ANALYSIS OF INTRINSIC FACTORS CONTRIBUTING TO URBAN ROAD CRASHES

S. RAICU, D. COSTESCU & S. BURCIU
Politehnica University of Bucharest, Romania.

ABSTRACT

The impact of risk associated with traffic is rarely included in the evaluation of projects for increasing urban traffic fluency, although the social costs of traffic crashes are estimated as very high. Lately, more frequently, the risk associated with urban road traffic is included, as a supplementary criterion, in the selection of the best urban planning scenarios, in order to a-priory minimize the number of crashes. Therefore, we aim to develop tools to enable the analysis of different intrinsic factors (characteristics of urban area and of road network) on traffic safety performances. The paper presents an analysis of the traffic crashes registered in Bucharest in the period 2008–2014. Following the analysis, the highest values of the average number of crashes were identified for signalized intersections that include tram infrastructure. Hence, the study is continued for this category of network features for which a model to estimate traffic crashes is proposed.

Keywords: crash analysis, crash prediction function, GIS modelling, road safety, spatial analysis.

1 INTRODUCTION

In all of the concerns to reduce the negative effects of urban traffic, particular attention is given – both locally and globally – to issues concerning road safety. Under the phenomeno-logical aspect, the urban area, where there is a risk of traffic crashes, is a complex space. In this system, identifying urban entities vulnerable to traffic crashes, understanding the causes of crashes, identifying priority actions to reduce the risk of accidents and their application at the stage of urban planning might lead to improved road safety [1–3].

This paper presents research conducted in order to identify solutions to mitigate the risk associated with road traffic in Bucharest, an urban space marked by radical changes in the past two decades. The radical changes of socio-economic life, major structural changes in the commercial sector - caused by the development of large shopping centres, new concentrations of residential areas and offices, located in the urban area or at its perimeter, structural and spatial changes of the points of interest (for work, education, leisure) have provoked an increase in traffic intensity and heterogeneity, and thus, in risk exposure.

The main objectives of this study are to identify the relationships between the frequency of road crashes and intrinsic factors that contributed to their production. The source data for the analysis consisted of serious road crashes records (those that have caused deaths, serious injuries and/or significant material damages) in Bucharest, for the 2008–2014 period. The *Geographic Information Systems* (GIS) tools were used for geocoding the points where there have been crashes and their allocation to the road network elements and to the spatial analysis units. Next, the developed data structures are being presented and their analysis stages. Finally, a function for predicting crashes within signalized intersections for road and tram traffic, the most vulnerable feature category, is proposed.

2 DATA AND METHODOLOGY

2.1 The spatial model of the study area

Previous studies have identified three main categories of variables considered in the analysis and estimation of crashes: (1) traffic characteristics, (2) road characteristics and (3) socio-demographic factors [4]. To analyse these variables, a GIS model that includes geodatabase to characterize road network and urban area divided into *Traffic Analysis Zones* (TAZ), was developed.

2.1.1 Urban road network
GIS procedures were used for inclusion in the digital model of the road network of the attributes that describe the characteristics of intersections and road sections [5]. These attributes will allow the identification of the most vulnerable features of urban infrastructure. For the major intersections codes were used that describe:

- Traffic lights existence;
- Configuration: 3-way intersections (T-, Y- or skewed Y-intersections); 4-way intersections (regular cross, skewed 4-legged intersections); more than 4-way intersections;
- Type of infrastructure: only road traffic intersections and intersections including tram infrastructure;
- Public transport stations existence;
- Pedestrian crossing existence/type,

and for the road sections:

- Section length;
- Number of lanes;
- Tram lines existence (simple, double);
- Bus public transport lines existence;
- The existence of street side parking.

Traffic characteristics, such as flow and speed can be considered as the most important factors influencing crashes [4]. Generally, the studies regarding crashes frequency use the *Annual Average Daily Traffic* (AADT) to characterise the flows, even though, in this manner, significant hourly non-uniformities are excluded from the analysis [6]. In this study, the AADT values were estimated based on the traffic flows values used in the Sustainable Urban Mobility Plan for Bucharest - Ilfov Region [7].

Speed analysis was not included in the study because no relevant data were available. However, given that traffic in Bucharest happens under quasi-permanent congestion, speed could be considered not an important factor at this level of analysis.

2.1.2 Traffic analysis zones
Several studies indicate that different socio-demographic variables such as population, working population, poverty, economic activity influence traffic crashes [8, 9]. To take into account these features, especially in the second part of the study, which aimed at developing a model to estimate crashes, the TAZ data set defined in the Sustainable Urban Mobility Plan for Bucharest - Ilfov Region [7] was used. In that data set the area of Bucharest (240 sq.km, with
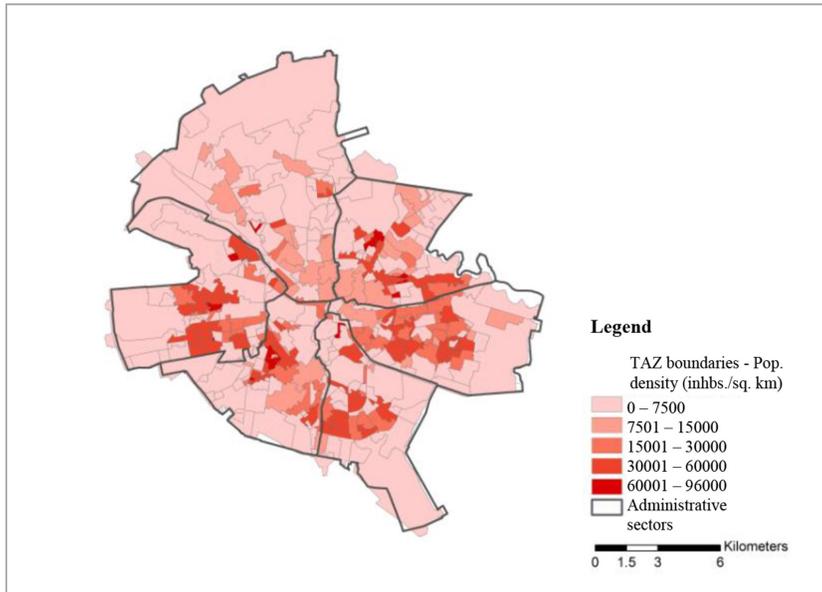
Figure 1: Traffic analysis zone geodatabase with socio-demographic attributes.

approximately 1,885,000 inhabitants) was divided in 366 TAZ. Based on these data (Fig. 1), for each TAZ, population density, employment density, school population density, road network density, the percentage area occupied by roads were determined, attributes that may influence the frequency of road crashes.

2.2 Crash data

Traffic crashes are complex events, in the production of which multiple factors are involved. Typically, researchers seek to establish links between road infrastructure, environment and socio-economic characteristics of spatial units based on the number of road crashes recorded at various regional levels. The study presented in this paper refers to the area of Bucharest, and the analyses refer only to intrinsic factors, those that characterize urban areas and road network elements (factors related to road infrastructure user behaviour and meteorology were not included in this study).

The data available for this study contain records of traffic crashes in Bucharest between 2008 and 2014. The processed list included data on severe crashes (involving injuries, fatalities and/or significant property damage), for which the address and the cause were specified (without including temporal or environmental aspects). Based on causes, the crashes were classified in two categories: (1) the ones with pedestrians involved and (2) the ones with only vehicles involved, Fig. 2.

To identify the properties of the most vulnerable road network features, the crash data were geocoded and geodatabase with the crash location within urban areas and on the road network was obtained. Then, each entity of the road network on which crashes occurred was assigned attributes of the TAZ where it is located, Fig. 3.

By adding the annual frequencies of crashes on the road network features (intersections and road sections), the total number of crashes on each network entity for 7 years was
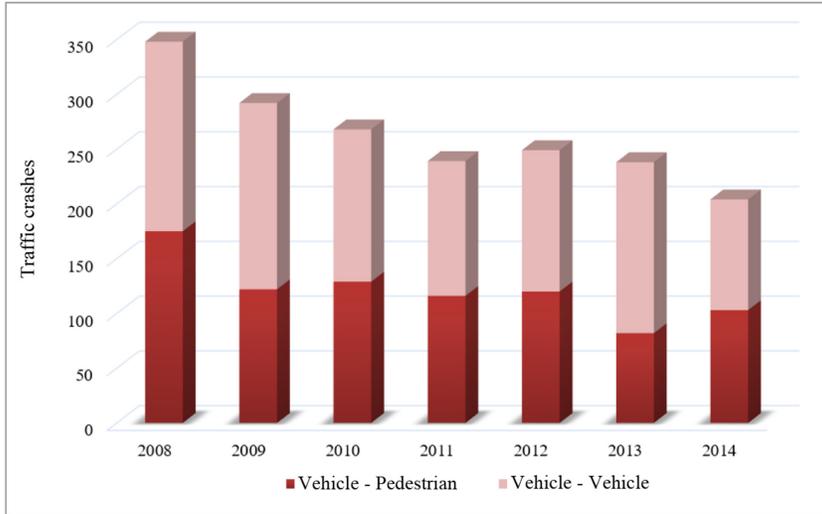
Figure 2: Evolution of traffic crashes in Bucharest (2008–2014).
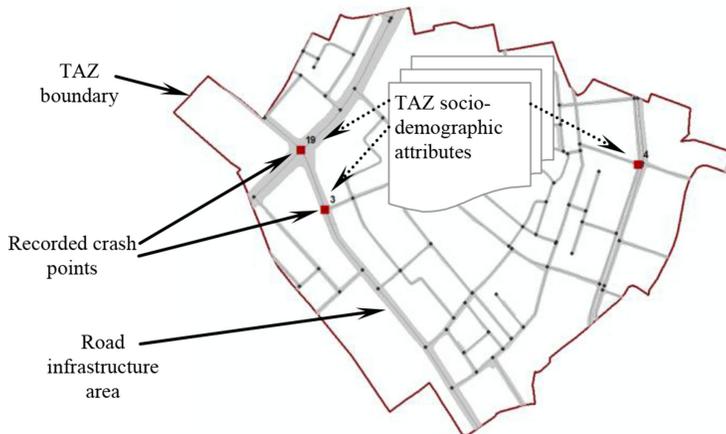
Figure 3: Allocation of socio-demographic attributes.

obtained, Fig. 4. This procedure led to the loss of the annual crash frequency non-uniformities, but consisted an advantage in obtaining larger sample data for each category of network features.

2.3 Data analysis

Databases resulting from processing have allowed the classification of the road network features, depending on their class and the type of crashes recorded, Table 1. The data show, as expected, that the percentage of crashes at intersections is significantly higher (83.5%) than the ones registered on road sections (22% of the points that recorded crashes are placed on road sections). Therefore, the analysis was continued for crashes recorded at intersections.
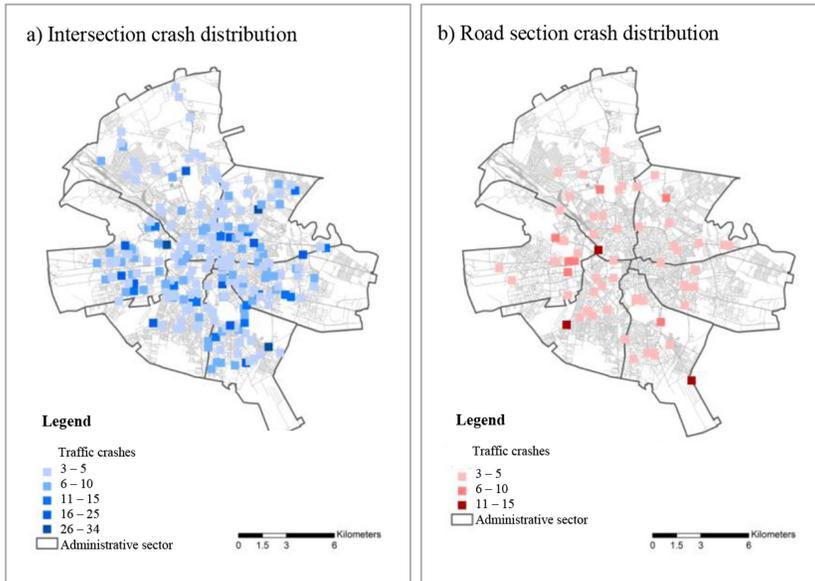
Figure 4: Traffic crashes spatial distribution in Bucharest (2008–2014).

Table 1: Allocation of the recorded traffic crashes on road network feature classes.

Road network feature class	Crash category	No. of points with recorded crashes	Recorded crashes	Crashes/point Min – Max - Mean – St. Dev.
Intersections	*Total*	240	1485	3 – 34 – 6.2 – 4.9
	– Vehicle - vehicle	135	839	3 – 34 – 6.2 – 5.1
	– Vehicle - pedestrian	127	646	3 – 22 – 5.1 – 3.3
Road sections	*Total*	66	293	3 – 15 – 4.4 – 2.5
	– Vehicle - vehicle	25	104	3 – 12 – 4.1 – 2.0
	– Vehicle - pedestrian	44	189	3 – 12 – 4.3 – 2.2
Total		306	1778	3 – 34 – 5.8 – 4.6

The data were classified according to traffic lights at intersections, their configuration, the type of infrastructure, Table 2. It is noted that 42% of crashes at intersections have been recorded within signalized intersections and that, of this percentage, the highest share is for the crashes in intersections with four legs. A significant share between points where there have been recorded crashes belongs to the unsignalized intersections category (53% of the total points with crashes registered, 68% of accident registered intersections). This category of unsignalized intersections will be subject to a subsequent study, which will identify any changes in the socio-economic structure of the areas where they are located, the traffic flows generated/attracted and the need for management adjustment and systematization of traffic.

However, it can be seen that the average number of crashes is higher for signalized intersections: 8.2 crashes/signalized intersection, compared to 5.2 injuries/unsignalized

Table 2: Classes of intersection traffic crashes.

Intersection characteristics			Crash category/No. of points with recorded crashes		Crashes	Crashes/point Min – Max - Mean – St. Dev.
Signalized intersection			Total	76	621	3 – 34 – 8.2 – 6.6
			Vehicle - vehicle		358	0 – 34 – 4.7 – 6.3
			Vehicle - pedestrian		263	3 – 22 – 3.5 – 4.2
	Layout	• 3-way: T-, Y- or skewed Y-inter-sections	Total	18	85	3 – 9 – 4.7 – 2.1
			Vehicle - vehicle		29	0 – 9 –1.6 – 2.7
			Vehicle - pedestrian		56	0 – 8 – 3.1 – 2.6
		• 4-way: regular cross, skewed 4-legged intersec-tions	Total	50	468	3 – 34 – 9.5 – 7.3
			Vehicle - vehicle		275	0 – 34 – 5.6 – 6.7
			Vehicle - pedestrian		193	0 – 22 – 3.9 – 4.9
		• more than 4-way intersections	Total	8	68	3 – 18 – 8.5 – 5.9
			Vehicle - vehicle		54	0 – 18 – 6.7 – 7.2
			Vehicle - pedestrian		14	0 – 5 – 1.7 – 1.8
Type of infrastructure		• With tram tracks included in the road infrastruc-ture	Total	50	504	3 – 34 – 10.1 – 7.3
			Vehicle - vehicle		284	0 – 34 – 5.6 – 7.4
			Vehicle - pedestrian		220	0 – 22 – 4.4 – 4.7
		• Designed only for road traffic (without tram facilities)	Total	26	117	3 – 10 – 4.6 – 2.4
			Vehicle - vehicle		74	0 – 9 – 2.9 – 2.5
			Vehicle - pedestrian		43	0 – 8 – 1.7 – 2.2
Unsignalized intersection			Total	164	864	3 – 25 – 5.2 – 3.5
			Vehicle - vehicle		481	0 – 25 – 2.9 – 3.9
			Vehicle - pedestrian		383	0 – 17 – 2.3 – 3.0
	Layout	• 3-way: T-, Y- or skewed Y-inter-sections	Total	32	205	3 – 25 – 6.4 – 4.5
			Vehicle - vehicle		70	0 – 25 –2.2 – 4.7
			Vehicle - pedestrian		135	0 – 16 – 4.2 – 3.8
		• 4-way: regular cross, skewed 4-legged intersec-tions	Total	132	659	3 – 19 – 4.9 – 3.2
			Vehicle - vehicle		411	0 – 19 – 3.1 – 3.7
			Vehicle - pedestrian		248	0 – 17 – 1.8 – 2.5

intersection. Also, one can see that from the signalized intersections subcategories, the highest average was recorded for intersections that include trams infrastructure (10.2 crashes/intersection). The available data only included crashes with road vehicles and/or pedestrians, but no trams.

To properly interpret the factors contributing to crashes, discrete examination of distinct categories of road infrastructure elements is recommended [10]. Given the large share that signalized intersections with tram lines hold, this category was chosen to develop a function to estimate the crashes.

3 CRASH ESTIMATION FUNCTION

Traffic crashes are complex events, difficult to estimate. Numerous studies have proposed variants of models to estimate crashes for different categories of road infrastructure features and have revealed several issues that can lead to incorrect definition and calibration of the crashes estimation functions [4, 11, 12].

Over-dispersion is a main characteristic of data regarding crashes. As these data are registered in relation to the point of record, over-dispersion can be explained by data clustering, insufficient amount of data to identify time correlations and omission of some variables that would explain the heterogeneity of neighbouring areas in which crashes were recorded.

Generally, negative binomial (NB) regression models are applied to establish the relation between the number of crashes and the factors leading to their occurrence [13–15]. NB models may take into account the unnoticed heterogeneous effects between neighbourhoods (due to omitted variables), but cannot take into account the spatial correlations between the zones. An alternative manner is to develop spatial econometric models, which may include (1) traditional econometric models for continuous data and (2) Bayesian hierarchical models, suitable for non-positive random data records [4, 16].

The analysis presented in this paper intends to develop a function for predicting crashes within signalized intersections with both road and tram infrastructure. Because of the geodatabase created, with the locations of the points where crashes occurred, GIS procedures were applied to determine the Moran's I Index and to check spatial correlations [17].

For the sample consisting of 50 signalized intersections including tram lines a small value of Moran's I Index was obtained, indicating a random spatial distribution and that over-dispersion in this case is not caused by data concentration. As a result, the development of the crashes estimation model may start for a NB regression equation:

$$A = \beta_0 \cdot Q_1^{\beta_1} \cdot Q_2^{\beta_2} \cdot Q_t^{\beta_3} \cdot e^{\sum_j a_j \cdot x_{kj}} \; crashes/year, \tag{1}$$

where A represents the expected number of crashes for the selected intersection class;

Q_1, Q_2-entering the traffic flow in intersection from the major, respective minor road section;

Q_t-entering tram flow in an intersection; recorded data about tram flow were not available, but we estimated the values according the tram schedule published by the urban transit operator;

β_1, β_2, β_3-effect of traffic volume on the expected number of crashes;

x_{kj}-explanatory variables (including traffic);

α_j-parameters to be estimated, representing the effects of variables j^{th} on crash occurrence;

β_0-intercept.

Subsequently, we have applied correlation tests for the available datasets, selected only for the category of intersections analysed (signalized intersections with tram lines) to identify the variables that might influence crash occurrence. By using correlation coefficient (R^2), adjusted correlation coefficient (AR^2), Akaike information criterion ($AICc$) and variance inflation factor (VIF) as evaluators [17], the following equation for the estimation of crashes at intersections with traffic lights and tram lines resulted:

$$A = 1.61 \cdot Q_1^{0.012} \cdot e^{0.00014 \cdot Q_1 - 0.003 \cdot x_1} \; crashes/year, \tag{2}$$

where x_1 is the number of lanes on the main artery entering the intersection.

For the available data no influence of the trams flow on crashes was identified. This might be explained by data over-dispersion caused in this case, by the heterogeneity of the intersections with traffic lights and crossed by tram lines (junctions with 3 and 4 arms with different layouts, roundabouts). At this stage of the study, given that the total number of intersections in this category was 50, we could not make estimates for subgroups of intersections. The data samples for subgroups of the intersections would have been too small to lead to accurate estimates.

4 CONCLUSION

Solutions to ensure mobility needs, dependent on size, shape and pattern of urban areas, are potentially responsible for the measure of risk associated with traffic. In regard to the demands of sustainable development, these solutions must harmonize the accessibility needs of residents with the requirements of the environment protection and with the quality of life, which includes the critical need to reduce the risk of traffic crashes. To apply appropriate solutions to improve road safety, it is necessary to identify the attributes of the road infrastructure elements and of the urban functions within low-performing road safety areas.

In this framework, the research presented has led to the development of a GIS model that allows identifying the particularities of the areas with risk of crashes occurrence in Bucharest and obtaining data series necessary to develop crashes estimation models. Literature recognizes the value of these models, which, in addition to the necessary traffic studies, allow the estimation of the risk associated with traffic and its inclusion as an additional criterion when analysing different development and configuring alternatives of the urban space.

Previous studies showing different variants of crashes estimation models highlight the difficulties in applying and improving them. The developed GIS model allows the elimination of one of the most common difficulties encountered, the over-dispersion caused by data spatial concentration. The main problems in this study consisted in the lack of a tradition to collect and integrate detailed data about traffic crashes in Bucharest.

In addition, the heterogeneity of road infrastructure elements in the study area has led to obtaining relatively low data samples for developing crashes estimation functions for subgroups of road infrastructure features. Data aggregation for the period 2008–2014, however, allowed the acquisition of samples enabling the identification of a series of variables that influence crash occurrence. The proposed crash estimation function and the processed data are an important step for the development and time verification of the tools necessary to evaluate the frequency of crashes, for various subcategories of intersections and urban arteries.

REFERENCES

[1] Fleury, D., (coord.), *Projets urbains de cohérence fonctions/réseaux*. PREDIT Groupe Opérationnel N°2: Qualité des systèmes de transport, 08 MT S 027 et 08 MT S 028, IFSTTAR, France, 2011.

[2] 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*, **1897**, pp. 191–100, 2004.
https://doi.org/10.3141/1897-25

[3] Pulugurtha, S.S., Duddu, V. D. & Kotagiri, Y., Traffic analysis zone level crash estimation models based on land use characteristics. *Accident Analysis and Prevention*, **50**, pp. 678–687, 2013.
https://doi.org/10.1016/j.aap.2012.06.016

[4] Quddus, M.A., Modelling area-wide count outcomes with spatial correlation and heterogeneity: An analysis of London crash data. *Accident Analysis and Prevention*, **40**, pp. 1486–1497, 2008.
https://doi.org/10.1016/j.aap.2008.03.009

[5] Raicu, S., Costescu, D. & Burciu, S., The evaluation of road safety performances in urban areas. *Urban Transport XX. WIT Transactions on The Built Environment 138*, ed. C.A. Brebbia, WIT Press, England, pp. 447–458, 2014.

[6] Lord, D. & Persaud, B.N., Estimating the safety performance of urban road transportation networks. *Accident Analysis and Prevention*, **36**(4), pp. 609–620, 2004.
https://doi.org/10.1016/S0001-4575(03)00069-1

[7] Sustainable Urban Mobility Plan 2016–2030. Bucharest - Ilfov Region. Rom Engineering Ltd., AVENSA Consulting SRL, Bucharest, 2016, available at http://pmud.ro/ (*In Romanian*).

[8] Kim, K., Brunner, I.M. & Yamashita, E.Y., The influence of land use, population, employment and economic activity on accidents. *Transportation Research Record,* **1953**, pp. 56–64, 2006.
https://doi.org/10.3141/1953-07

[9] Noland, R.B. & Quddus, M.A., A spatially disaggregate analysis of road casualties in England. *Accident Analysis and Prevention*, **36**(6), pp. 973–984, 2004.
https://doi.org/10.1016/j.aap.2003.11.001

[10] Lee, C. & Abdel-Aty, M., Comprehensive analysis of vehicle–pedestrian crashes at intersections in Florida. *Accident Analysis and Prevention*, **37**, pp. 775–786, 2005.
https://doi.org/10.1016/j.aap.2005.03.019

[11] Lord, D. & Mannering, F., The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transportation Research Part A*, **44**, pp. 291–305, 2010.
https://doi.org/10.1016/j.tra.2010.02.001

[12] Persaud, B. & Lyon, C., Empirical Bayes before–after safety studies: lessons learned from two decades of experience and future directions. *Accident Analysis ans Prevention*, **39**(3), pp. 546–555, 2007.

[13] Shankar, V., Mannering, F. & Barfield, W., Effect of road way geometrics and environmental factors on rural freeway accident frequencies. *Accident Analysis and Prevention*, **27**(3), pp. 371–389, 1995.
https://doi.org/10.1016/0001-4575(94)00078-Z

[14] Abdel-Aty, M.A. & Radwan, A.E., Modeling traffic accident occurrence and involvement. *Accident Analysis and Prevention*, **32**(5), pp. 633–642, 2000.
https://doi.org/10.1016/S0001-4575(99)00094-9

[15] Lord, D., The prediction of accidents on digital networks: characteristics and issues related to the application of accident prediction models, Ph.D. Dissertation, Department of Civil Engineering, University of Toronto, Toronto, 2000.

[16] Zeng, Q. & Huang, H., Bayesian spatial joint modeling of traffic crashes on an urban road network. *Accident Analysis and Prevention*, **67**, pp. 105–112, 2014.
https://doi.org/10.1016/j.aap.2014.02.018

[17] Mitchell, A., *The ESRI Guide to GIS Analysis*, Vol. 2, ESRI Press: Noida, 2005.