Accident Analysis and Prevention 55 (2013) 107–115

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

A crash-prediction model for road tunnels

Ciro Caliendo[a,*], Maria Luisa De Guglielmo[a], Maurizio Guida[b]

[a] Department of Civil Engineering, University of Salerno, 84084 Fisciano (SA), Italy
[b] Department of Information Engineering, Electrical Engineering, and Applied Mathematics, University of Salerno, 84084 Fisciano (SA), Italy

ARTICLE INFO

Article history:
Received 19 July 2012
Received in revised form 11 February 2013
Accepted 12 February 2013

Keywords:
Road tunnels
Crash-prediction model
Non-severe and severe crashes
Bivariate Negative Binomial regression
Random Effects Binomial regression
Negative Multinomial regression
Traffic flow
Trucks
Length
Number of lanes

ABSTRACT

Considerable research has been carried out into open roads to establish relationships between crashes and traffic flow, geometry of infrastructure and environmental factors, whereas crash-prediction models for road tunnels, have rarely been investigated. In addition different results have been sometimes obtained regarding the effects of traffic and geometry on crashes in road tunnels. However, most research has focused on tunnels where traffic and geometric conditions, as well as driving behaviour, differ from those in Italy. Thus, in this paper crash prediction-models that had not yet been proposed for Italian road tunnels have been developed. For the purpose, a 4-year monitoring period extending from 2006 to 2009 was considered. The tunnels investigated are single-tube ones with unidirectional traffic. The Bivariate Negative Binomial regression model, jointly applied to non-severe crashes (accidents involving material-damage only) and severe crashes (fatal and injury accidents only), was used to model the frequency of accident occurrence. The year effect on severe crashes was also analyzed by the Random Effects Binomial regression model and the Negative Multinomial regression model. Regression parameters were estimated by the Maximum Likelihood Method. The Cumulative Residual Method was used to test the adequacy of the regression model through the range of annual average daily traffic per lane. The candidate set of variables was: tunnel length (L), annual average daily traffic per lane ($AADT_L$), percentage of trucks (%Tr), number of lanes (N_L), and the presence of a sidewalk. Both for non-severe crashes and severe crashes, prediction-models showed that significant variables are: L, $AADT_L$, %Tr, and N_L. A significant year effect consisting in a systematic reduction of severe crashes over time was also detected. The analysis developed in this paper appears to be useful for many applications such as the estimation of accident reductions due to improvement in existing tunnels and/or to modifications of traffic control systems, as well as for the prediction of accidents when different tunnel design options are compared.

1. Introduction

Driving in road tunnels reduces travel times and enhances the convenience of users, since tunnels allow passing through crowded urban areas and overcoming environmental obstacles such as mountains and rivers. Tunnels may also preserve natural landscape by reducing the damage to nature and the environment. The increasing rate of tunnels construction has been rapid during the last few decades also because innovative technologies for removing and controlling smoke in the event of fire have been installed.

Road tunnels, however, impose a form of driving behaviour that is unique when compared to that on open roads. Driving inside a tunnel in normal situations may cause anxiety because tunnels are dark, narrow and monotonous (PIARC, 2008). Some drivers may also be frightened of hitting something (e.g. other vehicles or tunnel walls) or of dangerous situations (e.g. a fire or a tunnel collapse). Furthermore drivers in tunnels generally modify both their lateral position and speed in order to avoid the disturbing effects due to the tunnel wall being too close to the traffic lane, which may happen more especially when an emergency lane is absent. Also when approaching the tunnel portal drivers often change their driving style on account of narrower tunnel structures by increasing the distance from the side wall and reducing their speed, which may interfere with the traffic flow in the adjacent lane. Another effect at the tunnel entrance is that the sunshine reflected from the tunnel portal or direct sunshine might cause ocular blinding of drivers before entering the tunnel, while the darkness in the first part of the tunnel might also cause poor sight conditions due to slow adaptation of eyes to the dark. In addition another effect might include blinding light when drivers leave tunnel at sunset, as well as their being negatively taken by surprise at the exit side by unexpected conditions (rain, fog, snow, lateral winds, traffic jams) that might force them to change their behaviour again. Therefore, driving in tunnels in general requires more attention and a greater mental workload (see PIARC, 2008 for greater in-depth knowledge of human behaviour in tunnels).

* Corresponding author. Tel.: +39 89 964140; fax: +39 89 964045.
E-mail address: ccaliendo@unisa.it (C. Caliendo).

0001-4575/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.aap.2013.02.024

Dealing with crashes in road tunnels is generally very difficult because of the excessive number of factors involved. However, the influence of those that appear to play a prevalent role is worthy of attention. When an emergency lane is present the above-mentioned effects on lateral position and speed reduction are reduced. The emergency lane is also expected to have a positive effect on road safety because this provides an escape lane for cars that have broken down and/or for emergency services that can reach the location more quickly in the event of crash and/or fire. On rising slopes heavy vehicles sometimes slow down considerably, which results in greater speed differences between cars and trucks so that a higher probability of crash might be expected. Narrower tunnels (e.g. with fewer lanes) are generally thought, because of the limited space, to be less safe than wider ones, while longer tunnels are often considered to be more dangerous than shorter ones. More complicated horizontal alignments are generally considered to cause more crashes; which might be attributable to the presence of bends that is sometimes difficult to estimate given that the tunnel walls reduce the view of the road and possible vehicle queues. By increasing traffic and/or the percentage of trucks more crashes are often also expected. Finally, a single tunnel tube for traffic in two driving directions is considered to be less safe than two separate tunnel tubes with one driving direction because head-on collisions can also occur (e.g. crashes with other vehicles moving in the opposite direction).

Indeed, crashes in road tunnels have hitherto hardly ever been studied, except in the case of large fires where a vast literature exists (Caliendo et al., 2012). In addition, the likely impact is not yet clearly understood on crashes in road tunnels due to some predominant variables (e.g. tunnel length, traffic, percentage of trucks, number of lanes, and the presence of sidewalk and/or emergency lane), and whether these variables are statistically significant.

In fact, according to the few studies available in the literature longer tunnels are safer; whilst other studies point out that the number of crashes increases as tunnel length increases. This is attributed to the drivers' diminishing concentration with increasing tunnel length.

In addition, with an increase in traffic a greater number of crashes is not always expected. This depends on whether free-flowing or almost congested flow-traffic conditions are found. In free-flow conditions with an increase in traffic drivers still enjoy great freedom for changing lanes and overtaking, which might be associated with an increase in crash risk; whereas in almost congested conditions when traffic increases the drivers' freedom of manoeuvre begins to become ever more limited, and consequently might be associated with a decrease in crash risk.

With reference to the influence of trucks in a traffic stream, some studies show that as the percentage of trucks increases the number of crashes decreases, whilst others find that crashes increase as truck percentages increase. One possible reason might be that, for free-flow conditions, as the percentage of trucks increases the frequency of lane changing and overtaking movements increases due to the increased number of cars so that more crashes are expected compared with almost congested traffic conditions.

Adding lanes is motivated by the need to relieve traffic congestion. But it is often believed by engineers and planners that decreased congestion resulting from adding lanes is also associated with improved safety, whilst the majority opinion among researchers is that crashes increase with an increase in the number of lanes due to more opportunities for lane change and consequently to more conflicts.

The aforementioned different views are still subject of controversy. Additionally it is to be said that the most common crash indicator that has hitherto been used is the number of crashes per million vehicle-kilometres (crash rate); but the relationship between crashes and tunnel length, as well as between crashes and traffic may be not linear so that a crash indicator expressed in terms of the number of crashes per year (crash frequency) appears to be more appropriate as a dependent variable for crash-prediction models. Besides, given that the freedom of changing lanes and overtaking is better understood by considering traffic per lane, the use of annual average daily traffic per lane (AADT/lane) seems to be more suitable than AADT also for showing the effects of the percentage of trucks and/or number of lanes.

Additionally it is to be said that the studies available have been carried out prevalently in countries where traffic characteristics, driver behaviour, and geometry of tunnels differ from those in Italy. In this respect it is to be noted that according to a monitoring exercise regarding Italian motorway tunnels (Caliendo and De Guglielmo, 2012), the average accident rates were computed to be 0.25 accidents/10^6 veh-km for total crashes and 0.12 accidents/10^6 veh-km for severe crashes. In the same monitoring period, on the motorway sections containing these tunnels the average severe accident rate was estimated to be 0.09 severe accidents/10^6 veh-km. This means that regarding severe crashes, the average severe accident rate in Italian motorway tunnels is in general higher than that of the corresponding motorway sections; in other words the consequences of crashes in tunnels are often more severe. Therefore, there is evidence that crashes in road tunnels need to be investigated in greater detail.

As a result of the above, there are at least three main reasons for justifying this paper. The first is motivated by the need to quantify the effects on the crash frequency in road tunnels as a function of the following main variables: sidewalk, tunnel length, traffic, percentage of trucks, and number of lanes. In fact there is an a priori reason for believing that crashes in road tunnels are associated with these variables in different ways if compared to accidents occurring on open road sections. The second is to have a better understanding about the more significant variables that tend to increase or decrease accident frequency in road tunnels. Such knowledge may be helpful for suggesting measures for improving tunnel safety. Finally, given the lacuna regarding crash predictive models for Italian road tunnels, relationships that were not known before should now be developed.

The objective of this paper is to identify specific prediction models to estimate crashes in road tunnels as a function of traffic and of geometric infrastructure characteristics. For this purpose, the database of a 4-year monitoring period on Italian motorway tunnels was utilized. All tunnels investigated were single-tube ones with unidirectional traffic, and having two or three lanes. The number of crashes occurring in these motorway tunnels over the 4-year monitoring period was assumed as a dependent variable to be related both to traffic flow and road factors in the models proposed. Both non-severe crashes and severe crashes were jointly investigated in order to propose a prediction model. The Bivariate Negative Binomial distribution was used to model the random variation of the number of crashes. The likelihood function was maximized to obtain the estimates of model parameters. The Cumulative Residual Method was used to test the adequacy of the regression model across the range of the most significant variable, namely AADT.

Furthermore, in order to analyze the year effect, the number of severe crashes per year occurring in the motorway tunnels was subsequently assumed as dependent variable. In this respect, two specific prediction models suitable to deal with temporal data were proposed, namely the Random Effects Negative Binomial regression model, and the Negative Multinomial regression model.

In the light of the above considerations, the present paper is organized as follows: the next section contains a literature review concerning crashes in road tunnels, while the subsequent section deals with the data set used and the process of preparing it for analysis. Then the results of statistical modelling are presented and discussed, and two prediction models are proposed.

Finally, the conclusions and addresses for further studies are made.

2. Literature review

As far as the authors are aware the first studies on traffic accidents in road tunnels were devised by the PIARC Committee (1995) on Tunnel Safety. This Committee reported that: tunnels are safer than open roads except in the case of horizontal and/or vertical alignment defects; bidirectional tunnels cause more accidents than unidirectional tunnels; the average accident rate with injuries is 8 and 10 accident/10^8 veh-km for unidirectional and bidirectional tunnels, respectively. In studying driving behaviour in road tunnels, Törnoros (1998) found that the lateral distance to the nearest tunnel wall was greater when this wall was located to the left side of the drivers than to their right side. Amundsen and Ranes (2000) showed that traffic accident rates, which are higher in the entrance zone of tunnels, diminish as one proceeds inside the tunnel. These accident rates were also negatively associated with: tunnel length, AADT, and tunnel width. Lemke (2000) found that accident rates in unidirectional tunnels are fewer than those of the open roads (reduced by half); accident rates were also associated positively with the tunnel length and negatively with the presence of shoulder. Having found that the injury accident rates in unidirectional tunnels were both higher and lower than those of the corresponding open sections, SAFESTAR (2002) argues that one cannot generally conclude that safety in tunnels is better or worse than on open roads. Salvisberg et al. (2004) claims that the risk of an accident occurring in a tunnel is lower in longer tunnels than in shorter tunnels, and that this risk increases with the AADT and/or trucks. As a consequence of the catastrophic tunnel fires occurring in Europe over the last few years (Mont Blanc, Tauern, and Saint Gotthard tunnels) the European Parliament and Council (2004) adopted the European Directive 2004/54/EC. This European Directive set technical standards in order to prevent critical events that may endanger human life, and which regard a classification of road tunnels at five levels depending on traffic volume for lane and tunnel length, in which specific features should correspond to each level such as the number of tubes and lanes, unidirectional or bidirectional traffic, tunnel geometry, road signs, lightning, ventilation, and emergency exits. Manser and Hancock (2007) found that drivers' perception of speed and their subsequent behaviour are influenced by the visual pattern and the texture applied to the tunnel wall. Nussbaumer (2007) supported Salvisberg's aforementioned study and adds that the risk of accidents is higher in tunnels with bidirectional traffic than in tunnels with unidirectional traffic. Vashitz et al. (2008) showed that by means of improved communications in real time between drivers and tunnel through Intelligent Transport System (ITS) fewer accidents are expected. SWOV (2009) found that motorway tunnels have more injury accidents per vehicle/kilometre than open road sections in contrast with other international studies. Kircher and Ahlstrom (2012) showed how drivers' behaviour in tunnels is influenced by the illumination level and brightness of the tunnel walls.

The above-cited chronological presentation of the literature proves that different results have been obtained, and as a consequence the influence on crashes in road tunnels due more especially to traffic and geometry needs to be better understood. In addition these studies were all investigated in foreign countries where traffic and geometric conditions, as well as driving style, differ from those in Italy. This paper sets out to make a contribution to our knowledge by helping to understand more about this leading subject, also considering the development of crash-prediction models.

A wide variety of statistical methods have been applied over the years to crash-frequency analysis for developing prediction models. In this respect, Lord and Mannering (2010) and Savolainen et al. (2011) provided a listing of the methods applied in past studies with their strengths and weaknesses. In dealing with problems associated both with crash-frequency data and different statistical approaches, these authors also describe methodological alternatives in an attempt to improve the statistical validity of findings. With respect to the purposes of our paper, bivariate/multivariate regression models to analyze correlated accident counts such as crashes of different severity levels, and random-effects regression models to deal with temporal correlation in the data are here of special interest. In particular, bivariate/multivariate regression models were considered among others in Subrahmaniam and Subrahmaniam (1973), Marshall and Olkin (1990), Bijleveld (2005), Song et al. (2006), Park and Lord (2007), El-Basyouny and Sayed (2009a), and Anastasopoulos et al. (2012a). Regression models to account for temporal correlation in the data were proposed by Hausman et al. (1984) (the Random Effects Negative Binomial model) and by Guo (1996) (the Negative Multinomial model). Analyses of count data based on RENB and/or NM models were considered in Shankar et al. (1998), Shankar and Ulfarsson (2003), Chin and Quddus (2003), Hauer (2004) and Caliendo et al. (2007). In recent years, random-parameters models, which can be viewed as an extension of random-effects models, have been considered in Milton et al. (2008), Anastasopoulos and Mannering (2009), El-Basyouny and Sayed (2009b) and Anastasopoulos et al. (2012b). These models attempt to account for heterogeneity across accident observations by allowing some or all parameters to vary, in contrast with the assumption that parameters are considered to be constant across observations.

Nevertheless, as far as the authors are aware bivariate/multivariate or random-parameters models have hardly ever been applied to crashes in road tunnels. This paper makes this additional contribution possible.

3. Data description

A 4-year monitoring period extending from 2006 to 2009 was considered for Italian motorway tunnels. The database consisted of 260 tunnels with unidirectional traffic only, 232 of which were two-lanes tunnels while the remaining 28 were three-lane tunnels. The total length of the tunnels monitored was 303 km, with a total length of two-lane tunnels of 276 km and 27 km for three-lane tunnels, respectively. During the monitored period, crash data and traffic flow were collated. Accident data were extracted from the official reports of the Motorway Management Agencies (MMA) of these tunnels. For each accident a variety of details was recorded, including the name of tunnel in which accidents occurred, date, type and accident severity, number of vehicles and persons involved. Some 2304 accidents were considered in this study, 765 of which were severe crashes (i.e. including injury and fatal crashes). The total number of injured persons was 777, and in addition there were also 18 fatalities. In two-lane tunnels 1950 accidents were registered (670 of which were severe crashes) while in three-lane tunnels 354 accidents were counted (95 of which were severe crashes). Pedestrians and bicycles are forbidden to use motorway tunnels; thus pedestrians and bicycle were not involved in accidents. Table 1 gives accident count data observed during the 4-year monitoring period (severe crashes are given in brackets).

In the two-lane tunnels the width of lane was always 3.75 m; the presence of a side walk and an emergency lane are recorded in 147 and only 3 tunnels, respectively. In the three-lane tunnels the width of lane was 3.75 m in 26 tunnels and 3.50 m in the remaining 2 tunnels; the presence of side walk was recorded in 2 tunnels only, while the emergency lane is not present. With reference to an emergency lane, it is to be said that this lane was not generally present

Table 1
Accident count data observed during the 4-year monitoring period.

Year	Number of all (severe) accidents		Year's total
	Two-lane tunnels	Three-lane tunnels	
2006	671 (247)	119 (32)	790 (279)
2007	547 (181)	108 (33)	655 (214)
2008	388 (122)	72 (16)	460 (138)
2009	344 (120)	55 (14)	399 (134)
Total	1950 (670)	354 (95)	2304 (765)

in the existing tunnels of Italian motorways, in contrast with the corresponding open-motorway sections. This is due to reasons of tunnel construction costs. Since an emergency lane was present only in 3 of the investigated tunnels, this trait was not included among independent variables in subsequent statistical analyses.

Traffic flow was extracted from the traffic files of the Management Agencies of the aforementioned motorway tunnels. These files contained the annual average daily traffic (AADT) for one direction in each tunnel. AADT values ranging from 4500 to 40,760 vehicles per day were found for two-lane tunnels and AADT between 5030 and 32,260 vehicles per day were found for three-lane tunnels. The percentage of trucks was 14–31% and 17–23% for two and three-lane tunnels, respectively. Tunnel lengths also were extracted from files of the Motorway Management Agencies. Tunnel length values ranging from 387 to 3254 m were found for two-lane tunnels, while tunnel lengths between 524 and 4725 m were found for three-lane tunnels.

Summary statistics of tunnel length, AADT and percentage of trucks are given in Table 2.

4. Statistical analysis of data

4.1. Modelling accident counts

A basic way for describing the fluctuation of accident counts, say Y_i, occurring on a road section i during given time intervals (e.g. different years), is to assume that Y_i is a random variable (r.v.) with Poisson probability law. Let λ_i be the expected number of accidents per unit of time on section i. It is well known that a Poisson r.v. Y_i has $E(Y) = \lambda_i$ and $\mathrm{Var}(Y_i) = \lambda_i$. In many situations, however, accident counts appear to be "overdispersed" with respect to the theoretical variability consistent with the Poisson model. Hence, the Negative Binomial (NB) model,

$$f(y_i) = \frac{\Gamma(y_i+\varphi)}{y_i!\,\Gamma(\varphi)}\left[\frac{\lambda_i}{\lambda_i+\varphi}\right]^{y_i}\left[\frac{\varphi}{\lambda_i+\varphi}\right]^{\varphi} \qquad (1)$$

which has $E(Y_i) = \lambda_i$ and $\mathrm{Var}(Y_i) = \lambda_i\left(1+\lambda_i/\varphi\right)$, is often used, thus allowing for the variance of accident counts to be greater than the mean, provided that $1/\varphi > 0$.

A major objective of statistical analysis of accident counts is to estimate the expected number of accidents on a given section as a function of explanatory variables such as length, traffic flow, geometric design variables, and environmental conditions, which implies defining a regression model where the explanatory variables (and, possibly, combinations thereof) act as covariates. Let $\mathbf{x}$ be a vector of k covariates and $\boldsymbol{\beta}$ a vector of k (unknown) coefficients. For both the Poisson and the NB model, a regression model of the expected number of accidents is defined by $\lambda_i = g(\boldsymbol{x}_i;\boldsymbol{\beta})$, where $g(\cdot)$ denotes a certain function. In this paper the widely used regression model

$$\lambda_i = \exp\left(\boldsymbol{x}_i^T\boldsymbol{\beta}\right) \qquad (2)$$

was adopted.

Another relevant aspect concerns the form of building blocks of which the regression model is constituted. For this aim, in this paper, we used the approach proposed by Hauer and Banfo (1997), which makes use of the so-called Integrate-Differentiate (ID) method for recognizing a suitable functional form for building blocks and of the Cumulative Residuals (CURE) method for judging model adequacy. In particular, this approach was applied for identifying which functional form is more appropriate for the "AADT per lane" covariate. Furthermore, several traits can usually be taken into account as potential covariates. In particular, in the present application the following variables were considered: tunnel length, AADT per lane, percentage of trucks, number of lanes, presence of sidewalk, and year of observation. However, some of the variables chosen as potential covariates could actually have very little or even no effect on crash number. Thus, in this paper, in order to decide which subset of the full set of potentially explanatory variables should be included in the regression model, a procedure based on the Generalized Likelihood Ratio Test (GLRT) was used. For this purpose a Fortran code was implemented which maximizes the log-likelihood function by using the double precision routine DBCONF of the IMSL® MATH/LIBRARY (1989).

4.2. Jointly modelling of non-severe and severe accident counts

In crash data analyses it is often of interest to study the effect of potential covariates on multiple accident counts, e.g. when accident data by severity are considered. In such a case, the use of a univariate model for each count may lead to less precise estimates due to correlations that possibly exist among these observations, and multivariate regression models should be considered. In particular, in the present paper we are interested in modelling the number of "non-severe" and "severe" crashes, which are expected to be positively correlated random variables.

Let Y_{ji} denote crash counts of type j ($j = 1, 2$) observed on tunnel i ($i = 1, \ldots, n$) over the monitoring period. Assume that Y_{ji} ($j = 1, 2$) are independent mixed Poisson rv's with $E\left(Y_{ji}\right) = \lambda_{ji}\theta$ ($j = 1, 2$), where θ is a Gamma r.v. with $E(\theta) = 1$ and $\mathrm{Var}(\theta) = 1/\varphi$. By averaging the joint (conditional on θ) probability function of the Poisson r.v.'s Y_{ji} ($j = 1, 2$) over θ, the Bivariate Negative Binomial (BIVNB) model (Marshall and Olkin, 1990) is obtained

$$f(y_{1i}, y_{2i}) = \frac{\Gamma(y_i+\varphi)}{y_{1i}!\,y_{2i}!\,\Gamma(\varphi)}\,\frac{\lambda_{1i}{}^{y_{1i}}\lambda_{2i}{}^{y_{2i}}\varphi^{\varphi}}{(\lambda_i+\varphi)^{y_i+\varphi}} \qquad (3)$$

Table 2
Summary statistics of independent variables.

Variable		Mean	Mode	Standard deviation	Minimum	Maximum
Length (km)	Two-lane tunnels	1.188	0.529	0.637	0.387	3.254
	Three-lane tunnels	0.957	0.632	0.791	0.524	4.725
AADT/10,000 (veh/day)	Two-lane tunnels	1.512	0.613	0.837	0.450	4.076
	Three-lane tunnels	2.179	2.364	0.665	0.503	3.226
Percentage of trucks (%)	Two-lane tunnels	21.5	16	4	14	31
	Three-lane tunnels	22.1	23	2	17	23

where φ is the overdispersion parameter, and $y_i = y_{1i} + y_{2i}$, $\lambda_i = \lambda_{1i} + \lambda_{2i}$. This model has NB marginals with $E(Y_{ji}) \equiv \lambda_{ji} = \exp(\boldsymbol{x}_i^T \boldsymbol{\beta}_j)$ and $\mathrm{Var}(Y_{ji}) = \lambda_{ji}\left(1 + \lambda_{ji}/\varphi\right)$ $(j = 1, 2)$. Note that model (3) can account for a positive correlation only. This is not a limitation for the present case, however. We recall that there are many applications of bivariate/multivariate regression models in the literature, including among others: Subrahmaniam and Subrahmaniam (1973), Bijleveld (2005), Song et al. (2006), Park and Lord (2007), El-Basyouny and Sayed (2009a), Anastasopoulos et al. (2012a).

4.3. Modelling year effect on severe accident counts

From Table 1 it clearly appears that the number of accidents registered on tunnels are constantly and markedly decreasing during the 4-year monitoring period. Thus, an analysis of the number of severe accidents per year is also considered in order to capture the deterministic year effect.

Let Y_{ij} denote the number of accidents which occur on tunnel i $(i = 1, \ldots, n)$ in year j $(j = 1, \ldots, n_i)$. The yearly counts contributed by the same tunnel form a cluster. The potential problem for clustered count data is that the observations of the same tunnel may not be independent, so that the NB regression model might be inappropriate. Possible alternatives are models incorporating random effects, which are able to deal with temporal correlation in the data. In this paper we consider two of these regression models: the Negative Multinomial (NM) regression model (Guo, 1996), and the Random Effects Negative Binomial (RENB) regression model (Hausman et al., 1984).

Assume that each individual count Y_{ij} in the cluster i is a mixed Poisson r.v. with mean $\lambda_{ij}\theta_i$, where θ_i is assumed to be a Gamma variate with $E(\theta_i) = 1$ and $\mathrm{Var}(\theta_i) = 1/\varphi$. Note that in this model the random effect θ_i varies only across clusters, not the member of a cluster. Moreover, the model assumes that, conditional to θ_i, the count r.v.'s Y_{ij} $(j = 1, \ldots, n_i)$ are mutually independent. Then, on writing down the conditional joint probability function for cluster i, the unconditional joint probability function for Y_{ij} $(j = 1, \ldots, n_i)$ is obtained by integrating over θ_i. This joint probability function, given by

$$f(y_{i1}, \ldots, y_{in_i}) = \frac{\Gamma(y_i + \varphi)}{\Gamma(\varphi)}\left[\frac{\varphi}{\lambda_i + \varphi}\right]^{\varphi}\prod_{j=1}^{n_i}\frac{1}{y_{ij}!}\left[\frac{\lambda_{ij}}{\lambda_i + \varphi}\right]^{y_{ij}} \tag{4}$$

where $y_i = y_{i1} + \cdots + y_{in_i}$, $\lambda_i = \lambda_{i1} + \cdots + \lambda_{in_i}$, and $\lambda_{ij} = \exp(\boldsymbol{x}_{ij}^T \boldsymbol{\beta})$, is known as the Negative Multinomial distribution with $E(Y_{ij}) = \lambda_{ij}$, $Var(Y_{ij}) = \lambda_{ij}\left(1 + \lambda_{ij}/\varphi\right)$ and $Cov(Y_{ij}, Y_{ik}) = \lambda_{ij}\lambda_{ik}/\varphi$.

Alternatively, assume that Y_{ij} $(j = 1, \ldots, n_i)$ are NB r.v.'s with parameters $(\gamma_{ij} = \mathrm{e}^{\boldsymbol{x}_{ij}^T \boldsymbol{\beta}}, \delta_i)$ so that $E(Y_{ij}) = \gamma_{ij}/\delta_i$ and $Var(Y_{ij}) = \gamma_{ij}\left(1 + \delta_i\right)/\delta_i^2$, where δ_i is a random effect across clusters. Furthermore, assume that, conditional to δ_i, the count r.v.'s Y_{ij} $(j = 1, \ldots, n_i)$ are mutually independent, and the ratio $\delta_i/\left(1 + \delta_i\right)$ is distributed as a Beta r.v. with parameters (a, b). Then, on writing down the conditional joint probability function for cluster i, the unconditional joint probability function for Y_{ij} $(j = 1, \ldots, n_i)$ is obtained by integrating over δ_i. This joint probability function, given by

$$f(y_{i1}, \ldots, y_{in_i}) = \frac{\Gamma(a+b)\Gamma(a+\gamma_i)\Gamma(b+y_i)}{\Gamma(a)\Gamma(b)\Gamma(a+b+\gamma_i+y_i)}\prod_{j=1}^{n_i}\frac{\Gamma(y_{ij}+\gamma_{ij})}{y_{ij}!\,\Gamma(\gamma_{ij})} \tag{5}$$

where $y_i = y_{i1} + \cdots + y_{in_i}$, $\gamma_i = \gamma_{i1} + \cdots + \gamma_{in_i}$, and $\gamma_{ij} = \exp(\boldsymbol{x}_{ij}^T \boldsymbol{\beta})$, is known as the Random Effect Negative Binomial regression model (Hausman et al., 1984). Within this model, $E(Y_{ij}) = b\gamma_{ij}/(a-1)$, and $Var(Y_{ij}) = \gamma_{ij}b(a+b-1)[(a-1)(a-2)]^{-1}[1 + \gamma_{ij}/(a-1)]$ which can be recast as $Var(Y_{ij}) = (a+b-1)(a-2)^{-1}E(Y_{ij})[1 + E(Y_{ij})/b]$. An interesting property of model (5) that, at the best of our knowledge, was not previously stated, is its convergence to model (4) as $a \to \infty$. In such a case the parameter b plays the same role as the overdispersion parameter φ in the NM model. The RENB model is potentially a more flexible model than NM one, since it uses a two-parameters probability density to model the random effects.

4.4. Choosing the functional form for traffic volume

A preliminary exploratory data analysis was carried out in order to assign a suitable functional form to AADT per lane (AADT_L), which represents the most influential variable in predicting the number of accidents. By following the Integrate-Differentiate method in its generalization to more than one explanatory variable, the Empirical Integral Function (EIF) for the traffic volume per lane was derived in the case of severe accidents. Since this function showed to have two inflection points: one at an AADT_L of about 5000 vehicles per day and another one at an AADT_L of about 13,000 vehicles per day, simple forms such as x^{β} and $e^{\beta x}$ appeared to be inadequate for describing the dependence of the expected number of accidents on traffic volume over the entire range of the observed AADT_L values in the present situation. In fact, it was found that the relative error of prediction when using the power model, although smaller than that for the exponential model, was still high in the regions ($\mathrm{AADT}_L < 5000$) and ($\mathrm{AADT}_L > 13{,}000$), with underestimation in the former one and overestimation in the latter one.

Thus, in order to get a better fit to tunnel data, in the present paper we propose the following functional form for the AADT_L variable

$$g(AADT_L) = {AADT_L}^{\beta_1}\,\mathrm{e}^{\beta_2 D_1}\,\mathrm{e}^{\beta_3 D_2} \tag{6}$$

where D_1 is a dummy variable which is 1 when the AADT_L is fewer than 5000 vehicles per day and 0 otherwise, while D_2 is a dummy variable which is 1 when the AADT_L is greater than 13,000 vehicles per day and 0 otherwise. When using this model, the predicted number of accidents computed in the said intervals of traffic volume practically coincides with the observed number of accidents. Moreover, it was also found that, under this model, the CURE line is entirely contained within two standard deviations ($\pm 2\sigma^*$). Thus, the subsequent analysis of the dataset of 260 road tunnels was carried out on the basis of the modified power function (6).

5. Estimation results

The data set for road tunnels used in the subsequent statistical analysis consists of annual number of non-severe and severe accidents registered in $n = 4$ years from 2006 to 2009 in $N = 260$ tunnels. The candidate set of explanatory variables is: length (L), annual average daily traffic per lane (AADT_L), percentage of trucks (%Tr), number of lanes (N_L), presence of sidewalk. Moreover, since it appears that the number of accidents registered on tunnels are constantly and markedly decreasing during the 4-year monitoring period (see Table 1), dummy variables year 2007, year 2008, and year 2009 are also considered to capture the deterministic year effect for severe accidents only. The reference year for these dummy variables is year 2006.

5.1. Analysis of non-severe and severe accident counts

Non-severe and severe accident counts observed on tunnel i $(i = 1, \ldots, n)$ over the 4-year monitoring period are jointly analyzed by using the BIVNB regression model (3). For comparison purposes, the estimates obtained by applying the univariate NB regression model are also reported in Table 3. For both models the estimated coefficients have the expected sign. The number of both

Table 3
Parameters of bivariate and univariate negative binomial regression models for non-severe and severe accidents.

	Variables	Bivariate NB			Univariate NB		
		Point estimate	Standard error	LRT statistic	Point estimate	Standard error	LRT statistic
Non-severe accidents	Constant	3.64063	0.03226	77.21	3.56850	0.03128	86.12
	Log of length in km	0.65747	0.07146	53.66	0.65883	0.06914	62.02
	Log of AADT per lane/10,000	1.45384	0.05707	87.99	1.45969	0.05461	102.51
	D1 (1 if AADT per lane < 5000, 0 otherwise)	0.30106	0.07840	3.27	0.30687	0.07447	4.07
	D2 (1 if AADT per lane > 13,000, 0 otherwise)	−0.17063	0.09833	1.31	−0.15804	0.09799	1.48
	Log of percentage of trucks/100	1.04928	0.02057	20.24	1.01217	0.02000	22.92
	Number of lanes (0 two lanes, 1 three lanes)	0.69753	0.08413	29.20	0.72422	0.08030	36.94
	Sidewalk (1 if present, 0 if absent)	0.06850	0.04766	0.39	0.08939	0.04567	0.80
	Overdispersion parameter	–			14.82429	4.52452	
	Log-likelihood	–			−604.113		
Severe accidents	Constant	2.59513	0.04060	25.08	2.51813	0.04778	20.07
	Log of length in km	0.47560	0.09466	18.06	0.51289	0.10812	17.89
	Log of AADT per lane/10,000	2.08129	0.08492	112.38	2.02511	0.09270	83.64
	D1 (1 if AADT per lane < 5000, 0 otherwise)	0.61642	0.11962	7.20	0.55669	0.12666	5.32
	D2 (1 if AADT per lane > 13,000, 0 otherwise)	−0.63012	0.11231	12.85	−0.59154	0.14268	9.31
	Log of percentage of trucks/100	0.69250	0.02551	5.42	0.64402	0.03017	4.02
	Number of lanes (0 two lanes, 1 three lanes)	0.29956	0.11267	3.28	0.32081	0.13352	3.29
	Sidewalk (1 if present, 0 if absent)	−0.07900	0.06551	0.32	−0.0974	0.07371	0.42
	Overdispersion parameter	8.08584	1.48524		5.53584	1.45287	
	Log-likelihood	−1068.789			−486.921		

non-severe and severe accidents occurring in tunnels over the 4-year monitoring period increases with average daily traffic per lane, tunnel length, number of lanes, and percentage of trucks. Moreover, a multiplicative coefficient greater than 1 ($e^{\beta_2 D_1}$) is associated to a small traffic volume ($AADT_L < 5000$), whilst a multiplicative coefficient lower than 1 ($e^{\beta_3 D_2}$) is associated to a high traffic volume ($AADT_L > 13{,}000$). For both regression models all variables are significant at the 10% level except the presence of sidewalk, which does not appear to be statistically significant both for non-severe and severe crashes, and the dummy variable associated to an AADT per lane greater than 13,000 which does not appear to be statistically significant for non-severe crashes. Table 3 shows the parameter estimates at convergence of the log-likelihood maximization procedure, and the corresponding standard errors, for both the models considered in the analysis. The significance of the covariates was evaluated by using the likelihood-ratio statistic $\Lambda = 2\left[l(\hat{\boldsymbol{\beta}}, \hat{\varphi}) - l(\hat{\boldsymbol{\beta}}', \beta_k = 0, \hat{\varphi})\right]$ (as suggested, for example, by Lawless, 1987), where $l(\hat{\boldsymbol{\beta}}, \hat{\varphi})$ is the log-likelihood of the regression model containing all the covariates, and $l(\hat{\boldsymbol{\beta}}', \beta_k = 0, \hat{\varphi})$ is the log-likelihood of the regression model with the k-th covariate out. The Λ statistic is asymptotically distributed as a chi-squared distribution with 1 degree of freedom, thus the Λ threshold for a variable to be significant at 10% level is 2.71.

It can be observed from Table 3 that the bivariate and the univariate models give similar results in terms of the estimated model coefficients and their significance. It is to be noted, however, that the log-likelihood at the convergence of the bivariate model is markedly higher than the sum of the log-likelihoods of the univariate models for non-severe and severe accidents, thus suggesting that the bivariate model is more appropriate.

5.2. Exploratory analysis of unobserved heterogeneity for severe accident counts

In order to account for the possible presence of unobserved heterogeneity, an exploratory analysis was also carried out for severe crashes on the basis of the 4-year accident counts for assessing if random-parameters should to be considered in the regression model. For this aim, each parameter in the NB regression model was in turn assumed to be a normal random variable with unknown mean and variance. As, for the parameter variance tending to zero, the single random-parameter model converges to the fixed-parameters one, the presence of heterogeneity for each variable can be statistically assessed by means of the likelihood ratio test. This analysis (not reported here to save space) showed that the hypothesis of fixed parameters can be accepted in the investigated case for all of the variables considered, with the exception of the dummy variable associated to a small traffic volume (AADT < 5000 vehicles per day). While this exploratory investigation appears to support the fixed-parameters NB model, it is to be said that these results are based on specific observations that might not be directly transferable to other cases or geographic areas where traffic, tunnels geometry, and driving style are different from those in Italy. Therefore, the potential of random-parameters models should be investigated in greater depth by future studies. In particular, a possible direction of expansion of the present work could be an analysis based on combined random-effects and random-parameters models in which the estimated random parameters (as the random effects) are fixed within clusters (i.e., the same tunnel) and vary between them. In this respect it is to be mentioned, for example, that Greene (2007) has developed estimation procedures for incorporating random parameters in count-data models.

5.3. Analysis of year effect on severe accident counts

In order to analyze the effect of the year on severe accident counts, the RENB (Random Effects Negative Binomial) and the NM (Negative Multinomial) regression models were applied. For comparison purposes, the estimates obtained by applying the NB regression model are also reported. From Table 4, it results that all these models give similar results in terms of expected signs for regression coefficients and significant variables, confirming average daily traffic per lane, length, number of lanes, and percentage

Table 4
Parameters of random effects negative binomial regression model and Negative Multinomial regression model.

Variables	Negative Binomial			Negative Multinomial			RENB
	Point estimate	Standard error	LRT statistic	Point estimate	Standard error	LRT statistic	Point estimate
Constant	1.24167	0.03671	8.81	1.36744	0.04755	6.59	11.84326
Log of length in km	0.47161	0.08704	24.41	0.51118	0.10757	18.01	0.51142
Log of AADT per lane/10,000	2.10943	0.07945	159.38	2.04340	0.09238	90.97	2.04377
D1 (1 if AADT per lane < 5000, 0 otherwise)	0.67292	0.11167	11.01	0.59351	0.12378	6.46	0.59376
D2 (1 if AADT per lane > 13,000, 0 otherwise)	−0.61715	0.09398	22.16	−0.59533	0.12954	12.01	−0.59518
Log of percentage of trucks/100	0.49086	0.02300	4.17	0.55578	0.03006	3.21	0.55586
Number of lanes (0 two lanes, 1 three lanes)	0.36733	0.10392	6.78	0.33391	0.13306	3.64	0.33416
Sidewalk (1 if present, 0 if absent)	−0.04726	0.06071	0.16	−0.10432	0.07337	0.51	−0.10393
Year 2007 (dummy 0,1)	−0.25784	0.04388	7.74	−0.26022	0.07277	8.16	−0.26026
Year 2008 (dummy 0,1)	−0.70508	0.05087	46.32	−0.70455	0.08862	48.68	−0.70443
Year 2009 (dummy 0,1)	−0.70325	0.05406	44.03	−0.69846	0.08983	44.75	−0.69825
Overdispersion parameter	41.11389	72.485		5.64035	1.49851		
a							200,000
b							5.64140
Log-likelihood	−1016.762			−999.303			−999.306

of trucks as significant variables, and the presence of a sidewalk as not significant. In addition the year covariate is found to be highly significant. It is worth noting, however, that the log-likelihood at convergenge of both RENB and NM models is markedly higher than the log-likelihood of NB model, thus suggesting that models suitable to account for clustering effects are more appropriate in the present situation. It is also worth mentioning that RENB and NM models give in practice the same results in terms of both estimates of model coefficients and log-likelihood at convergence as can be readily seen in Table 4. In particular, note that the supremum of the log-likelihood function of the RENB model is attained at a huge value of parameter a. As noticed in a previous section, for very large values of a the RENB model converges to the NM model with $b \equiv \varphi$, which should imply $\hat{b} \cong \hat{\varphi}$ as has been verified in this case. Thus, we may conclude that in the present situation the RENB model and the NM model come to the same results. As such, for saving space in Table 4, standard errors and LRT for RENB model, which are in practice the same of NM model, are not reported.

The fraction of accidents in the years from 2007 to 2009 with respect to year 2006 predicted by all of the models considered via the dummy variables resulted 0.771, 0.494 and 0.497, which are very close to the corresponding observed fractions 0.767, 0.495 and 0.480.

5.4. Comments to estimation results

A year effect consisting in a systematic reduction in severe crashes over time was found in this paper for motorway tunnels. Likewise, a severe accident decreasing trend over time had been found on the motorways containing the tunnels investigated (Caliendo and De Guglielmo, 2012). The reduction of severe crashes both in tunnels and on the corresponding motorways containing these structures might be attributable to an increasing installation of electronic speed control systems (Tutor) on Italian motorways (Autostrade, 2012) that record the average speed of vehicle on the basis of the time spent to cover a given distance. The tutor makes it possible for the Road Police to check whether the speed limit along the motorway has been violated (if not otherwise indicated this speed limit on the Italian motorways is 130 km/h or 110 km/h when it is raining). Additionally, also the driving licence with a demerit point system, introduced in Italy in 2003, might have contributed towards reducing accidents over time. Point penalties are associated with violations of traffic laws and, when all points are exhausted, the driving license is revoked. Finally, a reduction in accidents also may be due to the training of drivers based on information campaigns to control speed and encourage respect of speed limits. However, severe crashes in the tunnels investigated if compared to severe accidents on motorways (see Caliendo and De Guglielmo, 2012) seem to decrease over time at a slightly greater rate than on motorway sections, which might be attributable to the implementation and/or reinforcement of some facilities in tunnels after October 2006, the date of the coming into force in Italy of the European Directive 2004/54/EC.

The results of the bivariate statistical analysis shows that the number of both non-severe and severe crashes occurring in unidirectional motorway tunnels increases with the tunnel length, the annual average daily traffic per lane, the percentage of trucks, and the number of lanes. These results tend to confirm the hypothesis that longer tunnels are associated with a greater number of accidents due to the drivers' diminishing concentration with increasing length; and that, in free-flowing conditions, as AADT_L and/or %Tr increase, the frequency of lane changing and overtaking movements increase so that more accident are expected. Additionally, by means of an increasing number of lanes, the opportunities for lane change increase so that more traffic conflicts and consequently more accidents should also be expected.

However, apart from the goodness-of-fit of the crash-prediction models that have been developed in this paper, the aforementioned effects of geometry and traffic on crash frequency in road tunnels are found also to have some support in the literature and more especially in Lemke (2000) with reference to tunnel length and Salvisberg et al. (2004) for the AADT and heavy vehicles travelling in tunnels. Additionally, it is to be said that a further support might also be found in other studies available in the literature even if they are related to open roads. For example Milton and Mannering (1998) found that the increasing of the section length on open roads, AADT per lane, and number of lanes leads to more accidents. Abdel-Aty and Essam Radwan (2000) also found that crashes increase with an increase in the section length and AADT per lane. Noland and Oh (2004) showed that the increasing in the number of lanes appears to be associated with increased accidents. Caliendo et al. (2007) showed that crash frequency increases with the section length and AADT. Wang et al. (2009) found that the segment length and the AADT are positively associated with accidents in all prediction models, and the number of lanes is positively associated with the frequency of slight injury accidents. Vadlamani et al. (2011) also found that crash frequency increases as the truck percentage increases.

In the light of the aforementioned results and considerations, the authors are confident that through carrying out a rigorous statistical analysis they have developed models for realistically predicting the number of non-severe and severe crashes in Italian motorway tunnels.

6. Summary and conclusions

The writing of this paper was primarily motivated by the need to quantify, for the first time for Italian motorway tunnels, the effects on the expected number of non-severe and severe crashes of all the following variables: tunnel length, traffic flow, percentage of trucks, number of lanes, and sidewalks, with a view to suggesting countermeasures for improving tunnel safety. A further point of interest was to detect the statistical regression tool for treating accidents in tunnels.

On the basis of the 4-year monitoring period, extending from 2006 to 2009, considered for $N = 260$ motorway tunnels with unidirectional traffic and having two or three lanes, the Bivariate Negative Binomial regression model (jointly applied to non-severe and severe crashes) has shown that the number of both non-severe and severe accidents occurring in tunnels increases with the tunnel length, the annual average daily traffic per lane, the percentage of trucks and the number of lanes. In contrast, the sidewalk variable was not found to be statistically significant. Results have confirmed that, due to the drivers' diminishing concentration with increasing L, more accident may be expected in longer tunnels. In free-flowing conditions, as AADT_L and/or %Tr increase, the frequency of lane changing and overtaking movements also increase, so that more accidents may be expected. Finally, given that an increase in the number of lanes increases the opportunities for lane change, it is not surprising to find more accidents in these cases. An exploratory analysis carried out for severe crashes over the 4 years accident data appears to support that the hypothesis of fixed parameters may be justified in the investigated case for all of the variables considered, with the exception of the dummy variable associated to a small traffic volume. However, it is to be said that the alternative approach based on random-parameters model should be investigated in greater depth by means of further study that make such additional progress possible.

A significant year effect consisting in a systematic reduction of severe crashes in motorway tunnels was also found by means of the Random Effects Negative Binomial regression model and the Negative Multinomial regress model. This reduction might be attributable to: an increasing installation of electronic speed control systems on the motorways containing the tunnels investigated; the positive effects of the introduction of the driving licence with the demerit point system in the event of violation of the Highway Code, and to the implementation and/or reinforcement of some facilities in tunnels after October 2006, the date of the coming into force in Italy of the European Directive 2004/54/EC.

In formulating the regression models, it was also found that, in the present situation, simple forms such as $\text{AADT}_L{}^{\beta}$ and $e^{\beta \text{AADT}_L}$ appear to be inadequate for describing the dependence of the expected number of accidents on traffic volume over the entire range of observed values, as shown by an ad hoc analysis based on the Cumulative Residual method, even if the power law function appears to be the more suitable of two. Thus, in order to get a better fit, a form of function which incorporates a correction for small ($\text{AADT}_L < 5000$) and high ($\text{AADT}_L > 13{,}000$) traffic volumes per lane has been proposed, which has been tested by Cumulative Residual method, showing a much better fit than the simple power function model.

In the light of the aforementioned results, the models developed for Italian motorway tunnels appear to be useful for many applications such as the estimation of non-severe and severe accident reductions due to improvements of the existing tunnels or to modifications of traffic control systems, as well as for the prediction of accident counts when comparing different tunnel design options. Thus, the authors are reasonably confident that this research may represent a point of reference for the engineers in adjusting or designing road tunnels. Despite the potential benefits of models for predicting crashes in road tunnels, the authors believe that the goal of having safer tunnels is expected to be achieved, among other means, also through novel Intelligent Transport System (ITS) that communicate to users in real time as they have to drive through the tunnel. Therefore, research also needs to be addressed towards further studies that make these additional developments possible.

Acknowledgement

The authors would like to thank the two anonymous referees whose comments proved to be invaluable.

References

Abdel-Aty, M.A., Essam Radwan, A., 2000. Modeling traffic accident occurrence and involvement. Accident Analysis and Prevention 32, 633–642.
Amundsen, F.H., Ranes, G., 2000. Studies on traffic accidents in Norwegian road tunnels. Tunnelling and Underground Space Technology 15 (1), 3–11.
Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modelling vehicle accidents frequencies with random-parameters count models. Accident Analysis and Prevention 41, 153–159.
Anastasopoulos, P.C., Shankar, V.N., Haddock, J.E., Mannering, F.L., 2012a. A multivariate tobit analysis of highway accident-injury-severity rates. Accident Analysis and Prevention 45, 110–119.
Anastasopoulos, P.C., Mannering, F.L., Shankar, V.N., Haddock, J.E., 2012b. A study of factors affecting highway accident rates using the random-parameters tobit model. Accident Analysis and Prevention 45, 628–633.
Autostrade, 2012. Safety objective. www.autostrade.it/pdf/Sicurezza_WEB_Mid_Res.pdf
Bijleveld, F.D., 2005. The covariance between the number of accidents and the number of victims in multivariate analysis of accident related outcomes. Accident Analysis and Prevention 37, 591–600.
Caliendo, C., Guida, M., Parisi, A., 2007. A crash-prediction model for multilane roads. Accident Analysis and Prevention 39, 657–670.
Caliendo, C., Ciambelli, P., De Guglielmo, M.L., Meo, M.G., Russo, P., September 2012. Numerical simulation of different HGV scenarios in curved bi-directional road tunnels and safety evaluation. Tunnelling and Underground Space Technology 131, 33–50.
Caliendo, C., De Guglielmo, M.L., 2012. Accident rates in road tunnels and social costs evaluation. SIIV – 5th International Congress – Sustainability of Road Infrastructures. In: Procedia-Social and Behavioral Sciences, Vol. 53, pp. 166–177, www.sciencedirect.com/science/article/pii/S1877042812043327
Chin, H.C., Quddus, M.A., 2003. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. Accident Analysis and Prevention 35, 253–259.
El-Basyouny, K., Sayed, T., 2009a. Collision prediction models using multivariate Poisson-lognormal regression. Accident Analysis and Prevention 41 (4), 820–828.
El-Basyouny, K., Sayed, T., 2009b. Accident prediction models with random corridor parameters. Accident Analysis and Prevention 41 (5), 1118–1123.
European Parliament and Council, 2004. Directive 2004/54/EC. Official Journal of the European Union. L.167, Bruxelles, 30 April.
Greene, W., 2007. Limdep, Version 9.0. Econometric Software Inc., Plainview, NY.
Guo, G., 1996. Negative multinomial regression models for clustered events counts. Sociological Methodology 26, 113–132.
Hauer, E., Banfo, J., 1997. Two tools for finding what function links the dependent variable to the explanatory variables. In: Proc. International Cooperation on Theories and Concepts in Traffic Conference, Lund, Sweden, November.
Hauer, E., 2004. Statistical road safety modeling. In: 83rd TRB Annual Meeting, Washington, DC, USA, January 11–15.
Hausman, J., Hall, B.H., Griliches, Z., 1984. Econometric models for count data with an application to the patents-R&D relationship. Econometrica 52 (4), 909–938.
IMSL® MATH/LIBRARY, 1989. Fortran Subroutines for Mathematical Applications. IMSL, United States.
Kircher, K., Ahlstrom, C., 2012. The impact of tunnel design and lighting on the performance of attentive and visually distracted drivers. Accident Analysis and Prevention 47, 153–161.
Lawless, J.F., 1987. Negative binomial and mixed Poisson regression. The Canadian Journal of Statistics 15 (3), 209–225.
Lemke K., 2000. Road safety in tunnel. Transportation Research Record 1740, Paper No. 00-0155.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation research Part A 44, 291–305.

Manser, M.P., Hancock, P.A., 2007. The influence of perceptual speed regulation on speed perception, choice, and control: tunnel wall characteristics and influence. Accident Analysis and Prevention 39, 69–78.

Marshall, A.W., Olkin, I., 1990. Multivariate distributions generated from mixtures of convolutions and product families. Lecture Notes-Monographs Series 16, 371–393.

Milton, J., Mannering, F., 1998. The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. Transportation 25, 395–413.

Milton, J., Shankar, V., Mannering, F., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accident Analysis and Prevention 40 (1), 260–266.

Noland, R.B., Oh, L., 2004. The effect of infrastructure and demographic change on traffic-related fatalities and crashes: a case study of Illinois county-level data. Accident Analysis and Prevention 36, 525–532.

Nussbaumer, C., 2007. Comparative analysis of safety in tunnels. Young Researchers Seminar 2007, Brno. www.ectri.org/YRS07/Papiers/Session-9/Nussbaumer.pdf

Park, E.S., Lord, D., 2007. Multivariate Poisson-lognormal models for jointly modelling crash frequency by severity. Transportation Research Record 2019, 193–197.

PIARC, 1995. Road safety in tunnels. PIARC Technical Committee C5 Road Tunnels. http://www.piarc.org

PIAR, 2008. Human factors and road tunnel safety regarding users. PIARC Technical Committee C3.3, Road Tunnel Operation, Report R17, Paris. ISBN 2-84060-218-0. http://www.piarc.org

SAFESTAR, 2002. Safety Standards for Road Design and Redesign. Final Report. Coordinated by SWOV, funded by the European Commission.

Salvisberg, U., Allenbach, R., Cavegn, M., Hubacher, M., Siegrist, S., 2004. Verkehrssicherheit in Autobahn- und Autostrassentunneln des Nationalstrassennetzes. BFU- Report, Bern. ISBN 3-908192-17-X.

Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accident Analysis and Prevention 43, 1666–1676.

Shankar, V.N., Albin, R.B., Milton, J.C., Mannering, F.L., 1998. Evaluating median crossover likelihoods with clustered accident counts: an empirical inquiry using the random effects negative binomial model. Transportation Research Record 1635, 44–48.

Shankar, V.N., Ulfarsson, G.F., 2003. Accident count model based on multiyear cross-sectional roadway data with serial correlation. Transportation Research Record 1840, 193–197.

Song, J.J., Ghosh, M., Miaou, S., Mallick, B., 2006. Bayesian multivariate spatial models for roadway traffic crash mapping. Journal of Multivariate Analysis 97 (1), 246–273.

Subrahmaniam, K., Subrahmaniam, K., 1973. On the estimation of the parameters in the bivariate negative binomial distribution. Journal of the Royal Statistical Society: Series B 35, 131–146.

SWOV, 2009. The road safety of motorway tunnels. SWOV Fact Sheet, Leidschendam, the Netherlands.

Törnoros, J., 1998. Driving behaviour in real and a simulated road tunnel – a validation study. Accident Analysis and Prevention 30 (4), 497–503.

Vadlamani, S., Chen, E., Ahn, S., Washington, S., 2011. Identifying large truck hot spots using crash counts and PDOEs. Journal of Transportation Engineering 137 (1), 11–21.

Vashitz, G., Shinar, D., Blum, Y., 2008. In-vehicle information systems to improve traffic safety in road tunnels. Transportation Research Part F 11, 61–74.

Wang, C., Quddus, M.A., Ison, S.G., 2009. Impact of traffic congestion on road accidents: a spatial analysis on the M25 motorway in England. Accident Analysis and Prevention 41, 798–808.