade crossing blackspots. Transp Res Rec 1862:127–135

Schluter PJ, Deely JJ, Nicholson AJ (1997) Ranking and selecting motor vehicle accident sites by using a hierarchical bayesian model. Statistician 46-3:293–316

Shankar VN, Ulfarsson GF, Pendyala RM, Nebergall MB (2003) Modeling crashes involving pedestrian and motorized traffic. Saf Sci 41

Washington SP, Karlaftis MG, Mannering FL (2003) Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC, Washington, District of Columbia