DOI: 10.2478/ace-2014-0002

ESTIMATING THE SAFETY PERFORMANCE FUNCTION FOR URBAN UNSIGNALIZED FOUR-LEGGED ONE-WAY INTERSECTIONS IN PALERMO, ITALY

O. GIUFFRÈ[1], A. GRANÀ[2], T. GIUFFRÈ[3], R. MARINO[4], S. MARINO[5]

Starting from consideration that urban intersections are sites with promise for safety and operational improvements, the paper describes the steps taken to develop a crash predictive model for estimating the safety performance of urban unsignalized intersections located in Palermo, Italy. The focus is on unsignalized four-legged one-way intersections widespread in Italian downtowns. The sample considered in the study consist of 92 intersections in Palermo, Italy. For the study were collected crashes occurred in the sites during the years 2006-2012, geometric design and functional characteristics and traffic flow. Results showed that data were overdispersed and NB1 distributed. In order to account for the correlation within responses Generalized Estimating Equations (GEE) were used under different working correlation matrices.

Keywords: safety performance functions, urban intersection, overdispersion, correlation

1. Introduction

Italian official statistics report a high number of crashes occurring yearly on the national road network. For instance, in 2011 in Italy there were 205,638 injury crashes whereas the number of deaths was 3,860 and that of injured was 292,019 (Istat [1]). Compared to 2010, there was a decrease in the number of total crashes (-2.7%) and injury crashes (-3.5 %). A more substantial reduction was observed in the number of deaths (-5.6 %); from 2001 to 2011, the reduction in the number of deaths amounted to 45.6 %.

[1] Department of Civil, Environmental, Aerospace, Materials Engineering, Università di Palermo, Viale delle Scienze Ed.8, Palermo, 90128, Italy, E-mail: orazio.giuffre@unipa.it

[2] Department of Civil, Environmental, Aerospace, Materials Engineering, Università di Palermo, Viale delle Scienze Ed. 8, Palermo, 90128, Italy, E-mail: anna.grana@unipa.it; (corresponding author)

[3] Faculty of Engineering and Architecture, Università Kore di Enna, Cittadella Universitaria, 26, Enna 94100, Italy, E-mail: tullio.giuffre@unikore.it

[4] Department of Energy, Information Engineering and Mathematical Models, Università di Palermo, Viale delle Scienze Ed. 8, Palermo, 90128, Italy, E-mail: roberta-marino@libero.it

[5] Department of Civil, Environmental, Aerospace, Materials Engineering, Università di Palermo, Viale delle Scienze Ed. 8, Palermo, 90128, Italy, E-mail: sergio.marino@unipa.it

The target set by the European Union in the 2001 White Paper provided for a reduction of 50% of road deaths by 2010. Although Italy has approached the hoped reduction, that target has not yet reached; however, the reduction of 45.6% recorded in Italy is higher than the European mean value equal to 44.5%.

This important safety problem has also been observed on the urban road network located in and around Italian cities and towns. According to official statistics (Istat [1]), a large percentage of crashes occur in urban areas; on Italian roads inside urban agglomerations, indeed, in 2011 there were 157,023 crashes (76% of all crashes in Italy), leading to 213,001 injuries (73%) and 1,744 deaths (45%), whereas crashes on freeways accounted for 5% of the total (11,007), with 18,515 injuries (6%) and 338 deaths (9%). On the other rural roads, except the freeways, there were 37,608 crashes, 65,503 injuries and 1,778 deaths. However crashes happening on urban roads are less severe, with the mortality index equal to 1.1 deaths/100 crashes, while this index is equal to 3.1 deaths/100 crashes and 4.7 deaths/100 crashes on freeways and the other rural roads, respectively.

Urban crashes are concentrated at road junctions and intersections where the potential for vehicle–vehicle and vehicle–pedestrian conflicts is high. Italian official statistics inform that in 2011 69.7% of deaths are drivers, 15.3% are passengers and 15.1 % pedestrians (Istat [1]). Moreover, in the context of improper driving behaviors, failure to comply with the give-way sign and/or the traffic signals at urban intersections is the first cause of crashes (19.6%); this doesn't happen in rural roads where distraction and high speed are prevalent, with percentages of 19,7% and 18,1%, respectively. It should be said, however, that some forms of intersection control are more effective than others in reducing conflicts (Ewing and Dumbaugh [2]).

Anyway, in many cities and towns of the national territory of Italy, a large number of intersections are considered sites with promise for safety and operational improvements. Several studies have been carried out in many countries to establish relationships between crashes and flow and non-flow explanatory variables, using statistical tools to investigate factors critical to road safety; see Poch and Mannering [3], Lord and Persaud [4], Oh at al. [5], Abdel-Aty et al. [6].

This paper documents one component of the safety tools that are currently under development (HSM [7]). Indeed the paper describes the steps taken to develop a crash predictive model for estimating the safety performance of urban unsignalized intersections located in Palermo, Italy. The focus is on four-legged one-way unsignalized intersections widespread in Italian downtowns. It has to be noted that there has been a significant amount of researches done on the development and application of safety performance functions for various types of road safety analyses, with reference to roadway segments and (signalized or unsignalized) intersections (Rodegerdts et al. [8], Turner et al. [9], Cafiso and D'Agostino [10]. However, the number of studies dedicated to the type of unsignalized intersections here examined is to date rather limited, since these intersections are more common in European cities than other places around the world, such as in North America.

The methodological path followed in this research allowed to handle issues associated with the estimation of a safety performance function (SPF) for this particular kind of unsignalized intersections. Several crash predictive models were developed using the NB1 and NB2 modeling framework. Both flow-only models and models also including non-flow covariates were estimated using crashes and other related data collected at a sample of 92 four-legged intersections for the years 2006–2012.

At last, results supply methodological insights that may be useful in the subsequent quantifying of benefits obtainable by engineering measures aimed at enhancing traffic safety in built up areas.

Next section will summarize the description of data here analyzed; then the methodology for estimating the model with flow and non-flow variables; at last results will be presented and discussed.

2. Data description

The examined sample consists of four-legged unsignalized intersections, each of them with one-way approaches. The data collected included geometric design and functional characteristics, crash data and traffic volumes. Except for the crashes all the data were obtained from on-site visits. In spite of the costs associated with the data collection process, the sample was limited to 92 sites.

Fig. 1. Sample location in the Palermo road network

Since seven years of crash data were considered as distinct observations, there were in total 644 observations over the period 2006–2012 for the selected four-legged intersections; crash observations were considered as repeated measurements.

Crash data were collected from the official crash statistics database available at the Municipal Police Force in Palermo and police reports. Given the scope of this study, only crashes classified as intersection-related were considered in the development of the statistical models. Studies reported in literature are not unanimous in giving the exact definition of crashes intersection-related and have used different criterion (e.g. Vieira Gomes et al. [11]). Thus, for this study it was decided to use a radius of 20 meters from the center of the intersection to classify crashes as intersection-related.

Figure 1 shows the graphical representation of the Palermo road network and the location of the intersections used in this study.

Traffic data surveys were carried out during 2012. Traffic flow counts at intersections were collected manually, with several operators that counted the inflow on each leg. Since traffic flow counts were all collected in weekdays, it was assumed that the values obtained represented the Average Annual Daily Traffic (AADT). The changes in major-road AADT and minor-road AADT over time were estimated for each intersection and for each year in accordance with the growth of vehicle registrations from 2006 to 2011 (Automobile Club of Italy [12]).

Table 1 summarizes important data characteristics for crash data occurring at four-legged intersections.

Table 1

Summary statistics of crash dataset

	Crashes				AADT on major road		AADT on minor road	
	min	mean	max	total	min	max	min	max
2006	0	0.92	7	85	6,400	30,430	3,800	18,410
2007	0	0.95	8	87	6,496	30,886	3,857	18,686
2008	0	0.98	8	90	6,593	31,350	3,915	18,966
2009	0	1.00	12	92	6,692	31,820	3,974	19,251
2010	0	1.09	15	100	6,793	32,297	4,033	19,540
2011	0	1.08	14	99	6,895	32,782	4,094	19,833
2012	0	1.02	13	94	6,998	33,273	4,155	20,130
2006-2012	0	1.00	15	647	6,400	33,273	3,800	20,130

Note: AADT values are in veh/d

Table 2 summarizes the key variable statistics of the intersections used in this study. The minimum and maximum AADT values on major road ranged from about 6,400 to little more than 33,000 vehicles per day whereas AADT values on minor road ranged from about 3,800 to little more than 20,000 vehicles per day. Among the variables collected, number of entering lanes per leg, average lane width, parking on road sides and visibility conditions on intersection approaches were included.

Table 2

Summary statistics of the dataset

Variable	Description		min	max	mean	std. dev.	frequency
F_1	AADT on major [10^3 veh/d]		6.40	33.27	16.05	4.06	92
F_2	AADT on minor [10^3 veh/d]		3.80	20.13	11.17	3.34	92
F_T	F_1+F_2 [10^3 veh/d]		10.20	53.40	27.22	6.50	92
F_R	F_2/F_T [10^3 veh/d]		0.25	0.50	0.41	0.06	92
F_Q	F_2/F_1 [10^3 veh/d]		0.34	0.99	0.71	0.18	92
NL_1	number of entering lanes on major > 1	1 – yes	-	-	-	-	13.04 %
		0 – no	-	-	-	-	86.96 %
NL_2	number of lanes on minor >1	1 – yes	-	-	-	-	1.09 %
		0 – no	-	-	-	-	98.91%
LW_1	average lane width on major [m]		3.50	10.50	4.89	1.39	92
LW_2	average lane width on minor [m]		3.00	7.00	4.05	0.74	92
P_1	parking on both sides of major road	1 – yes	-	-	-	-	91.30 %
		0 – no	-	-	-	-	8.70 %
P_2	parking on both sides of minor road	1 – yes	-	-	-	-	91.30 %
		0 – no	-	-	-	-	8.70 %
V_1	good visibility on major road approach	1 – yes	-	-	-	-	35.87 %
		0 – no	-	-	-	-	64.13 %
V_2	good visibility on minor road approach	1 – yes	-	-	-	-	38.04 %
		0 – no	-	-	-	-	61.96 %

3. Model selection

The first step to develop a Safety Performance Function (SPF) is to select which explanatory variables should be used and to set the model form; these tasks were discussed in previous papers (Giuffrè et. al, [13]; Giuffrè et. al, [14]; Giuffrè at al [15]). Covariates explored in the current study are listed in Table 3; they were selected looking at safety performance function for urban unsignalized intersections referred in literature over last 10 years (e.g. Bauer and Harwood, [16]; McGee, Taori, Persaud [17]) and considering the statistical significance of each variable.

First column in Table 3 reports all the variables collected for each site and considered in the study; in the third column are marked the variables that were found to be significant at the 15% confidence level; in the right column the variables included in the model specification.

Table 3

Variables explored and selected

Variables	Abbreviation	Significant variables	Selected variables
Annual Average Daily Traffic on major-road	F_1	✓	
Annual Average Daily Traffic on minor-road	F_2		
Sum of Annual Average Daily Traffic on major and minor-road (F_T)	$F_1 + F_2$	✓	✓
Ratio between Annual Average Daily Traffic on minor-road and Total Annual Average Daily Traffic entering in intersection	$F_2 / (F_T)$	✓	
Ratio between Annual Average Daily Traffic on minor and major-road	F_2 / F_1	✓	
Major-road number of lanes	NL_1	✓	✓
Minor-road number of lanes	NL_2		
Major-road lane width	LW_1	✓	
Minor-road lane width	LW_2		
Parking on major-roadside (0 if it is permitted only on one side, 1 otherwise)	P_1		
Parking on minor-roadside (0 if it is permitted only on one side, 1 otherwise)	P_2		
Visibility on major-road approaching intersection (0 if it is insufficient, 1 otherwise)	V_1		
Visibility on minor-road approaching intersection (0 if it is insufficient, 1 otherwise)	V_2		
✓ *significant at the 15% confidence level*			

Different model forms (see Table 4) were investigated considering the combinations of all the variables listed in Table 3. The results of this exploratory analysis revealed that the best functional relationship between crashes and the significant covariates was the *power function* for the variable $F_1 + F_2$ and the exponential function for the variable NL_1.

Table 4

Model forms investigated

Name	Model form
Power function	$y = \beta_0\mathbf{X}^{\beta}$
Exponential function	$y = \beta_0 e^{\beta\mathbf{X}}$
Gamma function	$y = \beta_0\mathbf{X}^{\beta}e^{\beta\mathbf{X}}$

It has to be noted that F_1, F_2/F_T and F_2/F_1 were excluded from the model for two reasons: i) they were partially accounted in the variable $F_1 + F_2$; ii) their introduction in the model either in the power or in the exponential function produced no appreciable benefits on the model performance.

Then the final selected model had the form:

$$y_{ij} = \beta_0 (F_1 + F_2)_{ij}^{\beta_1} e^{\beta_2 NL_{1j}} \tag{3.1}$$

where:

y_{ij} = expected number of crashes for the year *i* and the intersection *j*;

$(F_1+F_2)_{ij}$ = sum of Annual Average Daily Traffic on major and minor-road for the year *i* and the intersection *j*;

NL_{1j} = major-road number of lanes at the intersection *j*

β_0, β_1, β_2, = parameters to be estimated.

4. Development of safety performance function

The relationship between crash frequency and traffic/geometric variables for roadway segments and intersections has been the subject of study for many years. A wide number of research efforts have examined this relationship with the purpose of determining the effect of road and intersection design on the frequency of crashes. Technique of Generalized Linear Models (GLMs) has been recognized able to offer a soundly-based approach for analyzing this kind of data and fitting predictive crash models. Due to the nature of crashes occurrence, the assumption of a Poisson distribution for the crash frequency in a given time period at any one site has proven to be a good choice to model the process (Maher, Summersgill, [18]). Assuming the Poisson model, the functional forms of relationships can be estimated using the technique of Generalized Linear Models (GLMs) (McCullagh, Nelder [19]). However, crash data characteristics and methodological-technical issues may impair the efficient use of the Poisson model, which thus could produce considerable bias in parameters estimates and possible erroneous inferences (Lord, Mannering [20]; Poch, Mannering, [3]). The use of Poisson assumes that the mean and the variance of the distribution are equal, but this assumption is often too much restrictive for crash data. Evidence suggests that crash data counts may be overdispersed (the variance exceeds the mean of the crash counts on road entities), otherwise in few cases the data may be underdispersed (the mean is greater than the variance). The Poisson model, indeed, cannot take account of overdispersion (and underdispersion); in order to relax the Poisson assumption of equidispersion, quasi-likelihood methods represent a potential solution. Several authors have addressed the overdispersion issue by using the Negative Binominal regression model; see e.g. Poch, Mannering, [3] Miaou [21]. Properties of the traditional NB models have been illustrated by Cameron and Trivedi [22].

Because crash counts often consist of observations over several time periods it is also necessary to take into account the question of temporal correlation in the data. However considering correlation in the data, the likelihood function becomes very complicated to solve (Cameron and Trivedi [22]; Cafiso and D'Agostino [10]), but Generalized Estimating Equations (GEEs) overcome this problem (Liang and Zeger [23]). The failure of the independence hypothesis for the response variate is a serious issue in safety modeling, that is why elusion of the correlation within responses can lead to misleading conclusions in model interpretation on the basis of incorrect estimates of the variances and of an inefficient or biased estimate of the regression coefficients (Giuffrè et. al [13]; Diggle et al. 2002 [24]).

Starting from these considerations, the purpose of the study was to calibrate a Safety Performance Function for urban four-legged one-way intersections and to improve parameters estimates efficiency taking into account either dispersion and temporal correlation in the data. First in this section it is addressed the problem of dispersion in the data using quasi-likelihood methods in GLM context. In order to select the best model different goodness-of-fit methods have been used to evaluate predictive performance of models and to find the model that best explains the data among all estimated models. Second it is addressed the problem of temporal correlation in the data; therefore the model selected in GLM context has been recalibrated using Generalized Estimating Equations (GEEs) under different correlation structures of the data; again different goodness-of-fit methods have been used to evaluate predictive performance of the model.

4.1. Goodnes of fit

Technical literature suggests different goodness-of-fit methods to evaluate predictive performance of models and to find the model that best explains the data among all estimated models. The methods used in this paper include the following (where the subscript "i" denotes the generic observation at year t and at site j):

Mean Prediction Bias (MPB)

MPB gives a measure of the magnitude and direction of the average model bias (Oh et al. [5]). If the MPB is positive then the model over-predicts crashes and if the MPB is negative then the model under-predicts crashes. It is computed using the following equation:

(4.1)
$$\mathrm{MPB} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)$$

where N is the sample size, $\hat{y}_i$ and y_i are the predicted and observed crashes at site i respectively.

Mean Absolute Deviance (MAD)

MAD gives a measure of the average mis-prediction of the model (Oh et al. [5]). The model that provides MAD closer to zero is considered to be the best among all the available models. It is computed using the following equation:

(4.2)
$$\mathrm{MAD} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{y}_i - y_i\right|$$

Mean Squared Predictive Error (MSPE)

MSPE is typically used to assess the error associated with a validation or external data set (Oh et al. [5]). The model that provides MSPE closer to zero is considered to be the best among all the available models. It can be computed using the following equation:

(4.3)
$$\mathrm{MSPE} = \frac{1}{N}\sum_{i=1}^{N}\left(\hat{y}_i - y_i\right)^2$$

Akaike Information Criterion (AIC)

The AIC (Akaike [25]) is a measure of the goodness-of-fit of an estimated statistical model and is defined as:

(4.4)
$$\mathrm{AIC} = -2\log L + 2p$$

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

Quasilikelihood under the Independence model Criterion (QIC)
As above referred, since the GEE method is a quasi-likelihood based method, an extension of the Akaike's information criterion is needed to compare covariance matrices under GEE models to the covariance matrix generated under the independence hypothesis. So AIC statistic is replaced by the QIC statistic, defined as (Pan [26]):

$$\text{QIC} = -2\,Q + 2\,p \tag{4.5}$$

where Q is the quasi-likelihood function ($Q = L/\phi$) and p is the number of parameters in the statistical model. The model with the lowest QIC is considered to be the best model among all available models.

4.2. Development of SPF considering dispersion in the data

In order to relax the Poisson assumption of equidispersion model regression coefficients and the associate standard errors were estimated in GLM context assuming Negative Binomial distributions (*NB1* and *NB2*); in both cases *GenStat software* was used. These distributions are an effective tool to address problem of dispersion and to increase estimates efficiency (Lord [27]); these models take into account for overdispersion by means of a parameter called *overdispersion parameter* α (with $\alpha > 1$); wider considerations are reported in the cited papers Giuffrè et. al [15]; Lord [27]).

Table 5 shows coefficient estimates and goodness-of-fit for the model selected.

Table 5

Coefficients Estimates and Goodness-of-Fit in GLM context

Model form $y_{ij} = \beta_0 (F_1 + F_2)_{ij}^{\beta_1} e^{\beta_2 NL_{1j}}$

	NB1			NB2		
Variables	est	s.e.	t	est	s.e.	t
Constant (β_0)	-3.46	0.873	-3.96	-2.1	0.898	-2.34
F_1+F_2 (β_1)	0.978	0.266	3.68	0.562	0.276	2.04
NL_1 (β_2)	0.935	0.137	6.82	1.04	0.167	6.25
α	1.72	0.443	3.88	1.07	0.180	5.94
MPB	0.00			-0.007		
MAD	0.9931			1.00		
MSPE	2.26			2.31		
AIC	1.077			1.764		

From results showed in Table 5 it can be seen that the NB1 model fits the data better than the NB2 one as highlighted by standard errors that are lower for NB1 model. Also the goodness-of-fit indicators show that the NB1 model performs data better than the NB2:

– MPB values for both the model are almost the same and next to zero, denoting that the models have good prediction accuracy:
– there are slight differences in MAD and MSPE values for the two models; values closer to zero of both indicators for NB1 model show that the latter performs data better than NB2;
– the lowest AIC value clearly indicates that the NB1 model has to be considered the best between the two models.

4.3. Development of SPF considering temporal correlation in the data

In order to account for the correlation within responses NB1 model was fitted again in GEE context considering different working correlation matrices, that is, assuming that repeated observations were correlated in different ways. Again GenStat software was used for this purpose. Four forms of correlation were explored starting from the simplest one (*independence* structure) for which observations are thought (unrealistically) to be uncorrelated. In contrast to the hypothesis of independence was also assumed the *unstructured* structure to allow the free estimates on the within-site correlation from the data, the *exchangeable* structure that supposes no logical ordering for within-entity observations and a correlation structure of a stationary *(n-1)-dependent* process. The GEE regression results under the four named working correlation matrices are summarized in Table 6, in which are shown also the goodness of fit indicators.

The results in Table 6 show that unstructured working correlation matrix fits the data better than the other structures. MPB, MAD and MSPE values are slightly different, however they show a better performance of the model with unstructured correlation. Pan statistic, conversely, allows the best correlation structure to be determined clearly, in fact it can easily be seen that the best performance is supplied by assuming the unstructured hypothesis.

5. Conclusion

The paper describes methods applied to develop a SPF for urban unsignalized four-legged one-way intersections. The sample considered in the study consist of 92 intersections in Palermo, Italy. For the study were collected crashes occurred in the sites during the years 2006-2012, geometric design and functional characteristics and traffic flow. The fist step in developing the SPFs involved the selection of explanatory variables to be used and how variables could enter into the model to choice the best model form. Two covariates were selected: $F_1 + F_2$ and NL_1. With regards to the functional model

Table 6

Coefficients Estimates and Goodness-of-Fit in GEE context

	independence			unstructured			exchangeable			6-dependence		
variables	est	s.e.	t	est	s.e.	t	est	s.e.	t	est	s.e.	t
Constant (β_0)	-3.46	2.347	-1.474	-5.145	1.774	-2.900	-3.564	2.285	-1.560	-3.795	2.270	-1.672
$AADT_1$+$AADT_2$ (β_1)	0.978	0.704	1.389	1.491	0.534	2.792	1.009	0.684	1.475	1.079	0.682	1.582
NL_1 (β_2)	0.935	0.275	3.400	0.737	0.223	3.305	0.931	0.242	3.847	0.909	0.249	3.651
α	2.087			2.3[(*)]			2.083			2.09		
MPB	0.015			0.023			0.017			0.018		
MAD	0.99			0.98			0.99			0.99		
MSPE	2.26			2.22			2.26			2.25		
QIC	519			472			520			518		

[(*)] *the unstructured working correlation matrix allows a dispersion parameter varying over time in the observation period, the value reported is the mean value of those obtained* form, the power function seemed appropriate for the covariate $F_1 + F_2$ and the exponential function for covariate NL_1. In order to relax the Poisson assumption of equidispersion Negative Binomial models (NB1 and NB2) were implemented in GLM context. Results showed that data were overdispersed and that the NB1 model performed better than the NB2. In order to account for the correlation within responses NB1 model was fitted again in GEE context considering different working correlation matrices (independence, unstructured, exchangeable and 6-dependence). Results showed that the unstructured working correlation matrix fitted the data better than the other structures; Pan statistic allowed to choose the best correlation structure for the data better than the other goodness of fit indicators used.

REFERENCES

1. ISTAT, Italian National Institute of Statistics, Statistics on Highway accidents, 2012.
2. R. EWING, E. DUMBAUGH, *The built environment and traffic safety: A review of empirical evidence*, Journal of Planning Literature 23, 347-367, 2009.
3. M. POCH, F. MANNERING, *Negative Binomial Analysis of intersection-accident frequencies*, ASCE Journal of Transportation Engineering, 122, 2, 105-113, 1996.
4. D. LORD, B. N. PERSAUD, *Accident Prediction Models With and Without Trend: Application of the Generalized Estimating Equations Procedure*, Transportation Research Record, 1717, 102-108, 2000.
5. J. OH, C. LYON, S. P. WASHINGTON, B. N. PERSAUD, J. BARED, *Validation of the FHWA Crash Models for Rural Intersections: Lessons Learned*, Transportation Research Record, 1840, 1, 41-49, 2003.
6. M. ABDEL-ATY, H. SALKAPURAM, C. LEE, AND P. A. BRADY, A Simplistic, Practical Approach To Identify Traffic Crash Profiles at Signalized Intersections, ITE Journal, 76, 4, 33-28, 2006.
7. AASHTO, *Highway Safety Manual, 1st Edition*, ISBN 978-1-56051-477-0, Washington DC, 2010.
8. L. RODEGERDTS, M. BLOGG, E. WEMPLE, E. MYERS, M. KYTE, M. DIXON, G. LIST, A. FLANNERY, R. TROUTBECK, W. BRILON, N. WU, B. N. PERSAUD, C. LYON, D. HARKEY, *Application of roundabouts in the US*, NCHRP Report 572: Roundabouts in the US, TRB, Washington, DC, 2007.
9. S. TURNER, B. N. PERSAUD, M. CHOU, C. LYON, A. *Roozenburg, International crash experience comparisons using prediction models*, Proc. of the 2007 Annual TRB Meeting, Washington, D.C., USA, 2007.
10. S. D. CAFISO, C. D'AGOSTINO, *Safety performance function for motorways using generalized estimation*, Procedia-Social and Behavioral Sciences, 53, 901-910, 2012.
11. S. VIEIRA GOMES, S. R. GEEDIPALLY, D. LORD, *Estimating the safety performance of urban intersections in Lisbon, Portugal*, Safety Science, 50, 9, 1732-1739, 2012.
12. Automobile Club of Italy, *Car trend 2000-2012*, Statistic area, studies & search direction, Management control office, Automobile Club of Italy & National Institute of Statistics, Rome, Italy, 2000-2012.
13. O. GIUFFRÈ, A. GRANÀ, T. GIUFFRÈ, R. MARINO, *Improving reliability of road safety estimates based on high correlated accident counts*, Transportation Research Record, 2019, 197-204, 2007.
14. O. GIUFFRÈ, A. GRANÀ, F. CORRIERE, R. MARINO, *Handling under-dispersion in calibrating Safety Performance Function at urban, four-leg, signalized intersections*, Journal of Transportation Safety & Security, 3, 3, 174-188, 2011.
15. O. GIUFFRÈ, A. GRANÀ, T. GIUFFRÈ, R. MARINO, *Accounting for Dispersion and Correlation in Estimating Safety Performance Functions. An Overview Starting from a Case Study*, Modern Applied Science, 7, 2, 11-23, 2013.

16. K. M. Bauer, D. W. Harwood, *Statistical models of at-grade intersection accidents, addendum* (FHWA-RD-99-094), U.S. Department of Transportation, Federal Highway Administration Washington, DC, 2000.
17. H. McGee, S. Taori, B. N. Persaud, Crash experience warrant for traffic signals (NCHRP Report 491): Transportation Research Board, National Cooperative Highway Research Program Washington, DC, 2003.
18. M. J. Maher & I. Summersgill, *A comprehensive methodology for the fitting of predictive accident models*, Accident Analysis & Prevention, 28, 3, 281-296, 1996.
19. P. McCullagh, J. A. Nelder, *Generalized linear models*, Chapman and Hall: London and New York, 1989.
20. D. Lord, F. Mannering, *The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives*, Transportation Research Part A: Policy and Practice, 44, 5, 291-305, 2010.
21. S. Miaou, *The relationship between truck accidents and geometric design of road sections: Poisson versus negative Binomial regressions*, Accident Analysis and Prevention, 26, 4, 471-482, 1994.
22. Cameron, A. C., & Trivedi, P. K., *Regression Analysis of Count Data*, Cambridge University Press; Cambridge, 1998.
23. K. Y. Liang, S. L. Zeger, *Longitudinal data analysis using generalized linear models*, Biometrika, 73, 1, 13–22, 1986.
24. P. J. Diggle, P. Heagerty, K. Y. Liang, S. L. Zeger, *Analysis of Longitudinal Data*, New York Oxford University Press, 396 p. ISBN: 0198524846, 2002.
25. H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on Automat Control, 19(6): 716 – 723, DOI: 10.1109/TAC.1974.1100705, 1974.
26. W. Pan, *Akaike's information criterion in generalized estimating equations*, Biometrics 57, 120-125, 2001.
27. D. Lord, *Modelling 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 & Prevention 38, 4, 751-766, 2006.

Received: 26.09.2013
Revised: 25.02.2014