The Variability of Urban Safety Performance Functions for Different Road Elements: an Italian Case Study

Paolo Intini[a]*, Nicola Berloco, Gabriele Cavalluzzi, Dominique Lord[b], Vittorio Ranieri and Pasquale Colonna

[a]*Department of Civil, Environmental, Land, Building Engineering and Chemistry, Polytechnic University of Bari, Bari, Italy;*

[b]*Zachry Dept. of Civil and Environmental Engineering, Texas A&M University, College Station, TX, United States.*

*corresponding author: email: paolo.intini@poliba.it, phone: +39 0805963390, fax: +39 0805963414

Abstract

Urban safety performance functions are used to predict crash frequencies, mostly based on Negative Binomial (NB) count models. They could be differentiated for considering homogeneous subsets of segments/intersections and different predictors. The main research questions concerned: a) finding the best possible subsets for segments and intersections for safety modelling, by discussing the related problems and inquiring into the variability of predictors within the subsets; b) comparing the modelling results with the existing literature to highlight common trends and/or main differences; c) assessing the importance of additional crash predictors, besides traditional variables. In the context of a National research project, traffic volumes, geometric, control and additional variables were collected for road segments and intersections in the City of Bari, Italy, with 1,500 fatal+injury related crashes (2012-2016). Six NB models were developed for: one/two-way homogeneous segments, three/four-legged, signalized/unsignalized intersections. Crash predictors greatly vary within the different subsets considered. The effect of vertical signs on minor roads/driveways, critical sight distance, cycle crossings, pavement/markings maintenance was specifically discussed. Some common trends but also differences in both types and effect of crash predictors were found by comparing results with literature. The disaggregation of urban crash prediction models by considering different subsets of segments and intersections helps in revealing the specific influence of some predictors. Local characteristics may influence the relationships between well-established crash predictors and crash frequencies. A significant part of the urban crash frequency variability remains unexplained, thus encouraging research on this topic.

Keywords: *safety performance functions; injury crash; urban segments; urban intersections.*

INTRODUCTION

The use of Safety Performance Functions (SPFs) is crucial for road safety purposes. Several functions were developed for rural and urban roads (Greibe, 2003; Gomes et al., 2012; Hauer et al., 2004; Vogt and Bared, 1998; Persaud and Dzibik, 1992; AASHTO, 2010). Few of these studies were conducted in Europe, especially for urban areas. While functions may be calibrated for being applied in other countries/regions (AASHTO, 2010; Geedipally et al., 2017), their transferability is not without issues (Farid et al., 2016; Sacchi et al., 2012).

Different aspects related to driving behaviour, cultural, geographic variables (AASHTO, 2010) may affect the model transferability. Transferability issues may be solved by applying a locally derived calibration factor. However, the effect of some variables (e.g., traffic volumes, geometric characteristics) may depend on the geographic context thus, a single calibration factor may not solve transferability issues (Geedipally et al., 2017; Colonna et al., 2016). In fact, the reliability and variability of calibration factors with geographic and road-related variables should be studied in detail (see e.g. Intini et al., 2019). Another option consists in estimating local SPFs, which may be particularly important in countries where they are scarcely used (Colonna et al., 2019).

Some European urban predictive models were developed, e.g., for Danish arterial segments and intersections (Greibe, 2003), Portuguese intersections (Gomes et al., 2012); Italian segments and intersections (Fancello et al., 2018; Canale et al., 2005). Some other studies were focused on specific crashes, such as vulnerable users (e.g., Jonsson, 2005). However, some of these studies are old, limited to specific road elements (e.g., roundabouts, segments or intersections), and/or the considered predictors were limited. In parallel, some other recent studies were focused on developing macro-level SPFs (Lee et al., 2019; Montella et al., 2019), including high-level variables, not specifically related to segments and intersections.

Research questions

Given the presented background, this study is based on the following research questions, which are intended to contribute to the existing body of research:

- systematically explore the crash performances of both urban segments and intersections, with the related influential variables thus searching for the best subsets of segments and intersections with homogeneous characteristics for modelling purposes, among different possibilities.
- Compare the significant predictors highlighted in the modelling stages with the significant crash predictors retrieved in previous research, to reveal specific local differences which may be of interest for further studies.
- Explore the influence of several other potential crash predictors, which are usually not considered in safety prediction studies, besides the traditional geometric and traffic control variables used in previous research.

Note that the article is not focused on assessing the optimal model and functional form for urban safety predictions, since the above reported research questions are explored in the context of the application of NB count models, which are best practice for urban safety predictions (e.g. Gomes et al., 2012; Vogt and Bared, 1998; Persaud and Dzibik, 1992). However, as previously anticipated, it is important to remark that such applications are scarce in the European context, especially considering the research questions posed in this study.

METHODS

Main dataset

In the context of the Pa.S.S.S. (Scientific Park for Road Safety) National research project (main agency: City of Bari, granted by the Italian Ministry of Transport and Infrastructures), the City of Bari (Italy) was chosen for data collection.

Fatal and injury crash data were collected in the period: 2012-2016[1]. They are crashes provided with generic information (e.g., date, hour), exact localization, information about vehicles and persons involved, crash type and circumstances, road-related variables.

Available traffic data from the City of Bari were coupled with crash data on the main interconnected urban network within the considered urban area (see **Figure 1**). After, weekday peak hour traffic counts were manually conducted (during 2018-2019, then converted into average daily volumes) to fill gaps in data obtained and to check for inconsistencies due to old traffic volumes and new roads openings. Traffic volumes were assumed as constant in the period: 2012-2019, coherently with average traffic volume trends in Southern Italy.

(Figure 1 here)

Samples of sites

The selected network was further divided into segments and intersections. Crashes were then linked to each segment and intersection identified for this research. Stop/yield lines or intersection zebra crossings (if present) were identified as a preliminary boundary between segments and intersections. However, based on the crash localization and description, crashes

[1]Data provided by the Puglia Regional agency ASSET (Puglia Strategic Regional Agency for the Eco-Sustainable Area Development), in coordination with ISTAT (National Institute of Statistics). The City of Bari (to which crash data are referred) is the County Seat of the Italian Puglia Region. All crash data collected by ISTAT are fatal and injury crashes with at least one vehicle involved. Fatal crashes in the City of Bari dataset account for 0.5% of the total.

in segment-to-intersection transition zones were manually assigned to either segments or intersections. Manual data explorations were preferred to predefined distance-based thresholds since they could present some arbitrariness and they may depend on the specific local context. Traffic volumes were divided into volumes on the main and the secondary entering roads. Segments were divided into "homogeneous" segments (**Figure 2**), by considering internal geometric or traffic control differences. 447 road "sites" were initially investigated: 325 homogeneous segments and 122 intersections.

(Figure 2 here)

Crash predictors

Several crash predictors were considered; most of them derived from ad-hoc inspections and/or online sources. For the sake of comparison with similar European models (taken from a previous literature review: Colonna et al., 2019), the main variables considered in Greibe (2003) and Gomes et al. (2012) were used. Lengths, speed limits, paved widths, minor roads/driveways, parking, land-use were collected for segments; while the number/width of intersecting road lanes, medians, turning lanes, number of one-way legs were collected for intersections. Other variables were considered such as sidewalks (Canale et al., 2005), vertical signs on minor roads/driveways, maintenance of pavements and markings (visually inspected), cycle paths, bus stops, reserved lanes, critical sight distance at intersections. The "critical" sight distance is considered in this article as the minimum available sight distance measured on all the intersecting legs of a road intersection, considering the obstacles on the roadside.

The selected variables and associated descriptive statistics are listed in **Table 1**.

(Table 1 here)

Data analysis techniques

Negative Binomial (NB) count data models were used to link crash frequencies to predictors. These models can account for the over-dispersion of crash data (Lord and Mannering, 2010) and they were used in similar studies (e.g. Gomes et al., 2012; Vogt and Bared, 1998; Persaud and Dzibik, 1992). NB models were estimated in *R* (Mass library: Venables and Ripley, 2002). The model structures used are reported in the Appendix. One of the research questions concerned the most appropriate way of disaggregating segments and intersections into subsets. Hence:

- preliminary models for the whole datasets of segments and intersections were run;
- two sub-categories for each family of sites (segments and intersections) were selected as based on results from preliminary models;
- models for each sub-category were run.

Disaggregating the dataset for research purposes results in reducing the initial sample size. The chosen level of significance was then set to $p = 0.10$, given the exploratory purposes and the limited dataset (similarly to e.g., Gomes et al., 2012). Injury severity modelling was not considered due to the scarce number of fatal crashes and the absence of injury scales (e.g., slight/serious/incapacitating) in the dataset. The Akaike Information Criterion (AIC) was used to comparatively assess different models and the Nagelkerke R^2 as a goodness-of-fit measure. In general, for each subset, the model having the least number of all significant variables included among different candidate best fitting models was selected. Results from each model obtained were compared to the corresponding null and full models through likelihood ratio tests.

RESULTS

Predictive models for segments

A model was firstly developed for the whole dataset of segments.

The main interest in this stage was to understand if the initial dataset can be efficiently differentiated into subsets. Among the different attempts performed, the explanatory power of the variable “Type of lanes” seems to be promising for a one-way/two-way classification. On the other hand, the variable “One-lane” (one or multilane segments) is never found as a statistically significant predictor. Hence, the dataset was divided into: one-way and two-way segments, rather than one-lane and multilane segments. The possible classification: undivided/divided segments was not taken into account since the segments divided by medians (for all their length) were firstly divided into two one-way segments (one for each direction), since directional traffic counts were generally available.

The model developed for all the segments is reported in **Table 2**. Crash frequencies increase with traffic volumes, segment length, two-way segments (with respect to one-way segments), presence of vertical signs on intersecting minor roads, parking on both sides (with respect to no parking). Bad pavement maintenance is associated with a decrease in the crash frequency. However, the included predictors can only explain a limited part of crash frequency, as based on the Nagelkerke R^2 value.

(Table 2 here)

When differentiating into one-way and two-way segments, some predictors are confirmed, while others are highlighted as well. For one-way segments, crash frequencies increase with traffic volumes, lengths, number of driveways/minor roads, vertical signs on minor roads/driveways; while they decrease with bad pavement maintenance. For two-way segments, crash frequencies increase with lengths, traffic volume (not significantly), parking (especially at both sides compared to prohibited parking), visible markings (although marginally significant at the 5 % level), vertical signs on minor roads/driveways. Note that the model which includes traffic volume was selected among other possibilities to avoid further worsening the limited model fit (traffic volume coefficient is significant, $p < 0.10$, if traffic is the only predictor).

Predictive models for intersections

A model was firstly developed for the whole dataset of intersections.

The main interest in this stage was to understand if the initial dataset can be efficiently differentiated into subsets. In this case, two promising models for the whole intersections were selected (see **Table 3**). The first model indicates the number of legs as an important explanatory variable. However, when trying to exclude all other possible correlated variables (turning lanes, number of legs, intersection control), the variable signalized/unsignalized assumes a notable importance in the alternative model. Hence, based on this, specific models were developed for two pairs of subsets: three-legged and four-legged intersections, signalized and unsignalized intersections. The consideration of the signalized/unsignalized subsets can be important for practical use. Another choice was made between considering: the main and the secondary traffic volume (separated) or the total volume and the main-to- total volume ratio. The second alternative has generally led to a better goodness-of-fit.

Based on the overall models for intersections, crash frequencies increase with the total volume, the four-legged configuration, traffic signals, specialized turning lanes/cycle paths (first model in **Table 3)**, critical sight distance (alternative model in **Table 3**). Crash frequencies decrease with the main-to-total AADT ratio (thus the more the total AADT, the slower the crashes increase) and the bad pavement maintenance. The predictors can explain the crash frequency better than in the segments case, as based on the Nagelkerke R^2 values.

(Table 3 here)

For three-legged intersections, sight distance, turning lanes and cycle path crossing are confirmed as predictors (similar coefficients). The presence of traffic lights does not seem to be influential (except for traffic lights with dedicated turning lights). Moreover, more entering lanes (main road) results in a decrease of crashes ($p < 0.10$).

For four-legged intersections, sight distance and turning lanes are confirmed as significant predictors, while bicycle crossings are not. Traffic lights seem not influential, while the bad pavement maintenance is associated to a decrease in crashes.

For signalized intersections (highest R^2), four-legged intersections are comparatively less safe than three-legged intersections. As the critical sight distance increases, the crash frequency increases (similarly to three/four-legged intersections). Specialized turning lanes and bad pavement maintenance are confirmed, namely, with positive and negative coefficients. Bus stops close to signalized intersections are related to crashes decreasing.

For unsignalized intersections, other predictors result in crashes decreasing, besides of those already mentioned: median on the main road ($p < 0.10$) and sidewalks.

DISCUSSION

Subsets of road sites and associated predictors

The predictive models for urban segments were stratified into one-way and two-way models, since significant differences between these two conditions were found (overall model in **Table 2**), differently than Greibe (2003), who did not include the variable one-way/two-way. Two-way segments result as less safe compared to one-way segments, other conditions being equal; while significant differences between one-lane and multilane segments were not highlighted.

Moreover, since the variable “Type of lanes” was not included in the disaggregated models, it seems that the organization of one-way segments in one or more lanes does not seem influential on crash risk. However, on average, the sampled one-way roads are about 10 m wide, thus they could be practically two-lane operated, even if they are single-lane roads. In the two-way segments model, significant differences between one lane and more lanes per direction were not highlighted as well. However, in the overall segments model, the multilane two-way segments seem slightly less safe than two-way two-lanes segments ($e^{Typeoflanes3}/e^{Typeoflanes2}=1.20$; see Appendix for the model specification), coherently with AASHTO (2010).

On the other hand, the predictive models for urban intersections were stratified into signalized/unsignalized models and three-legged/four-legged models. In fact, significant differences were found between both categories. As expected from the high number of conflicts, four-legged intersections result as comparatively less safe than three-legged intersections, by a factor of $e^{0.443}=1.557$. The effect of signals on intersections is less clear, such as in Gomes et al. (2012), who developed three/four-legged models, in which the traffic signal variable was insignificant. However, very disaggregated intersection subsets were considered by Canale et al. (2005): three/four-legged no-control/stop-controlled, four-legged signalized intersections. In this study, differences in predictors by separating unsignalized from signalized intersections were found. Moreover, contrary to expectations, signalized intersections seem comparatively slightly less safe than unsignalized intersections (factor of $e^{0.268}=1.307$, alternative model), ceteris paribus. Four-legged intersections are comparatively less safe than three-legged intersections especially for unsignalized intersections, as expected. Note also that give-way/stop controlled three-legged intersections seems even less safe than no-control intersections, ceteris paribus. Specialized turning lanes seem consistently negative for safety, coherently with Gomes et al. (2012), in case of right turn on the major road. This could be explained by: a) the turning lanes variable being a surrogate measure for total conflicts, b) aberrant driving behaviours causing additional conflicts. Moreover, Canale et al. (2005), found mixed results for left/right turning lanes according to the intersection type.

Models for different subsets of urban intersections and segments are extremely useful for identifying predictors which are specifically only related to some subsets. For example, the increasing intersecting minor roads on segments are generally associated to crashes increasing (Greibe, 2003; Fancello et al., 2018). In this study, this variable is significant only for one-way segments, which is an important difference. In fact, one-way roads (especially if wide as in this dataset) may allow high speeds. Moreover, drivers should not care about other vehicles eventually crossing the travel direction from the other lane, such as on two-way undivided roads. Another difference relates to parking, which is generally associated to an increase in crashes with respect to rarely/prohibited parking (Greibe, 2003). In this study, this effect was found only for two-way segments (especially for parking on both sides, as expected due to the increased conflicts). Parking-related conflicts may even be more unexpected than in case of known minor roads (which seem less influential on two-way segments) and then drivers could not react in time. The traffic volume coefficient indicates a slower than linear increasing tendency for both subsets of segments; it is insignificant (close to zero) in case of two-way segments (for which the sample size is very limited). Hence, urban congestion seems more detrimental to the safety of one-way than of two-way segments. Moreover, note that the average segment length is included between 100 m and 200 m. This could explain the slower than linear increasing tendency of crashes with traffic, since several crashes on short segments may be influenced by the presence of intersections in case of high traffic volumes. Note that speed limits, road width and land use were not included in the segment models, differently than in Greibe (2003). However, note that speed limits are almost always equal to 50 km/h and land use is largely homogeneous in the central city area (**Figure 1**).

For what concerns intersections, the coefficients estimated for traffic volumes (and main-to-total ratio) are approximately similar between subsets. However, when the main-to-total AADT ratio increases by 10%, crashes increase with a factor of about 5 and 1.5, namely, for three-legged and four-legged intersections, ceteris paribus. Hence, especially for three-legged intersections, a severe unbalance between main and secondary volumes may increase crashes (see also Gomes et al., 2012). In fact, vehicles on the main road may not decelerate approaching to an intersection with a few entering vehicles. Moreover, in case of signalized intersections, drivers may even run red lights when few vehicles are entering from secondary roads. The usually considered predictor: median on the main road is significant ($p < 0.10$) and positive for safety, but only for unsignalized intersections. This is in line with previous results specific for three-legged intersections (Gomes et al., 2012), in particular stop-controlled (Canale et al., 2005). In fact, medians may help in channelizing the traffic flow. Moreover, road markings were previously found as negative for safety for three-legged stop-control intersections (Canale et al., 2005), such as here for two-way segments. Other predictors such as the lane balance, the number of intersecting one-way legs (Gomes et al., 2012), two-way operated major roads and the intersecting lane widths (Canale et al., 2005) were not confirmed here.

Assessment of additional variables

Some additional variables, usually not often considered for safety predictions, were considered in this study and some of them were actually included in the models. However, most of those present some unexpected trends, which may seem surprising at a first glance. For example, the presence of vertical signs on driveways/minor roads seem to be detrimental to safety. However, the number of driveways/minor roads provided with vertical signs in the sample is scarce (**Table 1**). Hence, this variable could be a surrogate measure for the driveway/minor road importance (i.e., considering the unlikely event that very-low volume driveways may have vertical signs); indicating that different driveways/minor roads may have variable effects on safety.

Maintenance-related variables (markings/pavement) are also worth to mention. It seems counterintuitive that deteriorated pavements may be positive for safety (one-way segments, four-legged/signalized intersections), as well as well-maintained markings (two-way segments). This could be explained by drivers being more cautious and driving at lower speeds on bad maintained pavements. However, a temporal displacement exists between the visual observations (mainly during 2018) and the crashes observation period (2012-2016). Hence, it is most likely that pavements (and markings) in good conditions in 2018 could have been resurfaced in the last years and vice versa. Thus, the estimated coefficient could also hide an opposite tendency.

The computed critical sight distance (the least value among all the intersecting legs) needs cautious interpretations as well. It was included in the models for three, four-legged and signalized intersections with similar positive coefficients. This can be explained by less cautious drivers (e.g., tending to speeding, see Colonna et al., 2016) when having more available sight distance, especially at signalized intersections. In fact, while sight distance is an important design pre-requisite; a longer sight distance could lead to a false sense of increased safety and possible aberrant behaviours, in specific cases such as running the red light. Manual explorations of crash circumstances at signalized intersections in the dataset seem to confirm this possibility.

Moreover, all the segments/intersections in the dataset were not originally designed with cycle paths/crossings, which were only recently implemented. Hence, having found that bicycle crossings are associated with crash increasing for three-legged and unsignalized intersections may indicate that such conflicts should be mitigated (especially at unsignalized intersections) e.g., by effective traffic calming measures, which are generally not present in the network studied. Moreover, bus stops close to signalized intersections seem to be positive for safety. This could be explained by drivers being forced to slow down for the combined presence of bus stops and intersections.

CONCLUSIONS

Safety performance functions for urban segments and intersections were estimated. The research aims of this study were dedicated to: a) explore possible subsets of segments and intersections for crash modelling, considering the predictors variability, b) find common trends and/or significant differences from the relevant literature, c) assess additional predictors often not considered for crash modelling.

The optimal subsets found were: one-way and two-way segments for the homogeneous segments; three-legged, four-legged, unsignalized, signalized intersections for the intersections. The division into three-legged/four-legged intersections seems the most effective, compared to the signalized/unsignalized division. Whereas, significant differences were not highlighted for the number of lanes on one-way and two-way segments. Predictors of intersection crashes share both commonalities and differences with similar studies (Gomes et al., 2012; Canale et al., 2005). Nevertheless, the segment model is largely different than the relevant reference study analysed (Greibe, 2003).

Some additional predictors often not included in prediction models were found as statistically significant. The effect of pavements/markings maintenance, critical sight distance at intersections, vertical signs on driveways/minor roads, cycle path crossings was discussed in detail. Their influence on crash predictions was demonstrated, even if requiring some additional explanations. These variables may be used in further urban safety studies.

The results shown in this study can be used for safety predictions in the same area in which data were collected. However, they could be generalized in other Italian/European cities with similar configurations, as based e.g., on some similarities shown with Portuguese models

(Gomes et al., 2012). However, the model transferability may provide a challenge. Besides practical aspects, this study provides new insights to overcome the problems and consequences of dividing urban intersections and segments into possible subsets and to increase the candidate crash predictors. Clearly, this study is based on a limited number of segments and intersections with a small number of crashes for some subsets (especially intersections), which may negatively influence crash predictions (Lord, 2006). Further data collection is currently in progress during the research project, which could help to enlarge datasets and to validate models/variables. Roundabouts were not considered since only a few roundabouts were present during the observation period. In some models, the explained variability of crash frequencies is someway small. Thus, there are several other variables which may be considered. A first attempt to enlarge them was conducted here, but further research is surely needed.

Acknowledgements. This work was supported by the City of Bari under Grant “Pa.S.S.S. research project” (Italian Ministry of Infrastructures and Transport).

Disclosure statement. No potential conflict of interest was reported by the authors.

Data availability statement. The data that support the findings of this study are available from the corresponding author, P. I., upon reasonable request.

References

American Association of State Highway and Transportation Officials (2010). Highway Safety Manual. 1st Edition. 2010.

Canale, S., Leonardi, S., & Pappalardo, G. (2005). The reliability of the urban road network: Crash forecast models. In *Proceedings of the III International Congress SIIV-People, Land, Environment and Transport Infrastructures, Bari, Italy*, (pp. 22-24.). 2005.

Colonna, P., Intini, P., Berloco, N., & Ranieri, V. (2016). The influence of memory on driving behavior: How route familiarity is related to speed choice. An on-road study. *Safety science*, *82*, 456-468. 2016.

Colonna, P., Intini, P., Berloco, N., & Ranieri, V. (2018). Integrated American-European protocol for safety interventions on existing two-lane rural roads. *European transport research review*, *10*(1), 5. 2018.

Colonna, P., Intini, P., Berloco, N., Fedele, V., Masi, G., & Ranieri, V. (2019). An Integrated Design Framework for Safety Interventions on Existing Urban Roads—Development and Case Study Application. *Safety*, *5*(1), 13. 2019.

Fancello, G., Soddu, S., & Fadda, P. (2018). A crash prediction model for urban road networks. *Journal of Transportation Safety & Security*, *10*(4), 387-405. 2018.

Farid, A., Abdel-Aty, M., Lee, J., Eluru, N., & Wang, J. H. (2016). Exploring the transferability of safety performance functions. *Accident Analysis & Prevention*, *94*, 143-152. 2016.

Geedipally, S. R., Shirazi, M., & Lord, D. (2017). Exploring the Need for Region-Specific Calibration Factors. *Transportation Research Record*, *2636*(1), 73-79.

Gomes, S. V., Geedipally, S. R., & Lord, D. (2012). Estimating the safety performance of urban intersections in Lisbon, Portugal. *Safety science*, *50*(9), 1732-1739. 2012.

Greibe, P. (2003). Crash prediction models for urban roads. *Accident Analysis & Prevention*, 35(2), 273-285. 2003.

Hauer, E., Council, F. M., & Mohammedshah, Y. (2004). Safety models for urban four-lane undivided road segments. *Transportation research record*, *1897*(1), 96-105. 2004.

Intini, P., Berloco, N., Binetti, R., Fonzone, A., Ranieri, V., & Colonna, P. (2019). Transferred versus local Safety Performance Functions: A geographical analysis considering two European case studies. *Safety Science*, *120*, 906-921.

Jonsson, T. (2005). *Predictive models for crashes on urban links-A focus on vulnerable road users*. Bulletin/Lund Institute of Technology, Department of Technology and Society, *226*. 2005.

Lee, J., Abdel-Aty, M., De Blasiis, M. R., Wang, X., & Mattei, I. (2019). International transferability of macro-level safety performance functions: a case study of the United States and Italy. *Transportation Safety and Environment*, 1(1), 68-78.

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 & Prevention*, *38*(4), 751-766. 2006.

Lord, D., & Mannering, F. (2010). 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.

Montella, A., Marzano, V., Mauriello, F., Vitillo, R., Fasanelli, R., Pernetti, M., & Galante, F. (2019). Development of Macro-Level Safety Performance Functions in the City of Naples. *Sustainability*, 11(7), 1871.

Persaud, B., & Dzbik, L. (1992). Crash prediction models for freeways. *Transportation Research Record*, *1401*, 55-60. 1992.

Sacchi, E., Persaud, B., & Bassani, M. (2012). Assessing international transferability of highway safety manual crash prediction algorithm and its components. *Transportation research record*, *2279*(1), 90-98. 2012.

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S*. Springer, New York, NY. 2002.

Vogt, A., & Bared, J. (1998). Crash models for two-lane rural segments and intersections. *Transportation Research Record*, *1635*(1), 18-29. 1998.

TABLE 1. Descriptive statistics of potential crash predictors for segments and intersections.

Variables [for segments]	Modalities	Homogeneous Segments (n = 320*)		One-way homogeneous segments (n = 190)		Two-way homogeneous segments (n = 130)	
		Mean (st. dev.)/ Counts	Min-max/ Percentages	Mean (st. dev.)/ Counts	Min-max/ Percentages	Mean (st. dev.)/ Counts	Min-max/ Percentages
Fatal+Injury (F+I) crashes	Numerical	689	-	283	-	406	-
F+I crashes/segment/year	Numerical	0.43 (0.86)	0-6	0.30 (0.68)	0-5	0.63 (1.03)	0-6
One-way	0 – No (two-way) 1 – Yes	0 – 130 1 – 190	0 – 0.41 1 – 0.59	0 – 0 1 – 190	0 – 0.00 1 – 1.00	0 – 130 1 – 0	0 – 1.00 1 – 0.00
One-lane	0 –No (multilane) 1 – Yes	0 – 226 1 – 94	0 – 0.71 1 – 0.29	0 – 96 1 – 94	0 – 0.51 1 – 0.49	0 – 130 1 – 0	0 – 1.00 1 – 0.00
AADT	Numerical	8637.2 (5889.3)	300.0-29960.0	6228.8 (4569.5)	300.0-22600.0	12157.1 (5843.5)	1160.0-29960.0
Length (m)	Numerical	138.3 (183.3)	30.0-1510.0	124.1 (175.4)	30.0-1510.0	159.1 (193.0)	30.0-1415.0
Speed Limit = 50 km/h	0 – Yes 1 – No (> 50 km/h)	0 – 309 1 – 11	0 – 0.97 1 – 0.03	0 – 190 1 – 0	0 – 1.00 1 – 0.00	0 – 119 1 – 11	0 – 0.92 1 – 0.08
Road width (m)	Numerical	10.5 (3.0)	4.1-28.0	9.9 (2.6)	5.5-20.0	11.3 (3.5)	4.1-28.0
Type of lanes	0 – 1 lane 1 – > 1 lane (up to 3) 2 – 1+1 lane 3 – > 1+1 lanes (up to 3+3)	0 – 116 1 – 74 2 – 89 3 – 41	0 – 0.36 1 – 0.23 2 – 0.28 3 – 0.13	0 – 116 1 – 74 2 – 0 3 – 0	0 – 0.61 1 – 0.39 2 – 0.00 3 – 0.00	0 – 0 1 – 0 2 – 89 3 – 41	0 – 0.00 1 – 0.00 2 – 0.68 3 – 0.32
Density of minor roads/ driveways per km	Numerical	26.9 (26.0)	0.0-143.0	22.4 (22.8)	0.0-142.9	33.3 (28.9)	0.0-143.0
Vertical signs on minor roads/ driveways	0 – No 1 – Yes	0 – 256 1 – 64	0 – 0.80 1 – 0.20	0 – 158 1 – 32	0 – 0.84 1 – 0.16	0 – 98 1 – 32	0 – 0.75 1 – 0.25
Bad maintenance of pavements	0 – No 1 – Yes	0 – 71 1 – 249	0 – 0.22 1 – 0.78	0 – 49 1 – 141	0 – 0.26 1 – 0.74	0 – 22 1 – 108	0 – 0.17 1 – 0.83

Road markings	0 – No/Partially absent or illegible 1 – Illegible 2 – Yes	0 – 60 1 – 124 2 – 136	0 – 0.19 1 – 0.39 2 – 0.42	0 – 24 1 – 82 2 – 84	0 – 0.13 1 – 0.43 2 – 0.44	0 – 36 1 – 42 2 – 52	0 – 0.28 1 – 0.32 2 – 0.40
Parking type	0 – Prohibited 1 – One side 2 – Both sides 3 – Mixed	0 – 85 1 – 109 2 – 119 3 – 7	0 – 0.27 1 – 0.34 2 – 0.37 3 – 0.02	0 – 26 1 – 74 2 – 87 3 – 3	0 – 0.14 1 – 0.39 2 – 0.46 3 – 0.02	0 – 59 1 – 35 2 – 32 3 – 4	0 – 0.45 1 – 0.27 2 – 0.25 3 – 0.03
Cycle paths	0 – No 1 – Yes	0 – 291 1 – 29	0 – 0.91 1 – 0.09	0 – 163 1 – 27	0 – 0.86 1 – 0.14	0 – 128 1 – 2	0 – 0.98 1 – 0.02
Sidewalks	0 – No 1 – Yes	0 – 31 1 – 289	0 – 0.10 1 – 0.90	0 – 4 1 – 186	0 – 0.02 1 – 0.98	0 – 27 1 – 103	0 – 0.21 1 – 0.79
Median	0 – No 1 – Yes	0 – 244 1 – 76	0 – 0.76 1 – 0.24	0 – 114 1 – 76	0 – 0.60 1 – 0.40	0 – 130 1 – 0	0 – 1.00 1 – 0.00
Bus stop	0 – No 1 – Yes	0 – 227 1 – 93	0 – 0.71 1 – 0.29	0 – 143 1 – 47	0 – 0.75 1 – 0.25	0 – 84 1 – 46	0 – 0.65 1 – 0.35
Bus/taxi lane	0 – No 1 – Yes	0 – 284 1 – 36	0 – 0.89 1 – 0.11	0 – 154 1 – 36	0 – 0.81 1 – 0.19	0 – 130 1 – 0	0 – 1.00 1 – 0.00
Land use	0 – Residential 1 – Mainly commercial 2 – Other	0 – 115 1 – 150 2 – 55	0 – 0.36 1 – 0.47 2 – 0.17	0 – 71 1 – 110 2 – 9	0 – 0.37 1 – 0.58 2 – 0.05	0 – 44 1 – 40 2 – 46	0 – 0.34 1 – 0.31 2 – 0.35
Variables [for intersections]	**Modalities**	**Intersections (n = 120*)**		**Three-legged intersections (n = 48)**		**Four-legged intersections (n = 72^)**	
		Mean (st. dev.)/ Counts	Min-max/ Percentages	Mean (st. dev.)/ Counts	Min-max/ Percentages	Mean (st. dev.)/ Counts	Min-max/ Percentages
Fatal+Injury (F+I) crashes	Numerical	906	-	232	-	674	-
F+I crashes/intersection/year	Numerical	1.51 (1.84)	0-15	0.97 (1.41)	0-15	1.87 (1.99)	0-12
Traffic lights	0 – No 1 – Yes	0 – 59 1 – 61	0 – 0.49 1 – 0.51	0 – 38 1 – 10	0 – 0.79 1 – 0.21	0 – 21 1 – 51	0 – 0.29 1 – 0.71
Type = 4 legs	0 – No 1 – Yes	0 – 48 1 – 72	0 – 0.40 1 – 0.60	0 – 48 1 – 0	0 – 1.00 1 – 0.00	0 – 0 1 – 72	0 – 0.00 1 – 1.00
Main AADT	Numerical	12192.4 (5313.2)	2475.0-23680.0	12588.4 (5350.8)	3550.0-22340.0	11928.4 (5309.1)	2475.0-23680.0

Secondary AADT	Numerical	5084.1 (4271.5)	300.0-17880.0	4008.3 (3897.4)	300.0-18520.0	5801.4 (4384.2)	345.0-17880.0
Total AADT	Numerical	17276.5 (7981.8)	4175.0-41560.0	16596.6 (7018.3)	5420.0-32300.0	17729.7 (8582.2)	4175.0-41560.0
Main AADT/ Total AADT	Numerical	0.53 (0.03)	0.50-0.62	0.55 (0.03)	0.50-0.62	0.53 (0.02)	0.50-0.61
Lane balance	0 – No 1 – Yes	0 – 81 1 – 39	0 – 0.67 1 – 0.33	0 – 38 1 – 10	0 – 0.79 1 – 0.21	0 – 43 1 – 29	0 – 0.60 1 – 0.40
Median on the main road	0 – No 1 – Yes	0 – 100 1 – 20	0 – 0.83 1 – 0.17	0 – 48 1 – 0	0 – 1.00 1 – 0.00	0 – 52 1 – 20	0 – 0.72 1 – 0.28
Median on the secondary road	0 – No 1 – Yes	0 – 107 1 – 13	0 – 0.89 1 – 0.11	0 – 43 1 – 5	0 – 0.90 1 – 0.10	0 – 64 1 – 8	0 – 0.89 1 – 0.11
Entering lanes on the main road	Numerical	1.5 (1.1)	0.0-5.0	1.1 (1.0)	0.0-4.0	1.8 (1.1)	0.0-5.0
Mean lane width at intersections (m)	Numerical	4.5 (1.1)	2.8 (8.7)	4.3 (1.0)	2.8-6.6	4.6 (1.2)	2.8-8.7
Critical sight distance (m)	Numerical	31.2 (26.0)	0.0-139.5	30.0 (28.1)	0.0-139.5	32.0 (24.6)	1.6-100.0
One-way legs	Numerical	1.7 (1.5)	0.0-4.0	1.4 (1.1)	0.0-3.0	2.0 (1.6)	0.0-4.0
Specialized turning lane	0 – No 1 – Yes	0 – 88 1 – 32	0 – 0.73 1 – 0.27	0 – 43 1 – 5	0 – 0.90 1 – 0.10	0 – 45 1 – 27	0 – 0.63 1 – 0.37
Traffic control	0 – No 1 – Give-way/Stop 2 – Traffic lights 3 – Traffic lights with dedicated turning	0 – 22 1 – 37 2 – 35 3 – 26	0 – 0.18 1 – 0.31 2 – 0.29 3 – 0.22	0 – 16 1 – 22 2 – 7 3 – 3	0 – 0.33 1 – 0.46 2 – 0.15 3 – 0.06	0 – 6 1 – 15 2 – 28 3 – 23	0 – 0.08 1 – 0.21 2 – 0.39 3 – 0.32
Bad maintenance of pavements	0 – No 1 – Yes	0 – 18 1 – 102	0 – 0.15 1 – 0.85	0 – 9 1 – 39	0 – 0.19 1 – 0.81	0 – 9 1 – 63	0 – 0.13 1 – 0.87
Cycle path crossing	0 – No 1 – Yes	0 – 103 1 – 17	0 – 0.86 1 – 0.14	0 – 36 1 – 12	0 – 0.75 1 – 0.25	0 – 67 1 – 5	0 – 0.93 1 – 0.07
Sidewalks	0 – No 1 – Yes (both sides)	0 – 8 1 – 112	0 – 0.07 1 – 0.93	0 – 3 1 – 45	0 – 0.06 1 – 0.94	0 – 5 1 – 67	0 – 0.07 1 – 0.93
Bus stops	0 – No 1 – Yes	0 – 111 1 – 9	0 – 0.93 1 – 0.07	0 – 46 1 – 2	0 – 0.96 1 – 0.04	0 – 65 1 – 7	0 – 0.90 1 – 0.10
Bus/taxi lanes on intersecting roads	0 – No 1 – Yes	0 – 92 1 – 28	0 – 0.77 1 – 0.23	0 – 43 1 – 5	0 – 0.90 1 – 0.10	0 – 49 1 – 23	0 – 0.68 1 – 0.32

*Note: Five segments were discharged from the initial dataset due to segment length minor than 30 m, which were deemed as irrelevant for safety modelling purposes. Two intersections were discharged from the initial dataset due to possible errors in the counts of traffic volumes, leading to unrealistic data.
^These intersections include one five-legs intersection.

TABLE 2. Selected models for segments

Predictor	Estimate	Std. Error	z-value	*p*-value
All segments				
(Intercept)	-6.054	0.728	-8.318	< 0.001
Ln(AADT)	**0.375**	**0.084**	**4.449**	**< 0.001**
Ln(L)	**0.292**	**0.060**	**4.907**	**< 0.001**
Type of lanes 1: >1 (reference: 1 lane)	0.079	0.167	0.470	0.638
Type of lanes 2: 1+1 (ref.: 1 lane)	**0.524**	**0.151**	**3.471**	**< 0.001**
Type of lanes 3: > 1+1 (ref.: 1 lane)	**0.707**	**0.181**	**3.906**	**< 0.001**
Vertical signs on driveways/minor roads (Yes - ref.: No)	**0.652**	**0.110**	**5.943**	**< 0.001**
Bad maintenance of pavements (Yes - ref.: No)	**-0.268**	**0.114**	**-2.348**	**0.019**
Parking type 1: At one side (ref.: No parking)	*0.218*	*0.131*	*1.668*	*0.095*
Parking type 2: At both sides (ref.: No parking)	**0.258**	**0.126**	**2.053**	**0.040**
Parking type 3: Mixed (ref.: No parking)	**-1.092**	**0.556**	**-1.964**	**0.050**
Cycle paths (Yes – ref.: No)	*0.308*	*0.182*	*1.694*	*0.090*
Goodness-of-fit measures				
LRT (reference: null model) = 200.96 (df = 11), p <.001; LRT (reference: full model) = 6.64 (df = 11), p = .828				
Nagelkerke R^2 = 0.158, AIC = 2609.9				
Over-dispersion parameter (theta) = 1.08 (std. error: 0.17)				
Interval of continuous predictors				
AADT: 300-29960 vehicles/day, L: 30-1510 m				
One-way segments				
(Intercept)	-8.794	0.961	-9.151	< 0.001
Ln(AADT)	**0.695**	**0.108**	**6.434**	**< 0.001**
Ln(L)	**0.333**	**0.095**	**3.505**	**< 0.001**
Driveways/minor roads per km	**0.010**	**0.003**	**3.163**	**0.002**
Vertical signs on accesses/minor roads (Yes - ref.: No)	**0.477**	**0.169**	**2.825**	**0.005**

Bad maintenance of pavements (Yes - ref.: No)	**-0.497**	**0.159**	**-3.137**	**0.002**
Goodness-of-fit measures				
LRT (reference: null model) = 101.68 (df = 5), p <.001; LRT (reference: full model) = 9.77 (df = 10), p = .460				
Nagelkerke R^2 = 0.153, AIC = 1226.8				
Over-dispersion parameter (theta) = 1.01 (std. error: 0.26)				
Interval of continuous predictors				
AADT: 300-22600 vehicles/day, L: 30-1510 m, Density of driveways/minor roads: 0.0-142.9 number/km				
Two-way segments				
(Intercept)	-3.764	1.122	-3.356	< 0.001
Ln(AADT)	0.134	0.117	1.140	0.254
Ln(L)	**0.310**	**0.079**	**3.944**	**< 0.001**
Vertical signs on accesses/minor roads (Yes - ref.: No)	**0.432**	**0.141**	**3.074**	**0.002**
Markings 1: Illegible (ref.: No/partially absent/illegible)	*0.317*	*0.177*	*1.791*	*0.073*
Markings 2: Yes (ref.: No/partially absent/illegible)	*0.309*	*0.175*	*1.767*	*0.077*
Parking type 1: At one side (ref.: No parking)	*0.287*	*0.156*	*1.836*	*0.066*
Parking type 2: At both sides (ref.: No parking)	**0.441**	**0.152**	**2.903**	**0.004**
Parking type 3: Mixed (ref.: No parking)	-0.973	0.640	-1.521	0.128
Goodness-of-fit measures				
LRT (reference: null model) = 57.69 (df = 7), p < 0.001; LRT (reference: full model) = 5.63 (df = 7), p = 0.584				
Nagelkerke R^2 = 0.102, AIC = 1369.2				
Over-dispersion parameter (theta) = 1.32 (std. error: 0.29)				
Interval of continuous predictors				
AADT: 1160-29960 vehicles/day, L: 30-1415 m				

TABLE 3. Selected models for intersections

Predictor	Estimate	Std. Error	z-value	*p*-value
All intersections				
(Intercept)	0.930	1.524	0.610	0.542
Ln(Total AADT)	**0.285**	**0.106**	**2.675**	**0.007**
Main AADT/Total AADT	**-6.870**	**2.021**	**-3.399**	**0.001**
Type: Four-legged (reference: Three-legged)	**0.443**	**0.105**	**4.210**	**< 0.001**
Critical sight distance	**0.005**	**0.002**	**2.762**	**0.006**
Specialized turning lane (Yes – ref.: no)	**0.496**	**0.114**	**4.362**	**< 0.001**
Bad maintenance of pavements (Yes - ref.: No)	**-0.397**	**0.115**	**-3.444**	**< 0.001**
Cycle path crossing (Yes – ref.: No)	**0.327**	**0.122**	**2.677**	**0.007**
Goodness-of-fit measures				
LRT (reference: null model) = 5.63 (df = 12), p = .934; LRT (reference: full model) = 161.11 (df = 12), p <.001				
Nagelkerke R^2 = 0.283, AIC = 1877.6				
Over-dispersion parameter (theta) = 2.84 (std. error: 0.52)				
Interval of continuous predictors				
Total AADT: 4175-41560 vehicles/day, Main/Total AADT: 0.50-0.62, Critical sight distance: 0.0-139.5 m				
Three-legged intersections				
(Intercept)	4.696	2.895	1.609	0.108
Ln(Total AADT)	*0.366*	*0.220*	*1.661*	*0.097*
Main AADT/Total AADT	**-16.376**	**3.711**	**-4.413**	**<0.001**
Number of entering lanes on the main road	*-0.211*	*0.119*	*-1.768*	*0.077*
Critical sight distance	**0.006**	**0.003**	**2.009**	**<0.001**
Specialized turning lane (Yes – ref.: no)	**0.927**	**0.228**	**4.061**	**<0.001**
*Traffic control 1: give-way/stop (ref.: no control)**	*0.359*	*0.214*	*1.675*	*0.094*
Traffic control 2: traffic lights (ref.: no control)*	0.377	0.289	1.306	0.192
*Traffic control 3: main+turning lights (ref.: no control)**	*0.556*	*0.313*	*1.777*	*0.076*

Cycle path crossing (Yes – ref.: No)	**0.900**	**0.235**	**3.830**	**<0.001**
Goodness-of-fit measures				
LRT (reference: null model) = 63.39 (df = 9), p <.001; LRT (reference: full model) = 5.79 (df = 9), p = .761				
Nagelkerke R^2 = 0.292, AIC = 610.4				
Over-dispersion parameter (theta) = 5.26 (std. error: 2.75)				
Interval of continuous predictors				
Total AADT: 5420-32300 vehicles/day, Main /Total AADT: 0.50-0.62, Critical sight distance: 0.0-139.5 m, Number of entering lanes on the main road: 0-4				
Four-legged intersections				
(Intercept)	-0.571	1.823	-0.313	0.754
Ln(Total AADT)	**0.381**	**0.131**	**2.905**	**0.004**
Main AADT/Total AADT	*-4.623*	*2.620*	*-1.764*	*0.078*
Critical sight distance	**0.005**	**0.002**	**2.263**	**0.024**
Specialized turning lane (Yes – ref.: no)	**0.418**	**0.140**	**2.981**	**0.003**
Bad maintenance of pavements (Yes - ref.: No)	**-0.529**	**0.152**	**-3.476**	**0.001**
LRT (reference: null model) = 77.98 (df = 5), p <.001; LRT (reference: full model) = 8.92 (df = 11), p = .630				
Nagelkerke R^2 = 0.225, AIC = 1266.7				
Over-dispersion parameter (theta) = 2.71 (std. error: 0.56)				
Interval of continuous predictors				
Total AADT: 4175-41560 vehicles/day, Main /Total AADT: 0.50-0.61, Critical sight distance: 1.6-100.0 m				
All intersections (alternative)				
(Intercept)	1.713	1.513	1.132	0.257
Ln(Total AADT)	**0.371**	**0.103**	**3.613**	**< 0.001**
Main AADT/Total AADT	**-9.502**	**2.161**	**-4.397**	**< 0.001**
Median on the main road (Yes – reference: No)	*0.226*	*0.117*	*1.930*	*0.054*
Critical sight distance	**0.004**	**0.002**	**2.359**	**0.018**
Type of intersection (Signalized – ref.: Unsignalized)	**0.268**	**0.104**	**2.580**	**0.010**
Bad maintenance of pavements (Yes - ref.: No)	**-0.321**	**0.118**	**-2.730**	**0.006**

Cycle path crossing (Yes - ref.: No)	**0.253**	**0.123**	**2.058**	**0.040**
Goodness-of-fit measures				
LRT (reference: null model) = 124.59 (df = 7), p <.001; LRT (reference: full model) = 8.29 (df = 8), p = .406				
Nagelkerke R^2 = 0.222, AIC = 1914.1				
Over-dispersion parameter (theta) = 2.40 (std. error: 0.41)				
Interval of continuous predictors				
Total AADT: 4175-41560 vehicles/day, Main/Total AADT: 0.50-0.62, Critical sight distance: 0.0-139.5 m				
Signalized intersections				
(Intercept)	2.390	2.293	1.043	0.297
Ln(Total AADT)	**0.407**	**0.142**	**2.863**	**0.004**
Main AADT/Total AADT	**-11.302**	**3.586**	**-3.152**	**0.002**
Type: Four-legged (reference: Three-legged)	**0.452**	**0.182**	**2.479**	**0.013**
Critical sight distance	**0.004**	**0.002**	**1.986**	**0.047**
Specialized turning lane (Yes – ref.: no)	**0.307**	**0.148**	**2.076**	**0.038**
Bad maintenance of pavements (Yes - ref.: No)	**-0.630**	**0.148**	**-4.243**	**< 0.001**
Presence of bus stops (Yes – ref.: No)	**-0.914**	**0.357**	**-2.559**	**0.010**
Goodness-of-fit measures				
LRT (reference: null model) = 95.60 (df = 7), p < 0.001; LRT (reference: full model) = 4.39 (df = 9), p = 0.884				
Nagelkerke R^2 = 0.320, AIC = 1078.2				
Over-dispersion parameter (theta) = 3.08 (std. error: 0.71)				
Interval of continuous predictors				
Total AADT: 5800-41560 vehicles/day, Main/Total AADT: 0.50-0.61, Critical sight distance: 1.6-139.5 m				
Unsignalized intersections				
(Intercept)	-2.604	2.185	-1.192	0.233
Ln(Total AADT)	**0.759**	**0.163**	**6.343**	**< 0.001**
Main AADT/Total AADT	**-7.868**	**0.173**	**4.389**	**0.001**
Type: Four-legged (reference: Three-legged)	**1.036**	**2.406**	**-3.270**	**< 0.001**
Median on the main road (Yes – reference: No)	**-0.998**	**0.329**	**-3.031**	**0.002**

Specialized turning lane (Yes – ref.: No)	**0.516**	**0.204**	**2.523**	**0.012**
Cycle path crossing (Yes - ref.: No)	**0.997**	**0.193**	**5.156**	**< 0.001**
Presence of sidewalks (Yes - ref.: No)	**-1.008**	**0.331**	**-3.042**	**0.002**
LRT (reference: null model) = 65.32 (df = 7), p <.001; LRT (reference: full model) = 10.46 (df = 11), p =.489				
Nagelkerke R^2 = 0.239, AIC = 769.6				
Over-dispersion parameter (theta) = 8.11 (std. error: 6.16)				
Interval of continuous predictors				
Total AADT: 4175-31500 vehicles/day, Main/Total AADT: 0.50-0.62, Critical sight distance: 0.0-100.0 m				

*Note: The variable traffic control device was further disaggregated into: no control, give-way, stop control for unsignalized intersections, and into: traffic lights, traffic lights with dedicated left turning lights, traffic lights with dedicated right turning lights for signalized intersections.

FIGURE 1. Definition of the study area: City of Bari, Italy

Legend. On the left: localization of the city of Bari within Italy. On the upper right: limit of the urban area studied (the main urban centre inside the City of Bari), highlighted through the dashed line. On the lower right: zoom of the area studied within the City of Bari, Italy.

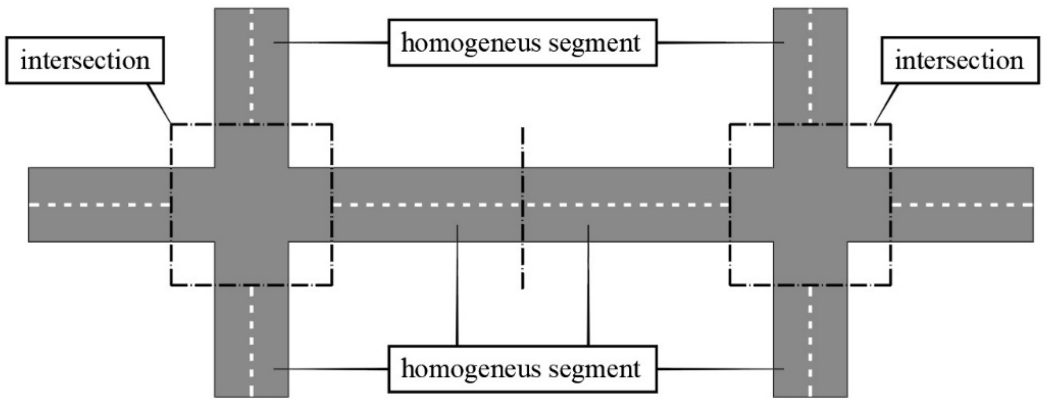

FIGURE 2. Division of the road network into sites: intersections and homogeneous road segments

APPENDIX

The general SPF model structures used are reported as follows:

$$N_{SPF,segments} = e^{\beta 0,S} * AADT^{\beta 1,S} * L^{\beta 2,S} * e^{\sum_{i=3}^{n} \beta i,S \, Xi,S} \tag{1}$$

$$N_{SPF,intersections} = e^{\beta 0,I} * (AADT_{maj} + AADT_{min})^{\beta 1,I} * e^{\sum_{i=2}^{n} \beta i,I \, Xi,I} \tag{2}$$

Where:

AADT = Annual Average Daily Traffic for segments;

$AADT_{maj}$ = AADT for the major intersecting road (carrying the highest amount of traffic);

$AADT_{min}$ = AADT for the minor intersecting road (carrying the lowest amount of traffic). Note that attempts at estimating separate coefficients for the major and minor traffic volumes were made, which however indicates the functional form in Eq. 2 as the most appropriate for the dataset;

L = segment length (m);

$X_{i,S}$ = other predictors for segments (numerical or categorical, in case of categorical variables they are transformed into binary dummy variables with modalities 0 and 1, 0: reference modality);

$X_{i,I}$ = other predictors for intersections (numerical or categorical, in case of categorical variables they are transformed into binary dummy variables with modalities 0 and 1, 0: reference modality);

$\beta_{i,S}$ = estimate of the coefficients associated to each crash predictor for segments through maximum likelihood estimation ($\beta_{0,S}$ is the estimate for the intercept).

$\beta_{i,I}$ = estimate of the coefficients associated to each crash predictor for intersections through maximum likelihood estimation ($\beta_{0,I}$ is the estimate for the intercept).