TRANSPORT PROBLEMS 2016 Volume 11 Issue 3
PROBLEMY TRANSPORTU DOI: 10.20858/tp.2016.11.3.4

road accidents; land use features; traffic analysis zone; spatial analysis; accidents estimation model

Serban RAICU, Dorinela COSTESCU*,
Stefan BURCIU, Florin RUSCA, Mircea ROSCA
University "Politehnica" of Bucharest
Splaiul Independentei 313, Bucharest 060042, Romania
**Corresponding author*. E-mail: dorinela.costescu@upb.ro

ROAD ACCIDENT ESTIMATION MODEL IN URBAN AREAS

Summary. Urban areas are significantly different in terms of traffic risk. In a decisive manner, they are the result of urban development policies. The shape, size and configuration of an entire urban area, the facilities to satisfy people and the need for mobility of goods, as well as behavioural attitudes of the population, are essential for a traffic pattern and its associated risks. In this framework, the purpose of this paper is to identify the effects of urban area characteristics on road accidents. Using specific spatial analysis, a model of accident estimation in the urban areas of Bucharest is developed. The study aims to provide useful tools for urban decision makers for a-priori analysis of the consequences of urban outline changes on traffic risks.

MODÈLE D'ESTIMATION DES ACCIDENTS DE LA ROUTE DANS LES ZONES URBAINES

Résumé. Les zones urbaines sont très différentes en termes de risque associé à la circulation. En manière décisive, ils sont le résultat des stratégies du développement urbain. La forme, la dimension et la configuration de l'ensemble de la région urbaine, l'offre pour satisfaire les besoins de mobilité des personnes et des marchandises, ainsi que les attitudes comportementales de la population sont essentielles pour le modèle de la circulation et le risque routier. Dans ce contexte, l'objectif de cet article est d'identifier les effets des caractéristiques des zones urbaines sur les accidents de la route. L'étude vise à fournir des décideurs dans la planification urbaine des instruments nécessaires pour analyser apriori les conséquences des changements de la structure urbaine sur le risque routier.

1. INTRODUCTION

From a phenomenological perspective, the urban area where traffic risk occurs is a complex system. Four main classes of inputs could be considered significant for risk circumstances [1]:

1) *Objects class*—defines the categories of elements, such as mobile entities (vehicles, cyclists, pedestrians, etc.) or features of technical and road urban infrastructure, which intervene in risk situations;
2) *Actors class*—defines different groups or organizations which could influence the system: urban authorities, urban network management companies, civil associations, population;
3) *Spatial structure class*—describes the topological configuration of the different elements and phenomena related to risk;

4) *Temporal structure class*—defines the position of system entities in time, intensity of daily/seasonal movement, etc.

The interactions between these classes generate a traffic flow characterized by the uncertainties related to traffic risk. In the last two decades, all four classes have suffered drastic transformations in Bucharest, the study area of the research presented in this paper. The rapid and radical changes of socioeconomic life of the Bucharest population, major changes in urban commerce structure caused by the emergence of large commercial centres, new concentration of urban and suburban residential areas, spatial and structural changes of places of interest (for work, education, leisure, etc.) have led to changes in the size and pattern of urban traffic flow. Consequently, the traffic risk exposure has increased. Understanding the road accident causes, identifying appropriate solutions and applying them since the urban planning phase can lead to traffic safety improvement [2, 3]. In this frame, the aim of our research is to identify the relationships between land use and road safety performances and to develop a road accident estimation model at urban zones level.

Several models are developed to estimate the number of accidents [4]. Most of the developed models include accident prediction function for different classes of junctions and road segments [5 - 11]. Reduced number of models have been elaborated for emphasizing the influence of the urban zone characteristics on accident occurrence. There are models developed for identifying the relationships between road accidents and compound road environments [12], for street network [13] and for land use [2, 14]. This paper presents the case study of the model developed to estimate accident function of the urban area attributes in Bucharest. Due to the peculiarities of Bucharest urban zones and the constraints given by the available data, new examination procedures have been necessary. The initial analysis has been made at the land use zone level, but unsatisfactory statistical measures resulted. Consequently, the analysis level has to be changed and adequate statistical measures have been obtained for the accident estimation model developed at the traffic analysis zones level.

2. STUDY AREA AND MODELLING SETTINGS

Our analysis of the relationships between road accidents and land use in Bucharest started from the following input data:

- Urban road network geo-database, which includes attributes of junction and road sections;
- Statistics of road accidents recorded in Bucharest during 2008 – 2012; the records offered by the Bucharest Road Police Department, including severe road accident with injuries and/or significant financial damages (Fig. 1), were edited in the geo-database and the data set with accident location on urban road network was obtained (Fig. 2). The GIS procedures were applied to obtain classes of different accident types (vehicle-vehicle accidents and vehicle-pedestrian accidents) on categories of road network features (junctions and road sections). The sample dimension of each accident class was too small to allow proper calibration of the estimation model [15]. Consequently, we have processed and calibrated the model for the aggregate set of all severe accidents recorded in Bucharest.
- Land use zones geo-database, whose available data includes the area of Bucharest city (around 228 sq. km with 1,890 mill. inhabitants) divided in N_{LU} = 9675 zones for 12 land uses (Tab. 1).

Because our aim is to identify the influence of land use on traffic risk, we eliminated from the land use dataset the areas of road infrastructure. Even if the accidents are located on road network features, the influence of land uses of the areas located around the accident location is analysed. After this step N_{LU}^* = 7650 land use zones are selected for further analysis and the recorded accidents are assigned to adjacent land use areas.

The land use and accident date sets are processed in the GIS environment to obtain the accident location distribution in urban zones. The total number of accidents A_i assigned to zone Z_i, $i=1..N_{LU}^*$ is calculated by the equation:

$$A_i = \sum_{I_k \subset Z_i} A_{I_k} + \sum_{XY \subset Z_i} A_{XY} \tag{1}$$

where A_{I_k} is the number of accidents that occurred at junction I_k located in zone Z_i and A_{XY} is the number accidents that occurred at section XY located in zone Z_i.

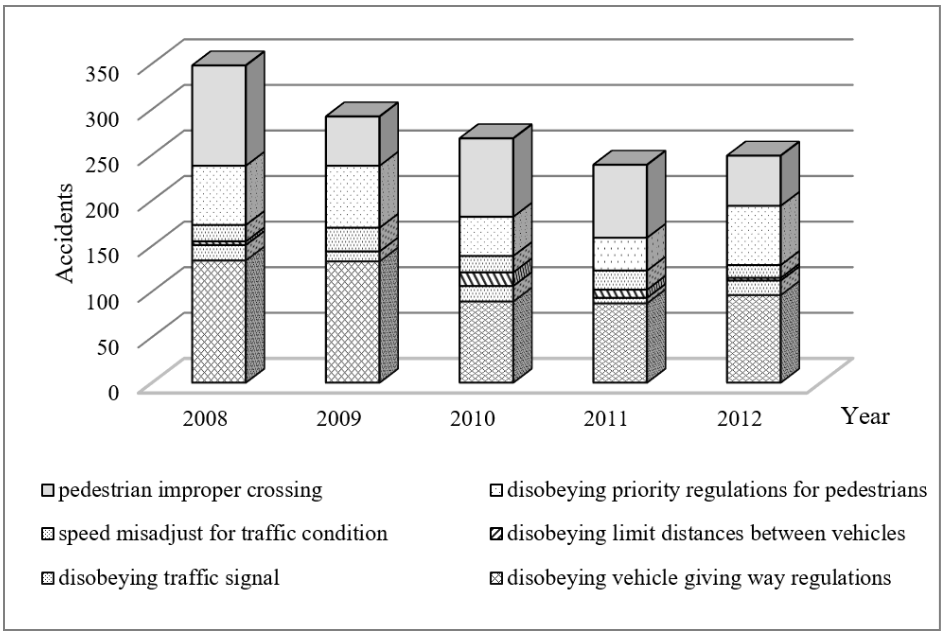

Fig. 1. Statistics of severe road accidents in Bucharest (2008-2012)
Fig. 1. Statistiques des accidents graves dans Bucarest (2008-2012)

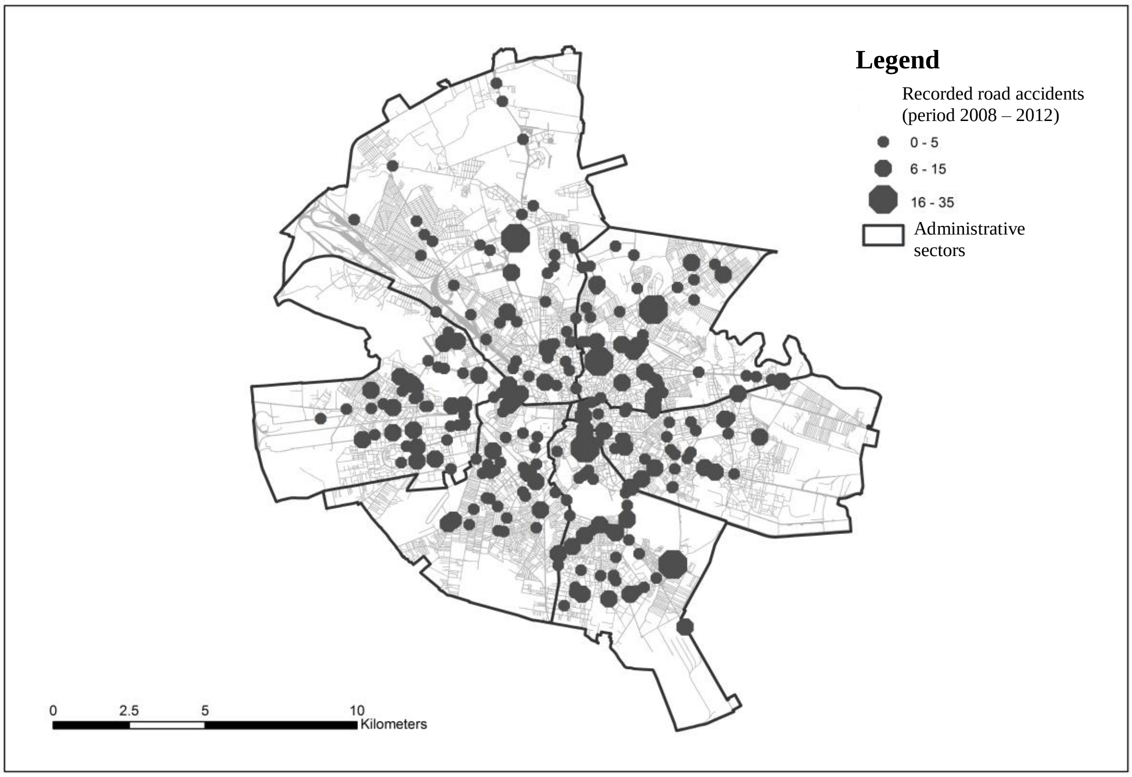

Fig. 2. Case study area of Bucharest city: Spatial distribution of recorded accidents on urban road network
Fig. 2. La zone d'étude de cas de Bucarest: la distribution spatiale des accidents enregistrés sur le réseau routier urbain

The results (Fig. 3) show that a significant rate of accidents is located in areas with a high density of buildings (88.71%), followed by areas with a low density of buildings (4.24%), educational use (3.33%) and commercial zones (1.69%).

The land use of each urban zone has a direct influence on the generated and attracted traffic flow and also on the activities and socioeconomic characteristics of the neighbouring zones. Therefore, it is necessary to evaluate the spatial autocorrelation of accident distribution before analysing the relationships between land use and traffic risk. The spatial autocorrelation was assessed using Moran's I Index [16, 17] computed function of the number of accidents assigned to urban zones. We obtained the value of 0.39 for Moran's I Index with a *z*-score of 264, which indicates a clustered pattern of accidents and spatial relationships between accidents located in adjacent zones. Hence, the accident analysis at this level of urban zonation is not suitable and we had to choose another level of analysis.

Table 1

Land use classification

i	Land use code	Description
1	High-density buildings development (HDU)	Areas with residential, business and retail uses, with high density of buildings
2	Low-density buildings development (LDU)	Areas with residential and retail uses, with low density of buildings
3	Governmental (GU)	Areas with governmental and administrative use
4	Commercial (COM)	Shopping centres, malls, recreational areas
5	Road infrastructure (TU)	Area of road infrastructure: streets, squares, parking
6	Passenger transport terminals (TTU)	Areas of airports, railway stations, regional bus stations
7	Industrial (IND)	Areas with industrial uses; distribution centres and freight transport terminals
8	Educational (EDU)	Areas with universities, schools, kindergartens
9	Cultural (CU)	Areas with museums, theatres, cultural centres
10	Parks (PKU)	Parks, entertainment areas, play fields
11	Stadiums (SU)	Areas with stadiums
12	Green land use (GRU)	Forests, green areas, greenhouses, farmland

We decided to analyse the accident distribution in the traffic analysis zone (TAZ) level for N_Z =80 zones (Fig. 4) defined by the Bucharest General Master Plan of Transport [18]. We applied the previous procedure (eq. 1) to assign accidents to the analysed zones and we again evaluated the spatial autocorrelation of the accident distribution on the new level of zoning. In this case we obtained the values 0.009 for Moran's I Index (very low) with a *z*-score of 0.68 and a *p*-value of 0.49, which indicate a random accidents distribution and the fact that the accidents located in a zone do not influence the accidents in adjacent zones. Therefore, the obtained data sets are appropriate to assess the relationships between urban zone characteristics and traffic risk.

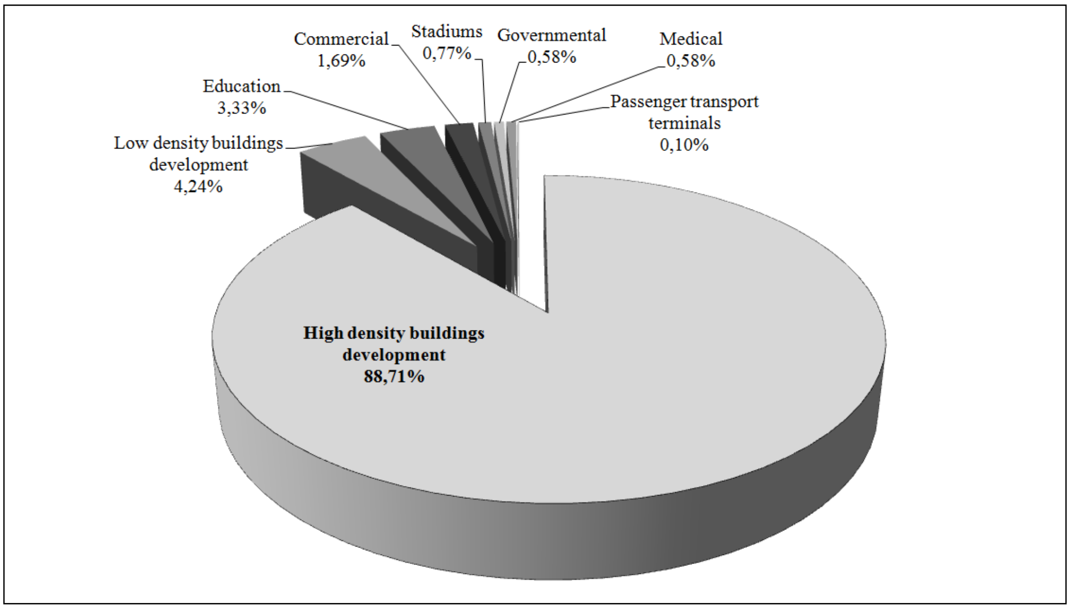

Fig. 3. Road accident percentage by land use in Bucharest during 2008 – 2012
Fig. 3. Le pourcentage d' accidents de la circulation par les fonctions urbaines à Bucarest entre 2008 - 2012

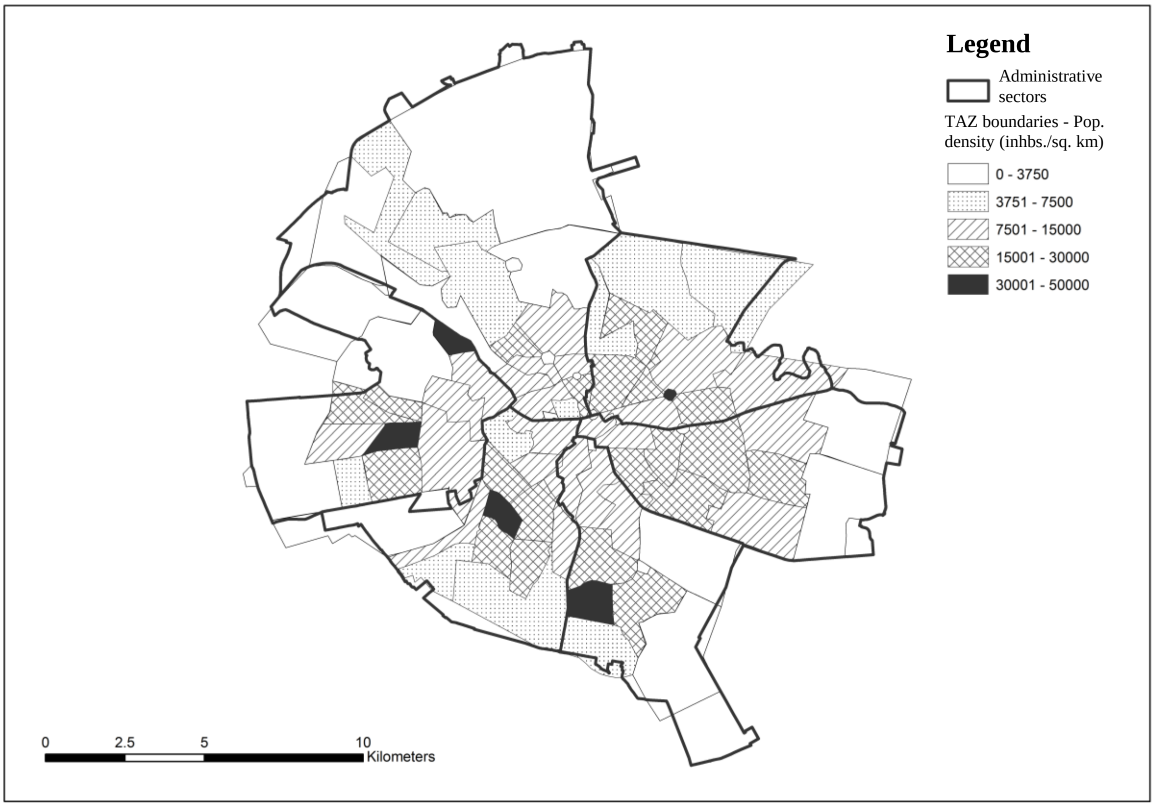

Fig. 4. Analysis units in the Bucharest city area
Fig. 4. Unités d'analyse dans la région de la ville de Bucarest

3. MODEL DEVELOPMENT

GIS spatial analysis procedures are applied on datasets of TAZs (including accident distribution), urban road network and land use to obtain features and attributes corresponding to accident analysis on selected urban zones (Fig. 5).

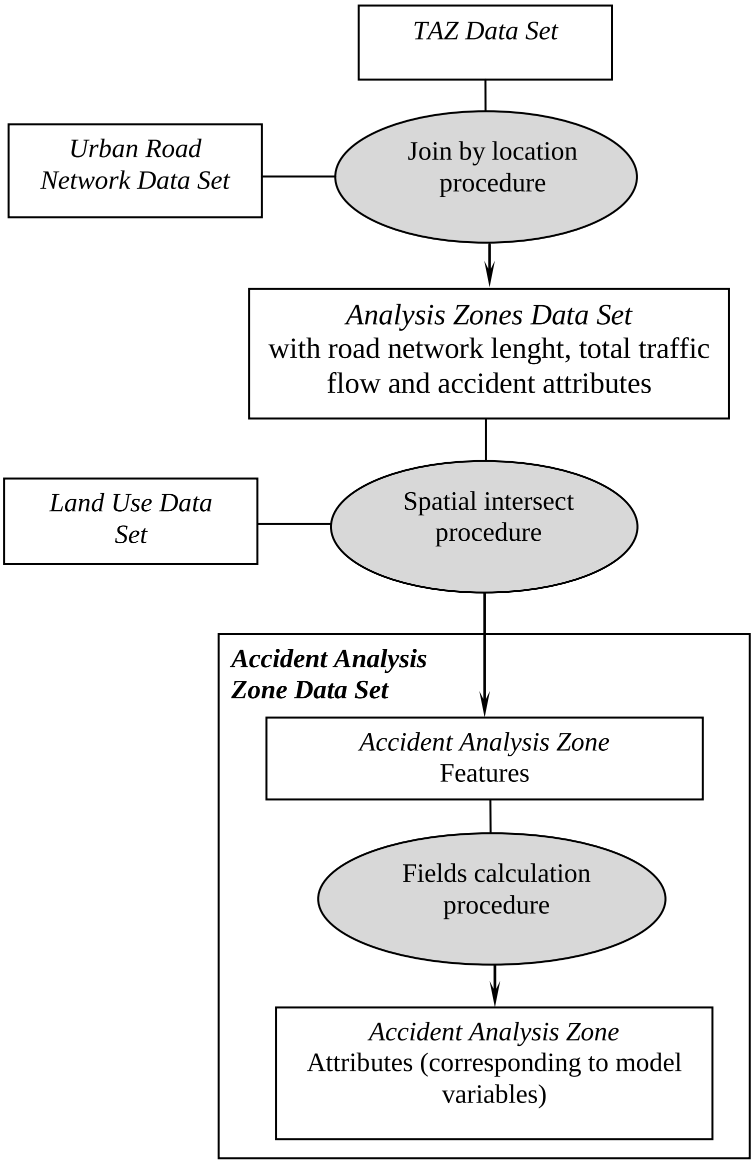

Fig. 5. Flow diagram of model data preparation process
Fig. 5. Schéma du procédé de préparation des données de modèle

The resulting data set allowed for the calculation of the variables Y_i, $i=1..12$ (Tab. 2) on the analysis zones used to estimate the relationships between urban characteristics and road accidents (the number of accidents represents the dependent variable, denoted by D_v). We considered the sum of the accidents on urban zones during 5 years in order to use a more representative data sample. Table 2 summarizes the minimum, maximum and mean values of urban characteristics for N_Z = 80 zones.

Table 2

Statistical measures of model zones characteristics

i	Variable Y_i	Description	No. of analysis zones N_Z	Average	Minimum	Maximum	Standard deviation
1	S_z	Area [sq. km]	80	3.2	0.5	28.5	3.7
2	P_z	Population [inhbs.]	80	23500	0	83220	20035
3	D_z	Population density [inhbs./sq. km]	80	11750	0	49950	10450
4	L_R	Road network length [km]	80	35.0	1.5	137.0	23.1
5	S_{TU}	Area of road infrastructure [sq. km]	80	1.4	0.3	3.12	0.65
6	D_{TU}	Density of road infrastructure [sq. km/ sq. km]	80	1.9	0.07	19.5	3.95
7	S_{HDU}	Area of high density building development use zones [sq. km]	80	1.3	0	3.5	0.8
8	S_{LDU}	Area of low density building development zones [sq. km]	80	0.2	0	3.8	0.5
9	S_{EDU}	Area of educational zones [sq. km]	80	0.06	0	0.90	0.16
10	S_{COM}	Area of commercial zones [sq. km]	80	0.16	0	0.55	0.13
11	P_{HDU}	Population of high density building development zones [inhbs.]	80	18345	0	70250	15760
D_v	A	Accidents/5 years	80	17	0	68	16.8

After several tests of correlation applied to different subsets of two up to five variables of the set $\{Y_i \mid i=1..12\}$ using correlation coefficient R^2, adjusted correlation coefficient R^2, Akaike information criterion *AICc* and variance inflation factor *VIF* as evaluators [16, 17], we identified two classes of significant variables in road accidents:

I. Population of high-density building development zone *(P_{HDU})*, area of road infrastructure *(S_{TU})*, area of educational zones, *(S_{EDU})* and area of commercial zones *(S_{COM})*; based on the ratio between average and variance of these variables [3, 19] used in the following equations to accident estimation:

$$\hat{A}(j) = \mu \cdot P_{HDU,j}^{\alpha} \cdot e^{\beta \cdot S_{TU,j}} \tag{2}$$

$$\hat{A}(j) = \mu \cdot P_{HDU,j}^{\alpha} \cdot S_{TU,j}^{\beta} \cdot e^{\gamma_1 \cdot S_{EDU,j} + \gamma_2 \cdot S_{COM,j}} \tag{3}$$

$$\hat{A}(j) = \mu \cdot e^{\sum_{k=1}^{4} \alpha_k \cdot X_{k,j}} , \tag{4}$$

with $j = 1..N_Z$, $X_k \in \{P_{HDU}, S_{TU}, S_{EDU}, S_{COM}\}$ and $\mu,\ \alpha,\ \beta,\ \gamma_1\ ,\gamma_2\ ,\alpha_k$ parameters to be calibrated.

II. Density of population of HDU zones (denoted by D_{HDU} and computed by $D_{HDU} = P_{HDU} / S_{HDU}$), density of road infrastructure area $(D_{TU} = S_{TU} / S_Z)$, density of educational zones, $(D_{EDU} = S_{EDU} / S_Z)$ and density of commercial zones $(D_{COM} = S_{COM} / S_Z)$ used in equations:

$$\hat{A}(j) = \mu \cdot S_{Z,j} \cdot D_{HDU,j}^{\alpha} \cdot e^{\beta \cdot D_{TU,j}} \tag{5}$$

$$\hat{A}(j) = \mu \cdot S_{Z,j} \cdot D_{HDU,j}^{\alpha} \cdot D_{TU,j}^{\beta} \cdot e^{\gamma_1 \cdot S_{EDU,j} + \gamma_2 \cdot S_{COM,j}} \tag{6}$$

$$\hat{A}(j) = \mu \cdot S_{Z,j} \cdot e^{\beta \cdot D_{TU,j} + \gamma_1 \cdot D_{EDU,j} + \gamma_2 \cdot D_{COM,j}} \tag{7}$$

with $j = 1..N_Z$ and $\mu,\ \alpha,\ \beta, \gamma_1, \gamma_2$ parameters to be calibrated.

Using GIS spatial regression analysis tools, we have tested the estimated results based on equations (2) – (7). Better assessments were obtained using the second class of variables (densities of characteristics of urban zones) than using the first class of variables describing the total characteristics (P_{HDU}, S_{TU}, S_{EDU}, S_{COM}).

After many statistical tests and steps of calibration, we obtained the accident estimation equation:

$$\hat{A}(j) = 4.03 \cdot D_{HDU,j}^{0.92} \cdot D_{TU,j} \cdot e^{0.04 \cdot S_{EDU,j} + 0.09 \cdot S_{COM,j}} \text{ Accidents/5 years} \tag{8}$$

or

$$\hat{A}(j) = 0.807 \cdot D_{HDU,j}^{0.92} \cdot D_{TU,j} \cdot e^{0.04 \cdot S_{EDU,j} + 0.09 \cdot S_{COM,j}} \text{ Accidents/ year,} \tag{9}$$

Applying eq. (8) on the available data, the values of 0.76 for correlation coefficient R^2 and 0.71 for adjusted correlation coefficient R^2 resulted, meaning that the model can be used to estimate accidents in the study area. Traffic accidents are composite events that can be caused by a variety of factors related to intrinsic features of the road network and urban environment but also to the behaviour of traffic infrastructure users, vehicle state, time and ambient conditions. Given that the proposed model aims to determine the influence of the urban patterns and land use on accidents, we consider that the values obtained for the correlation coefficients are satisfactory and, consequently, the identified independent variables have to be considered in urban planning and in traffic risk analysis.

The outcomes of eq. (8) lead to standard residual errors less than 1.5 for 70 zones (covering around 80% of the study area) and less than 0.5 for 42 zones (around 40% of the study area). Standard residual errors over 1.5 are obtained in zones with mixed use (residential and industrial) and land use changes in the last decade (zones which include business park development, large commercial areas). Additional variables (e.g. employees per activity categories and capacity of commercial areas) are necessary for enhanced calibration, but in this phase eq. (8) can be used to estimate accidents in the assessment of alternatives to urban planning.

4. CONCLUSION

Land use in each urban zone (residential, commercial, offices, mixed-use areas) is decisive for generated and attracted traffic flows. A new urban development has a direct influence on economic and social activities and implicitly on traffic flow and its associated risks. Thus, the model presented in this paper aims to identify the correlation between land use and traffic risk.

The decisions on urban planning (development of new residential areas, building of new business parks, location of large commercial centres) need rigorous validation of the consequences for road traffic and the associated accident risk. The size of traffic risk has to intervene as an additional criterion in these types of decisions.

The model presented in this paper represents a useful tool for assessments of the consequences of land use on road accidents. Using GIS procedures and spatial analysis models, the density of population of HDU zones and the density of road infrastructure area are identified as the most significant variables in road accident occurrence in the Bucharest city area. The density of educational zones and the density of commercial zones are also included as variables in the accident estimation model.

Better calibration could be obtained if the density of educational zones and density of commercial zones are computed as functions of their capacities instead of their surfaces. Furthermore, because HDU zones also include business parks, better estimation could be found if, beside the density of the population of HDU, the density of employees in categories and in urban zones will be analysed. Nevertheless, until a more detailed database is gathered, the developed model of the available data can provide significant results (with 0.76 for correlation coefficient R^2) on accident estimation function under urban zone characteristics and represent an important step in the development of a road accident estimation model for the Bucharest area.

References

1. Raicu, S. & Costescu, D. On estimate of risk associated with urban road traffic. *Advances in Automatic Control. Proceedings of the 16th International Conference on Automatic Control, Modelling & Simulation* (ACMOS 2014). Brașov – Romania. June 26-28. 2014. Mastorakis, N. & Udriste, C. & et al. (Eds.). WSEAS Press, Recent Advances in Electrical Engineering Series – 35. 2014. P. 92-97.
2. Pulugurtha, S.S. & Duddu, V. D. & Kotagiri, Y. Traffic analysis zone level crash estimation models based on land use characteristics. *Accident Analysis and Prevention*. 2013. Vol. 50. P. 678– 687.
3. Ladron de Guevara, F. & Washington, S.P. & Oh, J. Forecasting Crashes at the Planning Level: Simultaneous Negative Binomial Crash Model Applied in Tucson, Arizona. *Transportation Research Record*. 2004. Vol. 1897. P. 191–100.
4. Li, L. & Zhu, L. & Sui, D.Z. A GIS-based Bayesian approach for analyzing spatial–temporal patterns of intra-city motor vehicle crashes. *Journal of Transport Geography*. 2007. Vol. 15. P. 274–285.
5. Lord, D. & Mannering, F. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A*. 2010. Vol. 44. P. 291–305.
6. Lord, D. Issues related to the application of accident prediction models for the computation of accident risk on transportation networks. *Transportation Research Record*. 2002. Vol. 1784. P. 17–26.
7. Lord, D. & Persaud, B.N. Estimating the safety performance of urban road transportation networks. *Accident Analysis and Prevention*. 2004. Vol. 36 (4). P. 609–620.
8. Miaou, S.P & Lord, D. Modeling Traffic Crash-Flow Relationships for Intersections: Dispersion Parameter, Functional Form, and Bayes Versus Empirical Bayes Methods. *Transportation Research Record*. 2003. Vol. 1840. P. 31-40.
9. Hauer, E. Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation. *Accident Analysis and Prevention*. 2001. Vol. 33 (6). P. 799–808.
10. Lord, D. & Persaud, B. Accident Prediction Models With and Without Trend: Application of the Generalized Estimating Equations Procedure. *Transportation Research Record*. 2000. Vol. 1717. P. 102 – 108.
11. Hauer, E. & Harwood, D. W. & Council, F. M. & Griffith, M., S. Estimating Safety by the Empirical Bayes Method: A Tutorial. *Transportation Research Record*. 2002. Vol. 1784. P. 126-131.
12. Fernandes, A. & Neves J. An approach to accidents modeling based on compounds road environments. *Accident Analysis and Prevention*. 2013. Vol. 53. P. 39-45.
13. Marshall, W. E. & Garrick, N. Does Street Network Design Affect Traffic Safety? *Accident Analysis and Prevention*. 2011. Vol. 43 (3). P. 769-781.

14. Fleury, D. (Coord.) *Projets urbains de cohérence fonctions/réseaux*. 2011. PREDIT Groupe Opérationnel N° 2: Qualité des systèmes de transport. 08 MT S 027 et 08 MT S 028. IFSTTAR-France.
15. Lord, D. Modeling motor vehicle crashes using Poisson-gamma models: examining the effects of low sample mean values and small sample size on the Estimation of the fixed dispersion parameter. *Accident Analysis and Prevention*. 2006. Vol. 38 (4). P. 751–766.
16. Cameron, A. C. & Trivedi, P. K. *Regression analysis of count data*. Cambridge University Press. 1998.
17. Mitchell, A. *The ESRI Guide to GIS Analysis.* Vol. 2. ESRI Press. 2005.
18. EuropeAid/123579/D/SER/RO. *Bucharest General Master Plan for Urban Transport. Final Rapport*. 2008. [In Romanian: *Master Plan General pentru Transport Urban – Bucuresti, Sibiu si Ploiesti*. Raport Final Bucuresti. EuropeAid/123579/D/SER/RO].
19. Zhang, H. & Liu, Y. & Li, B. Notes on discrete compound Poisson model with applications to risk theory. *Insurance: Mathematics and Economics*. 2014. Vol. 9. P. 325–336.

Acknowledgments

The work has been funded by the Sectorial Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreements POSDRU/159/1.5/S/132395 and POSDRU/159/1.5/S/132397.

Received 17.01.2015; accepted in revised form 23.08.2016