THE BALTIC JOURNAL
OF ROAD AND BRIDGE ENGINEERING

ISSN 1822-427X print / ISSN 1822-4288 online
2014 Volume 9(2): 77–87

CONSISTENT APPROACH TO PREDICTIVE MODELING AND COUNTERMEASURE DETERMINATION BY CRASH TYPE FOR LOW-VOLUME ROADS

Francesca Russo[1]✉, Salvatore Antonio Biancardo[2], Gianluca Dell'Acqua[3]

Dept of Civil, Construction and Environmental Engineering, University of Naples Federico II,
Via Claudio 21, I-80125 Naples, Italy
E-mails: [1]francesca.russo2@unina.it; [2]salvatoreantonio.biancardo@unina.it; [3]gianluca.dellacqua@unina.it

Abstract. The object of this research is to develop one and only injury crash rate prediction model differentiable for three main crash types (head-on/side collisions, rear-end collisions, single-vehicle run-off-road crashes) observed on the selected Italian two-lane rural roads in low-volume conditions. An explanatory variable reflecting road "Surface" conditions (dry/wet), "Light" conditions (day/night), and geometric "Element" (tangent segment/circular curve) when the crash happened and referred to the police reports has been proposed within the safety performance function all together (Surface, Light and Element) with three other significant variables (lane width, horizontal curvature indicator and mean speed) as consistent factors to predict crashes and their degree of seriousness for different kind of crashes. Among different statistical approaches introduced in the past few years to deal with the data and methodological issues associated with crash-frequency data, a generalized estimating equation has been implemented to take into account over-dispersion of the crash data, with a negative binomial distribution additional log linkage equation. Residual plots were combined with the validation procedure and other goodness-of-fit measurements to determine the reliability of the results. Potential countermeasures have been proposed for the critical crash types surveyed on the studied roads; these countermeasures have had positive effects on the road segments where the serious crash types have occurred over an eight-year period of analysis.

Keywords: low-volume roads, predictive safety for crash type and road scenario, road structural strategies.

1. Introduction

Low-volume roads (LVRs), as presented in this paper, are characterized as roads with fewer than 1000 vehicles per day (vpd) for most times of the year, although the definition can differ. These roads need to be well planned, well-constructed, and properly maintained in order to have minimal adverse impact on the surrounding environment (Dell'Acqua *et al.* 2012).

Moya *et al.* (2011) proposed a management system for the Spanish low-volume road system based on an analysis of the present management of the country's road and low-volume road networks, the methodologies used in the inspection and evaluation of the latter's condition and on data from other countries. The proposed plan involved the stages of implantation, execution and monitoring, each one with its own objectives. The possible source of funding for such a management system was also discussed.

Stamatiadis and Hartman (2011) promoted a concept of the context-sensitive solution (CSS) as the best practice for project development. A CSS provides a systematic and comprehensive approach to project development from inception and planning through operations and maintenance. Recently, the economic constraints facing several state departments of transportation have created a new emphasis on financial issues as they relate to project development. Practical design and practical solutions are a process emphasizing design solutions that aim to achieve the maximum rate of return for the individual project and to maximize system returns.

Reducing crashes on highways has always been one of the most important concerns for transportation engineers during the processes of planning, design, construction, and maintenance. Providing a safe driving environment is indeed not only a responsibility, but also one of the top priorities for all highway projects (Chuo, Saito 2009).

Jasiūnienė *et al.* (2012) described the method for selecting and prioritizing road sections which have higher than the average accident saving potential in each road category. When selecting road sections for treatment, a potential reduction of accident costs shall be taken into

 doi:10.3846/bjrbe.2014.10

consideration. Road sections in each category were studied in Lithuania and classified by the factors related to road safety, such as the number of accidents, traffic flow and road characteristics.

Many authors dealt with specific crash types to identify causes, effects and appropriate countermeasures.

Srinivasan *et al.* (2009) showed, for example, that approx 27% of fatal crashes on the U.S. highway system occurred in horizontal curves of two-lane rural highways. Approx 70% of curve-related fatal crashes were single-vehicle crashes in which the vehicle left the roadway and struck a fixed object or overturned, and 11% of curve-related fatal crashes were head-on crashes.

Hu and Yang (2011) have studied the Golumd–Kunlun Mountain section of the low-volume Qinghai–Tibet highway. Their analysis showed that 72.5% of accidents occurred on straight sections and 9.2% on curves. Vehicle rollover occurred in 47.5% of all accidents, mostly on a combination of vertical and horizontal curves and a small-radius curve.

The Highway Safety Manual (HSM) published by AASHTO in 2010 gives safety performance functions (SPFs) to estimate the number of crashes over a specific roadway over a specific time period and safety countermeasures for those highways; in particular chapter 10 provides predictive models for rural two-lane highways giving estimates for total crashes. Because the SPF equations in the HSM were developed on the basis of data from a subset of states, HSM recommends that local agencies should either develop SPFs for their local conditions or use a calibration procedure to adjust the HSM SPFs to reflect local conditions.

The HSM identified on the rural two-lane two-way roadway segments six collision types for the single-vehicle accidents (collision with animal, collision with bicycle, with pedestrian, overturned, ran off road, other single-vehicle accidents) and five collision types for the multiple-vehicle accidents (angle collision, head-on collision, rear-end collision, sideswipe collision, other multiple-vehicle collision). HSM presents computational procedures for safety effectiveness evaluation methods of treatments as a percentage change in crashes, including the empirical Bayesian method, the comparison-group method, and the shift in proportions method for a specific target collision type. Steps are illustrated to assess whether a treatment significantly has affected the proportion of crashes of the collision type under consideration: nonparametric statistical method, the Wilcoxon signed rank test, is used to test whether the average difference between after and before proportions of crashes over all n treatment sites, is significantly different from zero at a predefined confidence level.

For several years now, the authors of this research work deal with roadway safety to evaluate how human, infrastructural and environmental factors can influence the collision types (Dell'Acqua *et al.* 2013; Russo *et al.* 2014). The objective of the study presented here is to develop an injury crash rate prediction model that is differentiable for three specific crash types identified on the studied two-lane rural roads in low-volume conditions (head-on/side collision, rear-end collision, single-vehicle run-off-road crash). A variable that reflects the type of crash, road element where the crash was detected (from the police report), and road surface/light conditions when the crash happened on the homogeneous road segment was introduced into the safety performance function. Because of over-dispersion, a generalized estimating equation was adopted. The research illustrated here follows a "network" approach for the safety analysis of the investigated road segments:

a) identification of the injury critic crash type on the studied road network with highest frequency of occurrence on the roadway segments during the analysis period;

b) identification of "black" roadway segments for a specific crash type, by using an injury crash rate prediction model, where the crash injury rate is higher than on the rest of the roads;

c) identification of accurate and precise countermeasures for the crash dynamics. These steps will be integrated in the future developments with the following future advances;

d) assessment of the difference between the after and before proportions of the injury crash rates at each treatment site for a specific target collision type;

e) assessment of the average difference between after and before proportions over all *n* treatment sites;

f) assessment of the statistical significance of the average shift in proportion of the target collision type.

The data set used for this study includes injury crashes from 2003 through 2010 on some two-lane rural roads falling within the road network in the Southern Italy. The experimental analysis was divided into two phases:

phase I – the calibration phase involving 300 km of analyzed two-lane rural roads and

phase II – the validation phase involving an additional 300 km not included in the first step on which to test the effectiveness of the model.

2. Literature review

Technical standards pervade commerce and society and allow the complexity of modern life to operate at all levels, global included. They also provide protection against many risks. A very common reason for standards is safety, which tends to be treated as an absolute, as an objective technical matter, and there is less oversight or quality assurance of standards set for safety reasons. But safety is a relative term and increases in safety will usually have costs. Judgment needs to be applied in a risk assessment but that raises institutional issues as to who is qualified to apply that judgment (Macrae 2011).

Preventing traffic crashes and reducing crash severity are of keen interest to transportation professionals and policy makers. Researches have been conducted to select countermeasures (D'Acierno *et al.* 2011) and a variety of causal factors have been associated with the mechanism of crash occurrence (Kweon, Oh 2011).

Kirk and Stamatiadis (2010) conducted a research to identify the specific driving maneuvers in specific types of crashes involving younger drivers. They explained that drivers' inexperience is the largest single contributor to their increased crash rates.

Zhu *et al.* (2010) specifically evaluated crash type as a categorical variable with binary logic models to predict the probability that a fatal crash will be a single-vehicle run-off-road fatal crash given roadway design characteristics, roadside environment features, and traffic conditions proximal to the crash site. In a model transferability assessment, the authors determined that lane width, horizontal curvature, and ambient lighting are the only three consistent significant variables for single-vehicle run-off-road crashes for all study locations.

Keall and Newstead (2009) identified some important features of crash occurrence useful for making choices of comparison crash types when controlling for exposure. The best sets of comparison crashes were found to be multi-vehicle crashes in which the vehicle was damaged in the rear or multi-vehicle crashes in which the driver was adjudged to be not at fault. Likely bias of induced exposure risk estimates was identified, even for these best sets of comparison crashes, according to vehicle size and owner age and gender.

Khan *et al.* (2012) evaluated the safety of horizontal curves through the use of curve geometric characteristics and sign data. The focus was on collecting a good-quality large data set to develop models and explore the relationship between safety at horizontal curves and sign types, specifically curve and turn signs. The crash prediction models displayed highly significant variables, which showed a positive relationship with Annual Average Daily Traffic (*AADT*, vpd), posted speed, and curve length, and a negative relationship with curve radius.

Xie *et al.* (2011) reviewed the use of the HSM calibration procedure and its application to the State of Oregon; they also presented an evaluation of crash severity distribution methods and an assessment of the significance of collision type distributions on the overall predicted crashes.

El-Basyouny and Sayed (2006) compared two types of regression techniques: the Traditional Negative Binomial (TNB) and the Modified Negative Binomial (MNB). While the TNB approach assumes that the shape parameter of the negative binomial distribution is fixed for all locations, the MNB approach assumes that the shape parameter varies from one location to another.

Qin (2012) applied an alternative crash modeling approach: Quantile Regression (QR) in the context of a count data model. QR model for crash count data confirmed that crash predictors have varying impacts on the different areas of the crash distribution and the marginal effects of covariates provide a more direct observation of changes in the quantity.

Kweon and Oh (2011) developed a modeling approach to identify promising road segments for safety improvement through speed management strategies and to illustrate how to select segments on the basis of model results. The study involved the application of four statistical techniques (generalized additive model, negative binomial model, linear model, and empirical Bayes method) in three sequential steps to data collected on a 190 km section of expressway in South Korea.

Chuo and Saito (2009) selected three two-lane rural highway sections in Utah for a sample application of Crash Prediction Model (CPM) to safety audits. The results of this evaluation showed that the CPM produces reasonably consistent crash predictions if appropriate input data, especially alignment data, reflect existing conditions with reasonable accuracy. The localization of the hot spots for detailed safety audit improves the safety audit task's focus and effectiveness.

3. Data collection

The crash data used in this research involved almost 600 km of two-lane rural roads in Southern Italy located in the flat area with a vertical grade of less than 6%; one half were used for the calibration procedure of the safety performance function, attempting to predict the number of injury crashes for year for km for 10^8 vehicles (crash frequency over traffic exposure) and the other half to check the effectiveness of the crash prediction model in the validation procedure. The roadways analyzed are in low-volume conditions with an *AADT* of less than 1000 vpd over a period of 8 years (2003–2010); 56 road segments with the same curvature indicator were adopted in phases I and II as illustrated below. In particular, these segments have a mean length of 3.5 km, mean speed of 65 km/h, mean road width (travel lanes plus shoulders) of 7.00 m with a minimum of 5.00 m and a maximum of 12.00 m, and a mean value for the curve radius of 150 m. Crash data was made available to the research group by the Administration of the Province of Salerno. Over the total length of the network analyzed, 168 injury crashes were observed from 2003 to 2010, with 256 injuries and 13 deaths, and a mean injury crash rate of 40 (the number of injury crashes per year per km per 10^8 vehicles on the horizontal homogeneous segment of two-lane rural roads). The relationship between injury crashes and driving behavior was analyzed from the point of view of the dynamic interactions between driver, vehicle, road, and environment. Three main crash types were identified:

– crash type I – head-on/side collisions with 92 total injury crashes;

– crash type II – single-vehicle crashes (vehicle exits the roadway and either strikes a fixed object or overturns) with 48 total injury crashes;

– crash type III – rear-end collisions with 28 total injury crashes.

Analysis showed that collisions crash type I occurred in 55% of cases with 68% of injury crashes on the circular curves and 32% on the tangent segment, while crash type II occurred in 29% of cases with 62% of injury crashes on the circular curves and 38% on the tangent segment. Crash type III made up only 16% of cases, with 40% of injury crashes on the circular curves and 60% on the tangent segment.

The geometric-environmental-light conditions of the crash scenario were enclosed in a single variable by the original acronym *SLE*. This acronym devised by the authors reflects the first letters of the words "Surface", "Light" and "Element" and it reproduces the identified road "Surface" (dry/wet), "Light" conditions (day/night), and geometric "Element" (tangent segment/circular curve) when the crash happened. *SLE* reflects specific combinations of the road geometric-environmental-light conditions which influence the consequences of a crash type when they are shared by a specific set of the Mean Speed (*MS*) + Roadway Width (*W*) + Curvature Indicator (*CI*) values.

SLE code is higher for the more dangerous investigated combinations in terms of collected injury and death, and lower for those less dangerous.

Table 1 gives an overview of the analyzed injury crashes varying the crash type and the crash context (scenario) during the total study period.

Table 1 shows that the maximum value for the injury crash rate for the analyzed period is for the crash type I in the WDC scenario, and the minimum value is for the crash type III in the DNT scenario. In addition, WDC scenario is the most dangerous scenario for all thre crash types. But the DNT scenario has for the crash type I alone the lowest injury crash rate while for the crash type II no low injury crash values rates were recorded. In the latter case, the minimum value is WNT.

A frequent injury crash is related on one hand to road geometric failures but on the other hand also to human factors such as degree of comfort, knowledge of the environment, and a driver's ability to perceive the coming road alignment. Knowing the critical injury crash type (id est, injury crashes with highest frequency of occurrence on a homogeneous segment during the analysis period), it is possible to calibrate an accurate safety predictive model that reflects all the features when the crash happened in order to identify detailed countermeasures. Roadway construction and maintenance procedures for LVRs are considered to be complicated because of their peculiar characteristics. Limited resources often exist for LVRs while huge sums of money are typically spent on ordinary roadways.

Since during the years analyzed, the roads in question had a high number of crashes of type I followed by crashes

Table 1. Overview of the analyzed injury crashes in 2003–2010

Scenario		Crash type* I			Crash type* II			Crash type* III		
		Distribution, %	Crash rate		Distribution, %	Crash rate		Distribution, %	Crash rate	
Dry Road Surface + Daylight + Curve	DDC	26	Min	15.55	17	Min	10.61	–	Min	–
			Mean	36.04		Mean	16.69		Mean	–
			Max	78.29		Max	22.77		Max	–
Dry Road Surface + Daylight + Tangent Segment	DDT	13	Min	17.61	10	Min	10.61	20	Min	1.47
			Mean	38.53		Mean	45.36		Mean	13.96
			Max	94.97		Max	80.11		Max	37.66
Dry Road Surface + Night + Curve	DNC	14	Min	9.54	20	Min	7.79	20	Min	1.47
			Mean	21.53		Mean	8.88		Mean	13.96
			Max	33.92		Max	9.97		Max	37.66
Dry Road Surface + Night + Tangent Segment	DNT	12	Min	7.79	17	Min	7.58	40	Min	0.50
			Mean	20.64		Mean	27.02		Mean	9.92
			Max	33.92		Max	46.69		Max	19.33
Wet Road Surface + Daylight + Curve	WDC	23	Min	10.61	10	Min	10.61	20	Min	0.98
			Mean	58.04		Mean	30.90		Mean	34.64
			Max	120.81		Max	118.05		Max	68.29
Wet Road Surface + Daylight + Tangent Segment	WDT	7	Min	18.99	6	Min	10.61	–	Min	–
			Mean	48.64		Mean	25.58		Mean	–
			Max	78.29		Max	33.94		Max	–
Wet Road Surface + Night + Curve	WNC	5	Min	7.79	15	Min	28.76	–	Min	–
			Mean	25.03		Mean	31.34		Mean	–
			Max	40.62		Max	33.92		Max	–
Wet Road Surface + Night + Tangent Segment	WNT	–	Min	–	50	Min	4.78	–	Min	–
			Mean	–		Mean	7.58		Mean	–
			Max	–		Max	19.34		Max	–

Note: * – crash type I – head-on/side collision, crash type II – single vehicle run-off-road crash, crash type III – rear-end crash.

of type II, and the crash type I recorded a mean value of 35.45 for the injury crash rate, which is higher than the estimated injury crash rate for the crash type II (29.40) and crash type III (21.48), the crash type I was designated as the critical crash type.

A careful analysis of the database has shown that a wide variety of factors appear to influence or be associated with the crash dynamic. It was determined by deep examination and several statistical studies that Surface (dry/wet) and Light conditions (day/night), the location of the detected crash (tangent/circular curve element), lane width, measurement of the curvature change rate, mean speed are consistent variables to explain the effects of a crash. Of the vehicles involved in the crashes, cars made up the highest proportion, i.e., 75%, and two-wheeled vehicles were second, i.e., 14%. The remaining 11% included trucks, tractors, buses, and agricultural vehicles.

4. Data analysis

4.1. Calibration data

An injury crash rates prediction model was developed to predict the number of injury crashes per year per km for 10^8 vehicles (crash frequency over traffic exposure) that is differentiable for three specific crash types identified on the studied two-lane rural roads (head-on/side collision, rear-end collision, single-vehicle run-off-road crash). In the calibration phase involving 5 years' traffic taken from the crash database (2003–2007) was used 300 km of the total length of the analyzed network for a total of 73 injury crashes (120 injuries and 4 deaths) of which 48% were of the crash type I (100 injuries and 3 deaths), 30% of the crash type II (16 injuries and 1 death), and 22% of the crash type III (4 injuries).

Table 2 synthetically shows the overall features of the crash database used in the calibration phase according to the scenarios listed in Table 1:

– mean value (μ) of crash rate for each homogeneous segments showing the number of injury crashes for year for km for 10^8 vehicles on the homogeneous road segment *i*;

– the coefficient of variation (CV equal to the standard deviation, σ, divided by mean value, μ, of the injury crash rate) is an indicator of measurement dispersion to compare different scenarios; all defined scenarios have a CV of less than 1 and this means that no scenario has a dispersion of crash data higher than another scenario and so each scenario involves really homogeneous crash data from the point of view geometric-environmental-light conditions;

Table 2. Descriptive statistics of crashes analyzed in 2003–2007

Scenario	Crash type*	Injuries number	Deaths number	μ	σ	CV	CI	Homogenous segments number	Total investigated road length, km	*MS*, km/h	*W*, m
	I	10									
DDC	II	3	0	31.20	20.85	0.7	2.71	8	55	60.00	6.00
	III	0									
	I	31	1								
WDC	II	1	1	58.04	40.52	0.7	2.36	12	50	62.00	7.00
	III	2	0								
	I	12									
WDT	II	1	0	40.95	32.50	0.8	2.67	3	15	58.00	6.00
	III	0									
	I	14	1								
DDT	II	1	0	38.85	30.05	0.8	2.42	7	51	63.00	7.00
	III	1	0								
	I	8									
WNC	II	2	0	27.13	11.28	0.4	2.31	6	25	64.00	7.00
	III	0									
	I	11	1								
DNT	II	5	0	22.37	12.68	0.6	2.49	12	64	64.00	7.00
	III	1	0								
	I	14									
DNC	II	3	0	18.15	10.42	0.6	2.66	8	40	62.00	6.50
	III	0									

Note: * – crash type I – head-on/side collision, crash type II – single vehicle run-off-road crash, crash type III – rear-end crash.

– *CI* (Curvature Indicator) is a measurement of the Curvature Change Rate (CCR, gon/km) of the homogeneous road segment. The homogeneous road segment is characterized by a CCR defined as the sum of the absolute values of angular changes in the horizontal alignment divided by the total length of the road section. A homogeneous roadway include more than one tangent segment ($length_{min}$ = 65.37 m; $length_{max}$ = 4699 m; $length_{mean}$ = 2225.92 m) and circular curve ($length_{min}$ = 22 m; $length_{max}$ = 218 m; $length_{mean}$ = 69 m; $Curve\_Radius_{min}$ = 25 m; $Curve\_Radius_{max}$ = 450 m; $Curve\_Radius_{mean}$ = 170 m). For each homogeneous segment a value between 1 and 3 was associated to define *CI*:

- 1 – for low road horizontal curvature (CCR < 50 gon/km);
- between 1 and 2 – for medium road horizontal curvature (50 gon/km ≤ CCR ≤ 300 gon/km;
- between 2 and 3 – for high road horizontal curvature (CCR > 300 gon/km);

– *MS* is the Mean Speed at each analyzed homogeneous roadway segment;

– *W* is the Width of the travel lanes plus shoulders.

From the above, regarding the importance of the influence of geometric-environmental-light conditions on the crash type that certainly affected driver behavior, a variable was performed by using *SLE* acronym.

The values of the *SLE* variable were set between 1 and 7 with increments of 0.5 from sensitive analyses checking the results in the validation procedure:

a) 1 for scenarios characterized by a lower crash rate (< 10) associated to specific combinations of crash type-infrastructural/environmental conditions;

b) 7 for scenarios characterized by a higher mean crash rate (≥ 90) for specific combinations of crash type-infrastructural/environmental conditions.

The values associated to the *SLE* variable varying the crash type, scenario and geometric-environmental-light

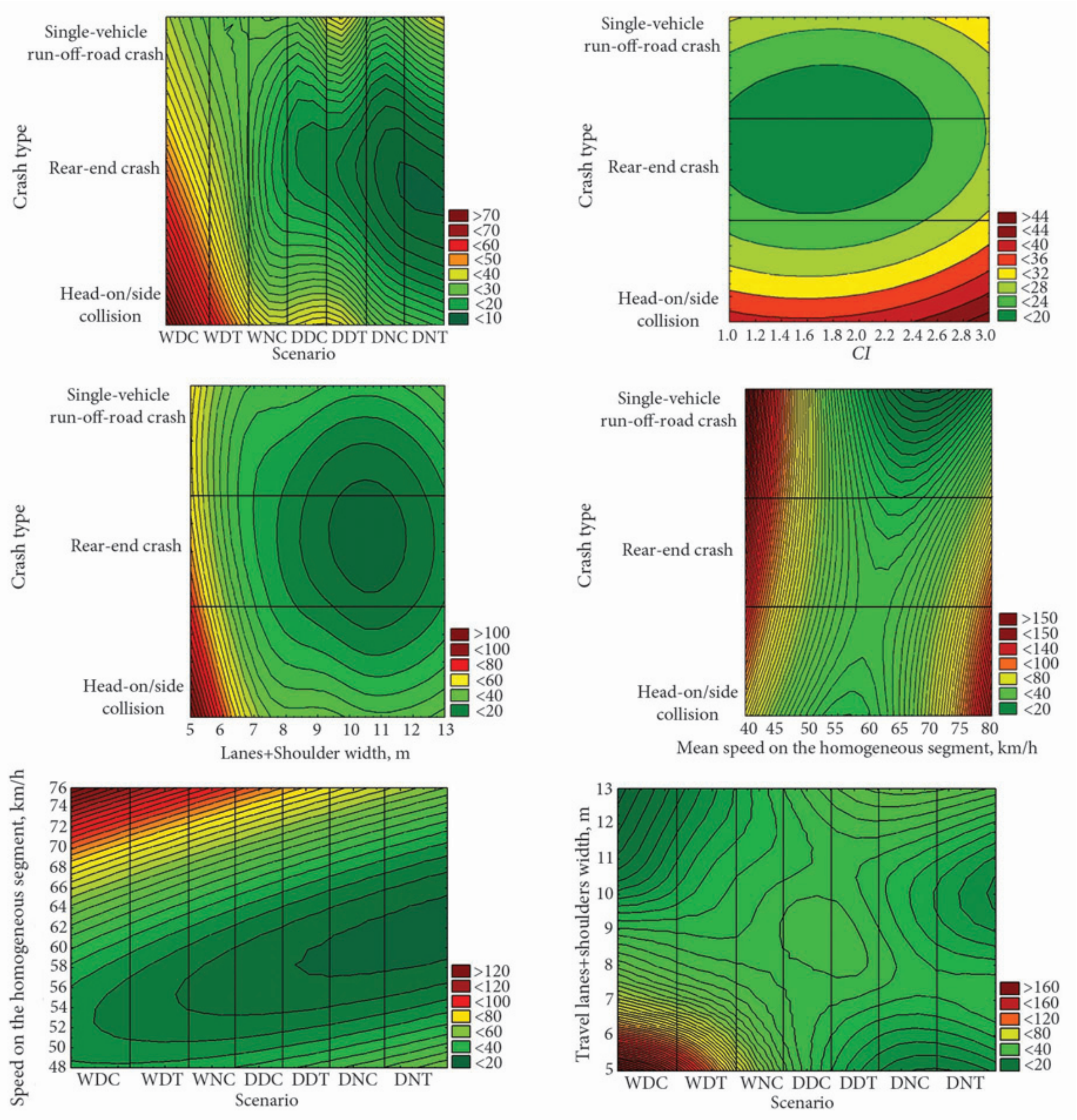

Fig. 1. Crash hazard maps

conditions are derived by careful analysis of the hazard density maps.

Fig. 1 shows some of the plotted risk type density diagrams. The diagrams (Fig. 1) illustrate the bands with different levels of injury crash rate for different combinations of crash types and geometric-environmental-light conditions. It is possible to move from a green band for low levels to a red band with high values. Analyzing all possible maps deriving from a plot of all possible combinations of cited features, the value for the *SLE* variable was chosen (Table 3).

Because of the crash data over-dispersion, a generalized estimating equation with an additional log linkage equation was adopted to calibrate the safety performance function. A preliminary analysis was conducted to identify the relationships between the possible independent variables by using a Pearson coefficient and the final variables considered for the crash rate prediction model were *SLE*, *MS*, *W*, and *CI* for each homogeneous road segment.

Before moving to the calibration phase, a technique for filtering anomalous crash calculated injury rates was adopted using the Vivatrat method widely used in geotechnical engineering.

The method is based on estimates of ranges of values considered fluctuations of the "regular" measures compared with values estimated as "abnormal". The procedure is summarized as follows:

1) divide the measures (injury crash rates) into homogeneous substrates according to CV values (in the case study, 7 scenarios have been identified);

2) order the scenarios in increasing the μ of the injury crash rates;

3) order the measurements of each scenario in increasing the injury crash rates measures;

4) determine μ and σ of injury crash rates distribution for each substrate;

5) determine the representative dispersion (S_r) for each scenario defined as the minimum value among the expressions (1–3):

$$S_r = \frac{1}{2}\left(S_{i+1} + S_i\right), \tag{1}$$

$$S_r = \frac{1}{2}\left(S_{i-1} + S_i\right), \tag{2}$$

$$S_r = \frac{1}{2}\left(S_{i+1} + S_{i-1}\right), \tag{3}$$

where S_{i-1}, S_i and S_{i+1} – the standard deviations (σ) for $(i-1)^{\text{th}}$ substrate, i substrate, and $(i+1)^{\text{th}}$ substrate, respectively.

6) for each scenario the measurements (injury crash rates) outside the range $\mu_i \pm AS_r$ were removed, where μ_i is the mean of the measures belonging to each scenario i, S_r is the σ and A is the coefficient characteristic that defines the amplitude of the semi-interval considered acceptable for the values assumed by the measures. It's need values of parameter A come between $2.5 > A > 0.5$. In the case study, no rate was rejected for each scenario because no values fall out of range $\mu_i \pm 2.5S_r$ and it was noted how more than 70% of calculated injury crash rates fall within the range [μ + 0.5σ; μ – 2.5σ] for each scenario.

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were mainly used to check the reliability of the model among many equations come out from several iterations. The preferred model was the one with the minimum AIC value. Hence AIC not only rewards goodness of fit, but also it includes a penalty that is an increasing function of the number of estimated parameters. This penalty discourages over fitting; generally, model selection performance improved as sample sizes increased. The ability of AIC to select a true model rapidly increases with the sample size and with the decrease of the σ of the sample size. Under unstable conditions such as small sample and large noise levels AIC outperforms BIC. The choice of the model that best fits the data is based on the AIC available values' comparison in small/medium samples as well as in the present study since the crashes are rare events, whilst on the BIC values in larger samples with high noise.

In addition to these criteria, other goodness-of-fit measures for GEE models were used: p-value to test significance of the regression coefficients (only coefficients with $p < 0.05$ were kept in the models), deviance, Pearson dispersion, Wald and Pearson χ^2 test. Eq (4) shows the best of the safety performance function

$$Y = e^{\left(-0.07W + 0.45SLE + 0.27CI + 0.034MS\right)}, \tag{4}$$

where Y – number of injury crashes for year for km for 10^8 vehicles on the road segment i calculated as follows:

$$Y = \frac{n_i \cdot 10^8}{L_i \cdot 365 \cdot \sum_{j=1}^{x}\left(AADT\right)_j}, \tag{5}$$

where n_i – number of injury crashes recorded during x years on the homogenous segment i; L_i – length of the homogeneous segment i, km; x – number of years studied to calibrate the safety performance function; $AADT_j$ – $AADT$ on an annual basis (for year j^{th}); W – roadway width as travel lanes plus shoulders, m; SLE – this variable assumes the value in Table 3 depending on the crash type to be studied, on a road segment to be analyzed (circular curve or tangent segment), on a dry or wet road surface, light conditions (day or night), on W, MS and CI.

The minimum AIC was 7.11, and BIC was –120.05, with a maximum log-likelihood value of – 195.224. The goodness-of-fit measures of the final model were deviance 0.04 and Pearson dispersion $5.94 \cdot 10^{-4}$. Eq (4) does not apply to areas near the intersections that require ad hoc models.

4.2. Validation data

A validation procedure was performed to test the reliability of the safety performance function: crash data covered three years (2008–2010). 95 injury crashes were analyzed with

Table 3. Assignment of the scenario variable code according to crash type

The values of the SLE	Crash type*	*MS*, km/h	*W*, m	*CI*	SLE code
WDC	I	> 65	$5 \le W \le 7$	3	7
WDC	I	> 65	$7 \le W \le 12$	3	5
WDC	I	< 65	$5 \le W \le 7$	2	3
WDC	I	< 65	$7 \le W \le 12$	2	2.50
WDC	I	< 65	$9 \le W \le 12$	2	1.50
WDT	I	> 65	$5 \le W \le 7$	3	4.50
WDT	I	< 65	$7 \le W \le 12$	2	1.50
WNC	I	> 65	$7 \le W \le 12$	$2 < CI < 3$	2
WNC	I	< 65	$9 \le W \le 12$	$1 < CI < 2$	1
DDC	I	> 65	$5 \le W \le 7$	$2 < CI < 3$	4.50
DDC	I	> 65	$7 \le W \le 12$	2	2.50
DDC	I	< 65	$5 \le W \le 7$	$2 < CI < 3$	2
DDC	I	< 65	$9 \le W \le 12$	$1 < CI < 2$	1.50
DDT	I	> 65	$5 \le W \le 7$	3	5.50
DDT	I	> 65	$5 \le W \le 7$	2	2.50
DDT	I	< 65	$7 \le W \le 12$	$2 < CI < 3$	2
DDT	I	< 65	$7 \le W \le 12$	2	1.50
DNC	I	> 65	$5 \le W \le 7$	$2 < CI < 3$	2.50
DNC	I	> 65	$7 \le W \le 12$	$2 < CI < 3$	1.50
DNC	I	< 65	$9 \le W \le 12$	$1 < CI < 2$	1
DNT	I	> 65	$5 \le W \le 7$	$2 < CI < 3$	2.50
DNT	I	> 65	$9 \le W \le 12$	$2 < CI < 3$	2
DNT	I	< 65	$5 \le W \le 7$	$1 < CI < 2$	1.50
DNT	I	< 65	$9 \le W \le 12$	$1 < CI < 2$	1
WDC	III	> 65	$7 \le W \le 12$	2	3
DDT	III	> 65	$7 \le W \le 12$	2	1.50
DNT	III	> 65	$5 \le W \le 7$	$1 < CI < 2$	1.50
DDT	II	< 65	$7 \le W \le 12$	3	4.00
DNT	II	> 65	$5 \le W \le 7$	3	3.50
DNT	II	> 65	$7 \le W \le 12$	$2 < CI < 3$	2.50
WDC	II	> 65	$7 \le W \le 12$	2	2.50
WNC	II	> 65	$7 \le W \le 12$	$1 < CI < 2$	2.50
WDT	II	> 65	$7 \le W \le 12$	$2 < CI < 3$	2
DDC	II	> 65	$7 \le W \le 12$	$2 < CI < 3$	1.50
DDC	II	> 65	$5 \le W \le 7$	$1 < CI < 2$	1.50
DNT	II	< 65	$7 \le W \le 12$	$2 < CI < 3$	1.50
DDC	II	< 65	$7 \le W \le 12$	$1 < CI < 2$	1
WNT	II	> 65	$7 \le W \le 12$	$2 < CI < 3$	1

Note: * – crash type I – head-on/side collision, crash type II – single vehicle run-off-road crash, crash type III – rear-end crash.

136 injuries and 9 deaths: 60% for the crash type I (61 injuries and 1 death), 27% for the crash type II (74 injuries and 7 deaths), 13% for the crash type III (1 injury and 1 death).

The validation procedure consisted of the assessment of the mean deviation (D_i) (differences between observed and predicted injury crash rates at each homogeneous road segment, the mean absolute deviation (MAD) on the investigated road network, the mean squared error (MSE), the value I equal to the square root of MSE divided by the mean predictive injury crash rate value over the total length of the analyzed network. The diagram of cumulative squared residuals plotted on the basis of an *AADT* showed the absence of vertical jumps (more correctly known as "outliers").

The values returned by the analysis of summarizing the statistical parameters are the following: mean deviation = 3.93; MAD = 5.69; MSE = 53.97; I = 0.13. It was noted how for injury crash rates residuals, more than 50% fall within the range μ ± σ; in fact 70% fall in the range [μ – σ; μ + σ] = [–10.30; 2.44], while 98% are within the range [μ – 2σ; μ + 2σ] = [–16.67; 8.82], where μ is the mean value and σ is the standard deviation of this difference.

In conclusion, the injury crash rate prediction model is statistically significant because the residuals fall within a limited range around the mean and this was confirmed by the low value of the MAD and I indicators, and the absence of jumps in the diagram of cumulated squared residuals.

5. Results

The research work aimed to calibrate an injury crash rate prediction model as a function of crash type, "Surface" conditions (dry/wet), "Light" conditions (day/night), and geometric "Element" (tangent segment/circular curve). Explanatory variables suggest potential countermeasures to reduce the injury crash rates of the most critical crash type on the investigated road network. For example, Fig. 2 refers to a specific scenario study: circular curves during daylight and wet road surface conditions when the crash of type I (head-on/side injury collision) happens.

In the diagram, the *y* axis shows the injury crash rate (in log scale), while the *x* axis shows an independent variable of the predictive model, for example, the MS on the geometric element. Fig. 2 presents a series of straight lines with a constant value for the remaining independent variables of the model in appropriate combination. The number of possible profiles of the Eq (4), and consequently the number of potential safety solutions, is equal to the number of available variables employed in the model.

For example, to lower the injury crash rate for circular curves on LWR's, variations on the speed, curvature indicator, and roadway width are carried out. In Fig. 2, by using the straight line, which represents the combination of width and curvature associated with the analyzed roadway segment, and knowing the MS on the roadway segment, it is possible to predict the number of injury crashes per year per km.

Among other things, it is possible to reduce the injury crash rate on the roadway segment without changing its width and curvature by reducing speed. One must move along the same straight line to the left. In the same way, it is feasible to reduce the crash density on the roadway segment with no change in speed but only in roadway width and curvature, moving downward along a vertical line. Therefore, a change in speed, curvature and width is needed.

For example, Fig. 2 shows that if a reduction in speed is planned to decrease the *MS* from 70 km/h to 50 km/h that is the posted speed limit on our roads moving on the first emerald green line leftwards characterized by 6 m of *W* and *CI* high, an adequate reduction of the injury crash rate per year per km is obtained as well as social cost.

The adjustment on a circular curve site considered dangerous from the viewpoint of road safety was actually carried out as it falls into the previously mentioned scenario; the crash trend was then monitored during 2007–2010 after the intervention to check the consistency of the operation and it was witnessed a real reduction in speed to 50 km/h and in crash rates.

Dell'Acqua (2011) illustrated in his previous work the adjustments carried out: the installation of gateways to slow the vehicles entering the built-up area, and the traffic-calming devices aimed to complement the gateway effect (Fig. 3). The gateways at the beginning of the homogeneous segment consist of a doorway in galvanized steel like a flag portal, preceding rumble strips and dragon's teeth forming a virtual narrowing of the lanes.

Weijermars and Wegman (2011) have proposed, for example, important traffic calming measures such as the construction of 30 km/h and 60 km/h zones as well as serious traffic enforcement. The authors obtained a positive effect on traffic safety through these measures.

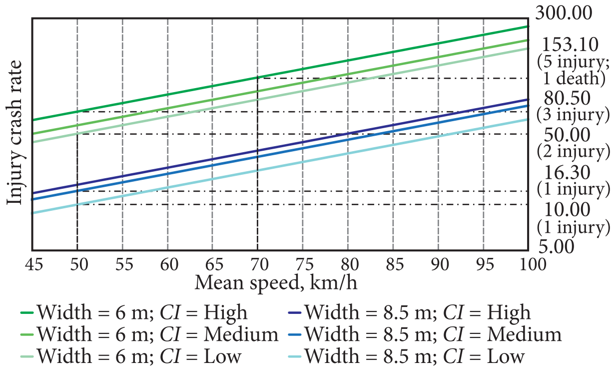

Fig. 2. Injury crash rates' abacus for a crash type and scenario

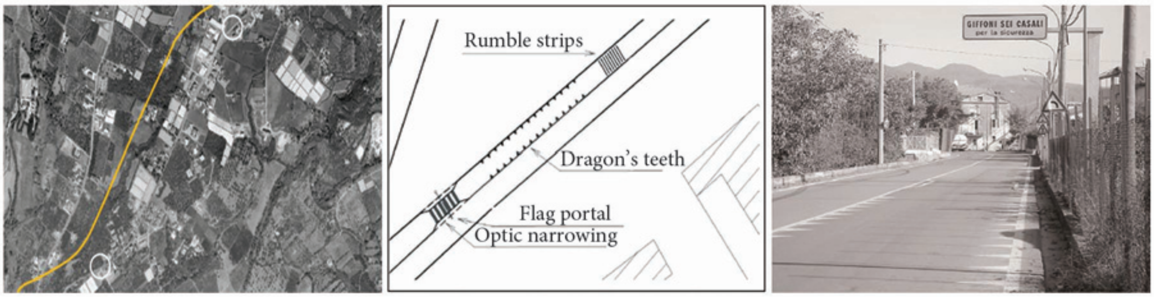

Fig. 3. Traffic calming design in the dangerous crash scenario (Dell'Acqua 2011)

Analyzing the social benefits on the circular curve segment according to the traffic calming adjustments proposed by Dell'Acqua (2011), a benefit-cost ratio of 17 was obtained: 1 since a death amounts to 1 377 933.00 EUR and 1 injury amounts to 26 688 EUR. By this adjustment aimed at a reduction of only the *MS*, a reduction in crash frequency was actually obtained, an increase in social benefits, a reduction in the number of injuries, and zero deaths.

Furthermore, be modifying the geometry (Fig. 2), according to the results of Eq (4), the social benefit increases (benefit-cost ratio 50:1) growing the reduction in the number of injuries and the crash rate with, obviously, a more expensive adjustment than in the first case.

6. Conclusions and future development

The research aimed to calibrate and validate an injury crash rate prediction model by varying crash type, geometric-environmental-light conditions to assess how candidate countermeasures help to reduce the frequency and social cost for specific types of crash and scenario on two-lane rural roads in low-volume conditions.

600 km of two-lane rural roads in low-volume conditions were used falling within the network of the Province of Salerno in Southern Italy for this study.

The research-study was divided into two phases: in phase I, a safety performance function was calibrated involving 5 years of the investigated crash database (2003–2007) with 55% of analyzed crashes being head-on/side collisions (crash type I) (100 injuries and 3 deaths), 29% single-vehicle run-off-road crashes (crash type II) (16 injuries and 1 death), and 16% rear-end collisions (crash type III) (4 injuries); in phase II, the results obtained previously were checked by using 3 years (2008–2010) of the whole crash database.

A network approach was used to create an injury crash rate prediction model by implementing a generalized estimating equation method with a negative binomial distribution and additional log linkage equation. An innovative variable by the original acronym *SLE* (first letters of the words "Surface", "Light", and "Element" and it reproduces the identified road "Surface" (dry/wet), "Light" conditions (day/night), and geometric "Element" (tangent segment/circular curve) when the crash happened) was enclosed in the Safety Performance Function that reproduces the identified road "Surface" (dry/wet), "Light" conditions (day/night), and geometric "Element" (tangent segment/circular curve) when the crash type happened. *SLE* code reflects the consequences of a crash type when a specific set of Mean Speed + Roadway Width + Curvature Indicator is shared. Goodness-of-fit measurements were used to determine the reliability of the results.

The application of the injury crash rate prediction model follows these steps:

1) homogeneous road segment identification in the analyzed network according to the Curvature Indicator;

2) identification of the number of curves and tangent segments for each homogeneous segment;

3) identification of the crash types during a study period;

4) characterization of environmental conditions (day/night) and road surface conditions (dry/wet) for which it is planned to predict the injury crash rates;

5) evaluation of width and mean speed for tangents and circular curves;

6) SLE code identification for specific crash type, geometric-environmental-light conditions;

7) sum of the injury crash rates for the circular curves and tangent elements falling within the same homogeneous segment after multiplying each crash rate for the whole extension of circular elements and tangent elements, respectively.

In this way, at each critical crash type, a consistent combination of interventions is defined and, subsequently, a reduction in the injury crash rates, severity and social cost for the more frequently expected and dangerous crash scenario are obtained. Thus, there are positive consequences on the less dangerous and frequent crash scenarios and cost reductions in the road maintenance-construction operations.

References

Chuo, K.; Saito, M. 2009. Applicability of FHWA Crash Prediction Module to Selecting Candidate Locations for Safety Audits of Two-Lane Rural Highways, *Transportation Research Record* 2137: 20–28. http://dx.doi.org/10.3141/2137-03

D'Acierno, L.; Ciccarelli, R.; Montella, B; Gallo, M. 2011. A Multimodal Multiuser Approach for Analysing Pricing Policies in Urban Contexts, *Journal of Applied Sciences* 11(4): 599–609. http://dx.doi.org/10.3923/jas.2011.599.609

Dell'Acqua, G.; Busiello, M.; Russo, F. 2013. Safety Data Analysis to Evaluate Highway Alignment Consistency, *Transportation Research Record* 2349: 121–128. http://dx.doi.org/10.3141/2349-14

Dell'Acqua, G.; De Luca, M.; Russo, F. 2012. Procedure for Making Paving Decisions with Cluster and Multicriteria Analysis, *Transportation Research Record* 2282: 57–66. http://dx.doi.org/10.3141/2282-07

Dell'Acqua, G. 2011. Reducing Traffic Injuries Resulting from Excess Speed Low-Cost Gateway Treatments in Italy, *Transportation Research Record* 2203: 94–99. http://dx.doi.org/10.3141/2203-12

El-Basyouny, K.; Sayed, T. 2006. Comparison of Two Negative Binomial Regression Techniques in Developing Accident Prediction Models, *Transportation Research Record* 1950: 9–16. http://dx.doi.org/10.3141/1950-02

Hu, J.-Bi.; Yang, Y. 2011. Safety of Driving Behavior on Low-Volume Roads in China Case Study of Qinghai–Tibet Highway, *Transportation Research Record* 2203: 100–105. http://dx.doi.org/10.3141/2203-13

Jasiūnienė, V.; Čygas, D.; Ratkevičiūtė K.; Peltola, H. 2012. Safety Ranking of the Lithuanian Road Network of National Significance, *The Baltic Journal of Road and Bridge Engineering* 7(2): 129–136. http://dx.doi.org/10.3846/bjrbe.2012.18

Keall, M.; Newstead, S. 2009. Selection of Comparison Crash Types for Quasi-Induced Exposure Risk Estimation, *Traffic Injury Prevention* 10(1): 23–29.

http://dx.doi.org/10.1080/15389580802383125

Khan, G.; Bill, A. R.; Chitturi, M.; Noyce, D. A. 2012. Horizontal Curves, Signs, and Safety, *Transportation Research Record* 2279: 124–131. http://dx.doi.org/10.3141/2279-15

Kirk, A.; Stamatiadis, N. 2010. Crash Rates and Traffic Maneuvers of Younger Drivers, *Transportation Research Record* 1779: 68–74. http://dx.doi.org/10.3141/1779-10

Kweon, Y.-J.; Oh, C. 2011. Identifying Promising Highway Segments for Safety Improvement through Speed Management, *Transportation Research Record* 2213: 46–52. http://dx.doi.org/10.3141/2213-07

Macrae, D. 2011. Standards for Risk Assessment of Standards: How the International Community Is Starting to Address the Risk of the Wrong Standards, *Journal of Risk Research* 14(8): 933–942. http://dx.doi.org/10.1080/13669877.2011.571782

Moya, M.; Gallego, E.; García, A. I.; Ayuga, F. 2011. A Proposed Methodology for the Management of Low-Volume Roads in Spain, *The Baltic Journal of Road and Bridge Engineering* 6(3): 153–162. http://dx.doi.org/10.3846/bjrbe.2011.20

Qin, X. 2012. Quantile Effects of Causal Factors on Crash Distributions, *Transportation Research Record* 2279: 40–46. http://dx.doi.org/10.3141/2279-05

Russo, F.; Biancardo, S. A.; Dell'Acqua, G. 2014. Road Safety from the Perspective of Driver Gender and Age as Related to the Injury Crash Frequency and Road Scenario, *Traffic Injury Prevention* 15(1): 25–33. http://dx.doi.org/10.1080/15389588.2013.794943

Srinivasan, R.; Baek, J.; Carter, D.; Persaud, B.; Lyon, C.; Eccles, K.; Gross, F.; Lefler, N. 2009. *Safety Evaluation of Improved Curve Delineation*. Report No. FHWA-HRT-09-045, Federal Highway Administration, Washington DC.

Stamatiadis, N.; Hartman, D. 2011. Context-Sensitive Solutions Versus Practical Solutions. What Are the Differences?, *Transportation Research Record* 2262: 173–180. http://dx.doi.org/10.3141/2262-17

Weijermars, W.; Wegman, F. 2011. Ten Years of Sustainable Safety in the Netherlands: an Assessment, *Transportation Research Record* 2213: 1–8. http://dx.doi.org/10.3141/2213-01

Xie, F.; Gladhill, F. K.; Dixon, K. K.; Monsere, C. M. 2011. Calibration of Highway Safety Manual Predictive Models for Oregon State Highways, *Transportation Research Record* 2241: 19–28. http://dx.doi.org/10.3141/2241-03

Zhu, H.; Dixon, K. K.; Washington, S.; Jared, D. M. 2010. Predicting Single-Vehicle Fatal Crashes for Two-Lane Rural Highways in Southeastern United States, *Transportation Research Record* 2147: 88–96. http://dx.doi.org/10.3141/2147-11

Received 13 November 2012; accepted 29 January 2013