Safety Science 50 (2012) 1732–1739

Contents lists available at SciVerse ScienceDirect

Safety Science

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

Estimating the safety performance of urban intersections in Lisbon, Portugal

Sandra Vieira Gomes [a,*], Srinivas Reddy Geedipally [b,1], Dominique Lord [c,2]

[a] National Laboratory of Civil Engineering, Transportation Department, Av. Brasil, 101, 1700-066 Lisbon, Portugal
[b] Texas Transportation Institute, 3135 TAMU, College Station, TX 77843-3135, United States
[c] Zachry Department of Civil Engineering, Texas A&M University, 3136 TAMU, College Station, TX 77843-3136, United States

ARTICLE INFO

Article history:
Received 9 August 2011
Received in revised form 5 March 2012
Accepted 30 March 2012
Available online 23 May 2012

Keywords:
Crash prediction models
Generalized linear modeling
Road accidents
Urban intersections
Lisbon

ABSTRACT

According to official statistics, a large percentage of crashes in Portugal are reported on urban roads. For instance, from 2004 to 2007, about 70% of all injury accidents and 43% of the fatalities occurred inside urban agglomerations. This important safety problem has also been observed on the urban network of Lisbon. Understanding this significant problem, the Government of the Portuguese Republic via its research grant agency – The Foundation for Science and Technology – funded a project whose primary objective consists of developing tools that would help estimating the safety performance of various components of the urban highway system in Lisbon. This paper documents one component of the safety tools that were developed and describes the steps that were taken to develop predictive models for estimating the safety performance of signalized and unsignalized intersections of Lisbon. Several crash predictive models were developed using the Poisson-gamma modeling framework. Two types of models were estimated: flow-only and models with covariates. They were estimated using crash and other related data collected at 44 three-legged and 50 four-legged intersections for the years 2004–2007, inclusively. It was found that some highway geometric design characteristics were associated with the crashes occurring at urban three- and four-legged intersections in Lisbon.

1. Introduction

According to official statistics, a large percentage of crashes in Portugal are reported on urban roads. For instance, from 2004 to 2007, about 70% of all injury accidents and 43% of the fatalities occurred inside urban agglomerations (DGV, 2004, 2005, 2006; ANSR, 2007). This important safety problem has also been observed on the urban network located in and around Lisbon. On the urban network in Lisbon, more than 2400 crashes occurred during the same time period, and about 30% of all crashes involved a pedestrian.

In order to understand this significant problem, the Government of the Portuguese Republic funded a project via its research grant agency – The Foundation for Science and Technology – which had, as the primary objective, the development of tools that would help estimating the safety performance of various components of the urban highway system in Lisbon. No such tools exist in Portugal for estimating the safety performance of urban networks, hence the funding of this project. The research project titled "IRUMS – Safer Roads in Urban Areas" is carried out by the National Laboratory of Civil Engineering (LNEC) jointly with the Department of Engineering at the University of Coimbra, Coimbra. This project intends to develop methods for managing the safety of urban road networks, particularly those applied in Lisbon. The methods focused on estimating the expected crash frequencies, the identification of hazardous sites (or sites with promise) and subsequently select effective countermeasures to reduce the number and severity of crashes.

This paper documents one component of the safety tools that are currently under development. More specifically, this paper describes the steps that were taken to develop predictive models for estimating the safety performance of signalized and unsignalized intersections located in Lisbon. Several crash predictive models (CPMs) were developed using the Poisson-gamma modeling framework. Two types of models were estimated: flow-only and models with covariates. They were estimated using crash and other related data collected at 44 three-legged and 50 four-legged intersections for the years 2004–2007.

The paper is organized as follows. The first section provides a brief background about existing statistical models developed in Portugal and elsewhere in Europe. The second section describes the methodology used for estimating the CPMs. The third section presents the characteristics of the data used in this study. The fourth section summarizes the modeling results. The last section provides a summary of the work accomplished so far in this project.

* Corresponding author. Tel.: +351 218443528; fax: +351 8443029.
E-mail addresses: sandravieira@lnec.pt (S. Vieira Gomes), srinivas8@tamu.edu (S.R. Geedipally), d-lord@tamu.edu (D. Lord).
[1] Tel.: +1 979 862 1651; fax: +1 979 845 6006.
[2] Tel.: +1 979 458 3949; fax: +1 979 845 6481.

0925-7535/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ssci.2012.03.022

2. Background

As discussed in previous work (see Lord et al., 2005 and Lord and Mannering, 2010), there has been a significant amount of research done on the development and application of crash prediction models for various types of highway safety analyses. Since design standards and operational characteristics (e.g., vehicle size, etc.) are obviously very different between Europe and other places around the world, especially in North America, this background section focuses on predictive models that have estimated and applied in European countries.

Hall (1986) estimated several CPMs for four-arm single carriageway urban signalized intersections in the UK The models were estimated using 177 intersections with a speed limit equal to or above 30 miles/h in urban areas. The author concluded that wider approaching lanes were associated with a larger number of right-angle crashes; a greater number of lanes were associated with higher pedestrian accident rates; and the sight distance was found to be significantly associated with left turn accidents. The author also concluded that the increased displacement of the opposite arm to the left or right was associated with lower accident rates.

Mountain and Fawaz (1996) created CPMs using data from 662 intersections with different types of traffic control systems in the UK. The authors developed models relating crashes with traffic flows and other variables, namely: method of control, road class, carriageway type (single or dual), number of arms and speed limit. The authors concluded that only the method of control had a significant effect on crash occurrence, although the best fitted models were ones with traffic flow as the only explanatory variable.

Greibe (2003) developed CPMs for road segments and urban intersections with three or four legs and with or without traffic signals in Denmark. For the intersections' models, the author found that the geometric variables were not significantly linked to the occurrence of crashes. This was attributed to the complicated internal correlations in the intersection design data, or a lack of good descriptive variables. The estimated accident prediction models for road links were capable of describing more than 60% of the systematic variation ('percentage-explained' value) while the models for junctions had lower values. The significant variables found in the study were: speed limit, road width, number of exits per km, number of minor side roads per km, parking and land use.

Brüde and Larson (1993) estimated models for pedestrian and cyclist crashes occurring at intersections in Sweden. Data from 285 intersections were used for modeling crashes involving pedestrians, and data from 432 intersections were used for modeling crashes with cyclists. The purpose of this study was to develop CPMs and illustrate their predictive capabilities. They used volumes of motorized vehicles, pedestrians and cyclists as explanatory variables. Interestingly, the authors concluded that the models with a low R2 value may have high predictive capabilities.

Reurings and Janssen (2006) developed CPMs for road segments in The Netherlands. Crash data from 524 km of roads on urban and rural areas were used for developing the models. The conclusions of this study were: carriageways inside urban areas with AADT <25000 generally have a lower crash rate than carriageways outside urban areas; carriageways with a speed limit of 50 km/h or 80 km/h and one driving direction have a lower crash rate than carriageways with the same speed limit but with two driving directions; the average crash rate of urban carriageways with a speed limit of 70 km/h is lower than the crash rate of carriageways with a speed limit of 50 km/h; and the average crash rate of rural carriageways with a speed limit of 60 km/h has almost the same crash rate as that of rural carriageways with a speed limit of 80 km/h and with two driving directions.

Although crash modeling is widely spread all over the Europe, Portugal has no such tools for estimating the safety performance of urban networks. The few models that were estimated by Azeredo Lopes and Cardoso (2007, 2007a, 2007b) in Portugal were related to road segments on motorways and rural areas. In their study, they used crash data collected from 1999 to 2004. Several exploratory variables were considered, such as number of lanes, type and condition of shoulders and medians, and the presence of an additional lane. The authors concluded that on Portuguese motorways all the variables except the number of lanes were significant. However, for single carriageway rural roads with a median, the number of lanes had a significant effect on crash occurrences.

3. Data description

The data collected at signalized and unsignalized intersections included the following: geometric design characteristics, crash data (severity, manner of collision, etc.) and traffic volumes. With the exception of the crash data, all the data were obtained from on-site visits. All the intersections which had missing traffic flow volumes were excluded from the analysis. Thus, given the costs associated with the data collection process, only 94 intersections were finally used (44 three-legged and 50 four-legged).

Fig. 1 shows the graphical representation of the Lisbon road network and the location of the intersections used in this study. The three-legged intersections are shown with a 'triangle' whereas four-legged intersections with a 'square'. The dark colored intersections indicate that traffic flows were either counted manually or using an automated system. The light colored intersections have the traffic flows that were estimated by the models developed by Martinez (2006).

Injury accident data were collected from the official Portuguese accident statistics database and police reports. The crashes were then geocoded in a GIS database. The main reason for using both sources of crash data is that the official accident statistics database includes only the street names, but not the exact location where the accident occurred within or near the intersection. To overcome this problem, the sketch of the accident location from the police report was projected onto the road network (intersections and segments). This allowed for identifying the location of each accident for most of the cases.

Fig. 2 presents the spatial distribution of all injury accidents that occurred in Lisbon between 2004 and 2007. Each dot represents an accident. From the figure, it can be seen that crashes were concentrated at intersections and segments located on major arterial roads (Av. Gen. Norton Matos and Eixo Norte-Sul are the major expressways of Lisbon, with a speed limit of 80 km/h and Av República and Av. Almirante Reis are two of the major arterial roads, located right in the heart of the city).

Given the scope for this part of the study, only crashes that were classified as intersection-related were considered in the development of the statistical models. Unfortunately, there is no exact definition about the radius from the center of the intersection to classify a crash as intersection-related. Previous studies have used different criterion. For instance, Mountain et al. (1998) considered that accidents are intersection related if they occur at a distance of 20 m from the curb line. Sayed and Rodriguez (1999) considered a distance of 30 m from the intersection (without specifying the point from which it is measured); Lord (2000) used a distance equal to 15 m measured from the center of the intersection because the data were collected and categorized in this manner by the Toronto Police and Engineering Department and Turner et al. (2006) used a distance of 50 m without specifying the point from which it is measured. For this study, it was decided to use a radius of 40 m from the center of the intersection to classify crashes as intersection-related. Injury accidents involving pedestrians were removed from the dataset since this part of the project focused

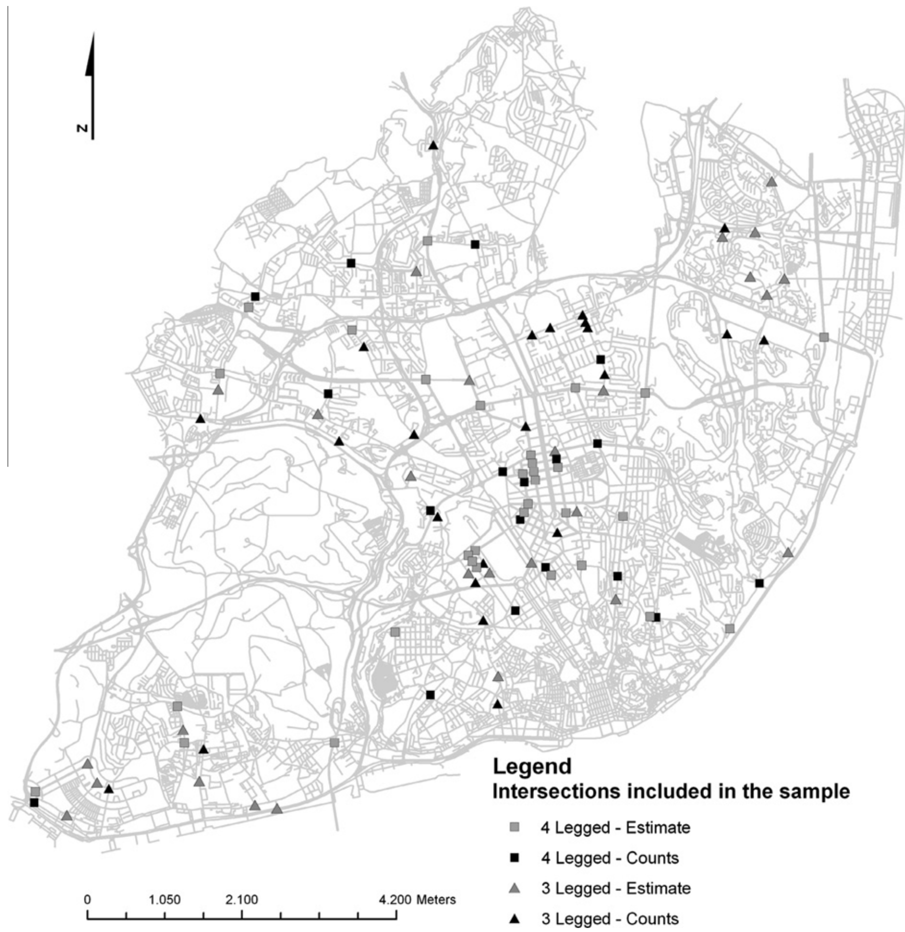

Fig. 1. Location of the intersections used in the study.

Fig. 2. Spatial distribution of injury accidents in Lisbon between 2004 and 2007.

Fig. 3. Road network used by Martinez in comparison with the total road network of Lisbon.

Table 1
Summary statistics of the crash dataset.

Type of intersection	Year	Accidents (min–max–total)	Major road flow (min–max)	Minor road flow (min–max)
Three-legged	2004	0–8–47	4963–77082	469–14166
	2005	0–5–31		
	2006	0–9–34		
	2007	0–5–35		
Four-legged	2004	0–11–94	5038–56066	299–31627
	2005	0–13–89		
	2006	0–13–92		
	2007	0–7–68		

on motor vehicle crashes only; pedestrian crashes will be analyzed separately in another part of the project.

Traffic flow counts at intersections were collected using three different sources:

- Automated: with a system that records traffic images and then processes it with a movement detection device.
- Manual: with several operators that counted the inflow on each leg.
- By estimates: Martinez (2006) developed a traffic assignment model for Lisbon, which covered about 55% of the road network length (see Fig. 3), all local roads were excluded from the model. Ten intersections of each type were included in the sample.

All the three sources for counting the traffic flow defined above had some limitations. The automated and manual counts were only done for a single day for each intersection. Since they were all collected in week days, it was assumed that the values obtained represented the Average Annual Daily Traffic (AADT). The changes in AADT over time were considered to be very small, and so, the same value of AADT was adopted for all four years. The output data obtained from the traffic assignment model estimated by Martinez (2006) was not fully adjusted to real traffic movements in some of the intersections included in the modeling sample. However, since the variables used relate to total inflow in each intersection, these values were still considered in the model development. Also, the estimated traffic data were supplied in Peak Hour Volume (PHV) from 8:00 to 9:00 am. From the manual and automated counts a conversion factor was calculated in order to transform the PHV into AADT (ratio between the number of vehicles from 8:00 to 9:00 over the AADT).

Table 1 summarizes important data characteristics for crash data occurring at three-legged and four-legged intersections. During the study period, a total of 137 crashes occurred at three-legged intersections, and 343 occurred at four-legged intersections.

Table 2 summarizes the key variable statistics of the intersections used in this study. The minimum and maximum AADT values ranged from about 400 to 77,000 vehicles per day. Among the variables collected, they included number of legs, number of lanes per leg, average lane width, median width, number of left turn lanes, number of right turn lanes and type of traffic control device (signalized or unsignalized).

4. Methodology

This section provides a brief description of the characteristics of the Poisson-gamma model used for estimating the crash prediction models.

4.1. Poisson-gamma model

The Poisson-gamma (or negative binomial or NB) model has the following modeling structure (Lord, 2006): the number of crashes 'Y_{it}' for site i and time period t when conditional on its mean μ_{it} is Poisson distributed and independent over all sites and time periods

$$Y_{it}|\mu_{it} \sim \text{Poisson}(\mu_{it}) \quad i = 1, 2, \ldots, I \text{ and } t = 1, 2, \ldots, T \tag{1}$$

The mean of the Poisson is structured as:

$$\mu_{it} = f(X; \beta)\exp(e_{it}) \tag{2}$$

where $f(\cdot)$ is a function of the covariates (X), β is a vector of unknown coefficients; and, e_{it} is the model error independent of all the covariates.

With this characteristic, it can be shown that Y_{it}, conditional on μ_{it} and α (i.e., the dispersion parameter of the Poisson-gamma distribution), is distributed as a Poisson-gamma random variable with a mean μ_{it} and a variance $\mu_{it} + \alpha\mu_{it}^2$ respectively. (Note: other variance functions exist for the Poisson-gamma model, but they are not covered here since they are seldom used in highway safety studies. The reader is referred to Cameron and Trivedi, 1998, and Maher and Summersgill, 1996 for a description of alternative variance functions). The probability density function (PDF) of the Poisson-gamma structure described above is given by the following equation:

$$f(y_{it}; \alpha, \mu_{it}) = \frac{\Gamma(y_{it} + \alpha^{-1})}{\Gamma(\alpha^{-1})y_{it}!}\left(\frac{\alpha^{-1}}{\mu_{it} + \alpha^{-1}}\right)^{\alpha^{-1}}\left(\frac{\mu_{it}}{\mu_{it} + \alpha^{-1}}\right)^{y_{it}} \tag{3}$$

where y_{it} is the response variable for observation i and time period t, μ_{it} the mean response for observation i and time period t and, α is the dispersion parameter of the Poisson-gamma distribution.

Note that if $\alpha \to 0$, the variance equals the mean and this model converges to the standard Poisson regression model.

The term α is usually defined as the "dispersion parameter" of the Poisson-gamma distribution or model (Note: that in some published documents, the variable α has also been defined as the "over-dispersion parameter"). This term has traditionally been assumed to be fixed and a unique value applied to the entire dataset in the study. As described above, the dispersion parameter plays an important role in safety analyses, including the computation of the weight factor for the Empirical Bayes method and the estimation of confidence intervals around the gamma mean and the predicted values of models applied to a different dataset than the ones employed in the estimation process.

4.2. Goodness-of-fit statistics

The method used to assess the goodness of fit was the Akaike Information Criterion (AIC). This measures of the goodness of fit of an estimated statistical model and is defined as (Akaike, 1974)

$$\text{AIC} = -2\log L + 2p \tag{4}$$

where L is the maximized value of the likelihood function for the estimated model, and p is the number of parameters in the statistical model. The AIC methodology attempts to find the model that best explains the data with a minimum of free parameters and thus it penalizes models with a large number of parameters. The model with the lowest AIC is considered to be the best model among all available models.

5. Analysis results

This section presents the estimation results of the models that were developed for three- and four-legged intersections. Flow-only models and models with covariates were developed for each of the intersection type.

5.1. Flow-only models

Flow-only models were developed for the three- and four-legged intersections. These models reflect the average conditions found in the data. They can be used for cases where the user has limited information about the geometric design features for the particular project under study. It should be pointed out that these models may be subjected to the omitted variables bias (Lord and Mannering, 2010).

Although different functional forms were examined, the functional form which provided the best fit for estimating flow-only models for intersections was the following (see Mountain and Fawaz (1996) and Miaou and Lord (2003), for a discussion about various functional forms for flow-only models):

$$\mu_{it} = \beta_0(F_{1it} + F_{2it})^{\beta_1}(F_{2it}/(F_{1it} + F_{2it}))^{\beta_2} \tag{5}$$

where μ_{it} is the estimated number of crashes for intersection i and year t, F_{1it} the entering traffic flows in vehicles per day (AADT) on the major approaches for intersection i and year t, F_{1it} the entering traffic flows in vehicles per day (AADT) on the minor approaches for intersection i and year t, β_0, β_1, β_2 is the estimated coefficients.

Table 3 summarizes the modeling results for the flow-only models for three- and four-legged intersections. This table shows that the ratio between F_2 and $F_1 + F_2$ was not found to be significant for four-legged intersections.

5.2. Models with covariates

For the model with covariates, the relationship between crashes and geometric design features is captured via the covariates inside the statistical model. The selection of the covariates to be included into the model can be governed by various statistical criteria, such as the statistical significance of each variable and AIC, as well as the statistical significance of the coefficients.

The functional form used for estimating the models with covariates for intersections is the following:

$$\mu_{it} = \beta_0(F_{1it} + F_{2it})^{\beta_1}(F_{2it}/(F_{1it} + F_{2it}))^{\beta_2} e^{\sum_{k=3}^{n}\beta_k x_k} \tag{6}$$

where μ_{it}, F_{1it} and F_{2it} are as defined above. x_k is the model covariates (e.g., right-turning lanes, median, etc.) and, $\beta_0, \ldots, \beta_k$ is the estimated coefficients.

In the initial step, a Poisson-gamma model was estimated with all the variables documented in Table 2 for three-legged intersections. Most of the variables were found to be insignificant or counterintuitive. This was attributed to small dataset (44 observations). In the end, only six variables were found to be significant at the 10% confidence level. Considering the relatively small sample size, we opted for a 10% level of significance. Previous studies have also used 10% level for this kind of dataset (e.g. Oh et al., 2006).

Table 4 presents the estimation results for three-legged intersections. The variables that were found to be significant were the

Table 2
Summary Statistics of the Dataset.

Variable	Description		Three-legged intersections				Four-legged intersections			
			Min.	Max.	Average (std. dev)	Frequency	Min.	Max.	Average (std. dev)	Frequency
F_1	AADT on major		4963	77082	19100 (13229)	44	5038	56066	19505 (12419)	50
F_1	AADT on minor		469	14166	4034 (3056)	44	299	31627	9839 (6739)	50
FT	$F_1 + F_1$		8793	78977	23134 (14368)	44	8104	80211	29344 (17864)	50
FR	F_1/FT		0.02	0.46	0.19 (0.12)	44	0035	0.50	0.33 (0.11)	50
FQ	F_1/F_1		0.02	0.84	0, 26 (0.21)	44	0037	0.98	0.53 (0.24)	50
LB	Lane balance	1 – yes	–	–	–	15 (34.09%)	–	–	–	12 (24.00%)
		0 – no	–	–	–	29 (65.91%)	–	–	–	38 (76.00%)
LMAJT7	Total number of entering lanes on major = 3 or more	1 – yes	–	–	–	26 (59.09%)	–	–	–	32 (64.00%)
		0 – no	–	–	–	18 (40.91%)	–	–	–	18 (36.00%)
LMINT7	Total number of entering lanes on minor = 2 or more	1 – yes	–	–	–	7 (15.91%)	–	–	–	27 (54.00%)
		0 – no	–	–	–	37 (84.09%)	–	–	–	23 (46.00%)
LWMAJ	Average lane width on major (m)	2.75	4.70	3.44 (0.53)	44	2.4	5.93	3.70 (0.78)	50	
LWMIN	Average lane width on minor (m)	2.35	7.03	3.71 (0.99)	44	2.25	5.84	3.71 (0.69)	50	
MMAJ1	Median presence on one leg of major direction	1 – yes	–	–	–	3 (6.82%)	–	–	–	–
		0 – no	–	–	–	41 (93.18%)	–	–	–	–
MMAJ2	Median presence on two legs of major direction	1 – yes	–	–	–	17 (38.64%)	–	–	–	35 (70.00%)
		0 – no	–	–	–	27 (61.36%)	–	–	–	15 (30.00%)
MMIN	Median presence on two legs of minor direction	1 – yes	–	–	–	17 (38.64%)	–	–	–	29 (59.18%)
		0 – no	–	–	–	27 (61.36%)	–	–	–	20 (40.82%)
LTPMAJ	Left turn presence on major	1 – yes	–	–	–	15 (34.09%)	–	–	–	9 (18.75%)
		0 – no	–	–	–	29 (65.91%)	–	–	–	39 (81.25%)
LTPMIN	Left turn presence on minor	1 – yes	–	–	–	8 (18.18%)	–	–	–	8 (17.02%)
		0 – no	–	–	–	36 (81.82%)	–	–	–	39 (82.98%)
RTPMAJ	Right turn presence on major	1 – yes	–	–	–	11 (25.00%)	–	–	–	12 (25.00%)
		0 – no	–	–	–	33 (75.00%)	–	–	–	36 (75.00%)
RTPMIN	Right turn presence on minor	1 – yes	–	–	–	13 (29.55%)	–	–	–	15 (32.61%)
		0 – no	–	–	–	31 (70.45%)	–	–	–	31 (67.39%)
TCD	Traffic control device	1 – signals	–	–	–	23 (52.27%)	–	–	–	35 (70.00%)
		0 – all others	–	–	–	21 (47.73%)	–	–	–	15 (30.00%)
LOW	Number of one way legs		–	–	–	–	0	1	0.10 (0.30)	50

total traffic flow entering the intersection, the ratio between the traffic flow entering on the minor direction and the total traffic flow entering the intersection, lane balance (note: lane balance means that the major and minor approaches entering the intersection have the same number of lanes), the median presence on one leg on major direction, the presence median on two legs for the major direction, the average median width on the minor direction and the presence of a right turn lane in the major direction.

Similar to the three-legged intersections analysis, a Poisson-gamma model was developed with all the variables documented in Table 2 for the four-legged intersections. Only six variables were found to be significant at the 10% confidence level. Table 5 presents the estimation results for four-legged intersections. The significant variables were the total traffic flow in the intersection, the lane balance, the total number of lanes on major direction higher than three, the average lane width on minor direction, the presence of a right turn lane in the major direction and the number of one way legs in the intersection.

Although it is usually believed that the use of signals as traffic control devices is safer, this variable was not significant for this dataset. The same happened with the variable for the presence of a left turn (on major or minor direction). The fact that these variables are not statistically significant could be attributed to the small sample size.

The coefficients of the models were compared with those produced from models that were documented in the literature (see, e.g., Greibe, 2003 and Bauer and Harwood, 2000). Very few variables could be compared, since most have not been examined by other researchers. Among those that were similar, all the coefficients had the same sign, but the magnitude was different (which is expected given that data are different). For instance, the coefficient for the variable "average lane width on minor direction" in the Portuguese four-legged intersection model was 0.5256. Greibe (2003) had a similar variable, desegregated by classes and obtained the following values for the coefficients: 5.0 to 7.5 m – 0.83; 8.0 to 8.5 m – 0.68, and 9.0 to 15.0 m – 0.80. A similar comparison was possible with the variable "right turn lane presence on major direction." The Portuguese coefficient was 0.7350, whereas Bauer and Harwood's (2000) coefficient was equal to 0.209 for the model developed for stop controlled intersections.

6. Summary and conclusions

The purpose of this study was to develop crash prediction models for urban intersections located in Lisbon, Portugal which would describe the expected number of accidents as a function of a range of explanatory variables, namely traffic flow counts and highway geometric design features. Flow-only models and models with covariates were estimated using data collected at 44 three-legged intersections and 50 four-legged intersections. The coefficients were estimated using the GLM modeling approaches with the Pois-

Table 3
Estimates for three and four-legged intersections.

		Ln (β_0)	$F_1 + F_1$ (β_1)	$F_1/(F_1 + F_1)$ (β_2)	α	AIC
Three-legged intersections	Estimates	−12.5848 (3.014)[a]	1.1486 (0.305)	−0.3551 (0.208)	0.7069 (0.544)	188.66
	t-value	−4.18	3.76	1.71	–	
Four-legged intersections	Estimates	−8.6325 (2.096)	0.8909 (0.205)	[b]	0.5188 (0.530)	286.27
	t-value	−4.12	4.35	[b]	–	

[a] Standard error.
[b] Insignificant.

Table 4
Estimates for three-legged intersections.

Variable	Estimates	*t*-value
Ln (β_0)	−14.2276 (2.3474)[a]	−6.06
$F_1 + F_1$ (β_1)	1.2635 (0.2257)	5.60
$F_1/(F_1 + F_1)$ (β_2)	−0.6980 (0.1787)	−3.91
LB (β_3)	−0.6462 (0.3468)	−1.86
MMAJ1 (β_4)	−1.1501 (0.4489)	−2.56
MMAJ2 (β_5)	−1.2357 (0.3163)	−3.91
MMIN (β_6)	0.4837 (0.2358)	2.05
RTPMAJ (β_7)	0.7350 (0.2362)	3.11
α	0.104 (10.1)	–
AIC	176.03	–

[a] Standard error.

Table 5
Estimates for four-legged intersections.

Variable	Estimates	*t*-value
Ln (β_0)	−6.9152 (2.6285)[a]	−2.63
$F_1 + F_1$ (β_1)	0.4288 (0.2645)	1.62
LB (β_2)	0.6672 (0.2524)	2.64
LMAJT7 (β_3)	0.9871 (0.3122)	3.16
LWMIN (β_4)	0.5256 (0.1644)	3.20
RTPMAJ (β_5)	0.5850 (0.2862)	2.04
LOW (β_6)	−0.7687 (0.3745)	2.05
α	0.2932 (3.41)	–
AIC	274.36	–

[a] Standard error.

son-gamma distribution. The total traffic inflow was found to be highly significant both for three- and four-legged intersections, as expected. Some highway design geometric variables influenced safety. For three-legged intersections, four variables were found to have a positive effect on the safety of three-legged intersections:

- the ratio between the traffic flow entering on minor direction and the total traffic flow entering the intersection,
- the lane balance, which can be explained by the fact that intersections with the same number of entering lanes on all the legs lead to linear movements and less conflicts
- the presence of a median on one leg on major direction or the presence of a median on two legs on major direction, meaning that the presence of a median, either on one or two legs of major direction is associated with less crashes

Although it was not expected, a negative influence on safety was found with the presence of a median on the minor direction and with the presence of a right turn lane on the major direction. This can be partly explained by endogenous effects, meaning that they are not unsafe solutions, but they were introduced because the sites were unsafe.

For four-legged intersections, only the number of one-way legs in the intersection was found to reduce accident occurrences, which is related with the lower number of conflicts that characterize one-way streets. The variables associated with an increase in the number of accidents were:

- Lane balance on four-legged intersections had the opposite effect than the one observed in three-legged intersections. It was found that the intersections where the variable had value “1” (5 in a total of 12), the observed number of accidents was very high, being most likely associated with large scale of these intersections, since in these cases, the variable LMAJT7 was also found to be equal to “1”. This is a possible reason that led to a positive coefficient for the LB variable.
- Total number of entering lanes on major equal to three or more: usually as the number of entering lanes increases, traffic volume also increases, meaning more conflicts in the intersection.
- Average lane width on minor: a larger width allows a higher variance on lateral positioning that might be related with lateral collisions.
- The presence of a median on two legs on major direction: partly justified by endogenous effects (see explanation for the three legged model).

It is recognized that a statistically significant association between crash frequency and explanatory variables does not necessarily explain a cause-effect relationship. However, this study can be considered a good starting point about gaining knowledge for better understanding the relationships between crashes and roadway characteristics in Lisbon. This study was part of an overall project lead by the LNEC, which included other predictive models for estimating the safety performance of roadway links and roundabouts as well as for estimating pedestrian collisions.

References

Akaike, 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control AC-19, 716–723.
ANSR, 2007. Autoridade Nacional de Segurança Rodoviária. Ano de 2007 – Sinistralidade rodoviária. Observatório de Segurança Rodoviária.
Azeredo Lopes, S., Cardoso, J.L., 2007. Accident Prediction Models for Portuguese Motorways. ICT Informação Científica, LNEC, Lisbon.
Azeredo Lopes, S., Cardoso, J.L., 2007a. Accident Prediction Models for Portuguese Single Carriageways Roads. ICT Informação Científica, LNEC, Lisbon.
Azeredo Lopes, S., Cardoso, J.L., 2007b. Accident Prediction Models for Bidirectional Data on Portuguese Motorways. ICT Informação Científica, LNEC, Lisbon.
Bauer, K., Harwood, D.W., 2000. Statistical models for at-grade intersection accidents – Addendum. Relatório FHWA-RD-99-094, US Department of Transportation, Fedearl Highway Administration.
Brüde, U., Larson, J., 1993. Models for predicting accidents at junctions where pedestrians and cyclist are involved. How well do they fit? Accident Analysis and Prevention 25 (5), 499–509.
Cameron, A.C., Trivedi, P.K., 1998. Regression Analysis of Count Data, Econometric Society Monograph. Cambridge University Press, New York, NY (No. 30).
DGV, 2004. Direcção Geral de Viação. Sinistralidade rodoviária 2004 – Elementos Estatísticos. Ministério da Administração Interna. Observatório de Segurança Rodoviária.
DGV, 2005. Direcção Geral de Viação. Sinistralidade rodoviária 2005 – Elementos Estatísticos. Ministério da Administração Interna. Observatório de Segurança Rodoviária.
DGV, 2006. Direcção Geral de Viação. Sinistralidade rodoviária 2006 – Elementos Estatísticos. Ministério da Administração Interna. Observatório de Segurança Rodoviária.
Greibe, P., 2003. Accident prediction models for urban roads. In: Accident Analysis and Prevention, vol 35. pp. 273–285.

Hall, R.D., 1986. Accidents at four-arm single carriageway urban traffic signals. Contractor Report 65, Transportation and Road Research Laboratory, United Kingdom.
Lord, D., 2000. The prediction of accidents on digital networks: characteristics and issues related to the application of accident prediction models. PhD Thesis, Toronto University.
Lord, D., Washington, S.P., Ivan, J.N. Poisson, 2005. Poisson-gamma and zero inflated regression models of motor vehicle crashes: balancing statistical fit and theory. Accident Analysis and Prevention 37 (1), 35–46.
Lord, D., 2006. Modeling motor vehicle crashes using Poisson-gamma models: examining the effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter. Accident Analysis and Prevention 38 (4), 751–766.
Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives. Transportation Research – Part A 44 (5), 291–305.
Maher, M.J., Summersgill, I., 1996. A comprehensive methodology for the fitting predictive accident models. Accident Analysis and Prevention 28 (3), 281–296.
Martinez, L., 2006. TAZ Delineation and Information Loss in Transportation Planning Studies. MSc Thesis, IST, Portugal.
Miaou, S.-P., Lord, D., 2003. Modeling traffic-flow relationship at signalized intersections: Dispersion paramet, function Form and bayes vs empirical bayes. Transportation Research Record 1840, 31–40.
Mountain, L., Fawaz, B., 1996. Estimating accidents at junctions using routinely-available input data. Traffic Engineering and Control 37 (11), 624–628.
Mountain, L., Maher, M., Fawaz, B., 1998. The influence of trend on estimates of accidents at junctions. In Accident Analysis and Prevention 30 (5), 641–649.
Oh, J., Washington, S.P., Nam, D., 2006. Accident prediction model for railway-highway interfaces. Accident Analysis and Prevention 38 (2), 346–356.
Reurings, M., Janssen, T. 2006. Accident prediction models for urban and rural carriageways. R-2006-14, SWOV.
Sayed, T., Rodriguez, F., 1999. Accident prediction models for urban unsignalized intersections in British Columbia. In: Transportation Research Record: Journal of the Transportation Research Board, No. 1665. pp. 93–99.
Turner, S.A., Roozenburg, A.P., Francis, T., 2006. Predicting Accident Rates for Cyclists and Pedestrians. Land Transport New Zealand Research Report 289, Christchurch, New Zealand.