Middle-East Journal of Scientific Research 18 (12): 1859-1866, 2013
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2013.18.12.21519

The Prediction Models of Motorcycle Accidents on Surabaya Arterial Roads Using Generalized Linear Models

Machsus, Harnen Sulistio, Achmad Wicaksono and Ludfi Djakfar

Department of Civil Engineering,
Faculty of Engineering, University of Brawijaya, Malang, East Java, Indonesia

Abstract: In Indonesia, traffic accident grows 11.64 % per-year, which is dominated by motorcycles traffic accidents. Similarly in Surabaya, the second biggest city in Indonesia, traffic accident increased by 17.71% per-year and motorcycles traffic accident occupied 70.7 % of the total traffic accident. This paper aims to develop a predictive model of the accident, in order to overcome motorcycle accidents on arterial roads in Surabaya, Indonesia. In developing the model, the researchers used generalized linear models approach, with a Poisson distribution and logarithmic link function. The results of this study indicate that the traffic volume, length of roads, number of access points and the speed of traffichave a significant effect in describing the motorcycle accidents on arterial roads, where an increasing value of those variables lead to the increasing risk of accidents. These findings can help many stakeholders in traffic engineering solve the problems related to motorcycle accidents on arterial roads in Surabaya and many other cities.

Key words: Motorcycle accidents · Prediction model · Traffic flow · Road length · Access points · Traffic speed

INTRODUCTION

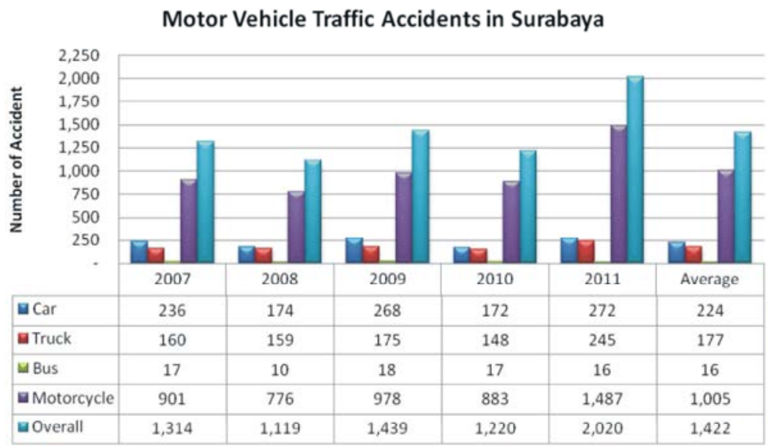
Transportation plays a major role in almost everyone's daily life, where the quality of life in community can be affected by the road trip experience [1]. In Indonesia, the number of traffic accidents reaches 109,776 incidents in 2011. Vehicles involved were approximately 239,257 units with the total victims were 176,763 people, which consisted of: 31,185 fatalities, 36,767 serious injuries and 108,811 minor injuries. Estimated materialloss caused by the accidents reachedRp 86.09 billion, (US$ 1≈ Rp11,500) [2]. The number of traffic accidents during 2006 to 2011 had been fluctuated, but the number has grown 11.64% ever since. Similarly, the number of vehicles involved in the accidents, the number of casualties and the material losses have also increased.

Analyzing traffic accidents in urban areas are interesting to be reviewed, because its number is greater than the accidents occurred outside urban areas [3]. Similarly, accidentsoccurred in major cities, such as Surabaya, is greater than accidentsoccurred in small towns. It is reported that in Surabaya, motorcycles dominate the proportion of vehicle type involved in the accidents. The number of motorcycles involved in traffic accidents in Surabaya is higher than other types of vehicles [4]. Figure 1 shows the proportion of motorcycle involvement in traffic accident.

Fig. 1 shows that from 2007 to 2011, the involvement of various types of vehicles accidents in Surabaya was on average of, as follows:70.7% (1,005 units) for motorcycles, 15.8% (224 units) for cars, 12.5% (177 units) for trucks and 1.1% (16 units) for the bus.Although the number was fluctuating, the motorcycle accidents in Surabaya from 2007 to 2011 were prone to increase with an average growing number of 17.71% annually.

Furthermore, the characteristics of traffic accidents can be seen from several parameters as follows: the frequency of incident, the vehicles involved, the number of casualties (fatalities, serious and minor injuries) and the amount of material loss. Generally, the characteristics of traffic accidents in Surabaya also reflect a phenomenon in Indonesia [2]. The number of motorcycle accident victims was higher than other types of vehicles [4]. Table 1 shows the characteristics of traffic accidents in Surabaya.

Corresponding Author: Machsus, Department of Civil Engineering, Faculty of Engineering, University of Brawijaya, Malang, East Java, Indonesia.



Source: Traffic Accident Unit, SurabayaPolice, Indonesian National Police

Fig. 1: Motor Vehicle Traffic Accidents in Surabaya

Table 1: Traffic accidents in Surabaya from 2007 to 2011

No	Description	Unit	Years					Average Growth %
			2007	2008	2009	2010	2011	
1	Frequency of incident	incident	739	643	855	711	1,119	15.13%
2	Vehicles involved	unit	1,314	1,119	1,439	1,220	2,020	16.03%
3	Fatalities	person	239	202	229	324	361	12.70%
4	Serious injuries	person	83	78	176	240	580	74.41%
5	Slight injuries	person	642	574	689	377	680	11.13%
6	Total casualties	person	964	854	1,094	941	1,621	18.74%
7	Loss of material	billion Rp.	0.88	0.32	0.94	0.58	0.85	34.80%

Source: Traffic Accident Unit, SurabayaPolice, Indonesian National Police

In Table 1, the number of traffic accidents in Surabaya in 2011 was around 1,119. The vehicles involved were 2,020, with the casualties were 1,621 people, comprised of: 361 fatalities, 580 seriously injured and 680 slightly injured. The material loss estimation caused by the accident reached Rp. 0.85 billion. Although it had fluctuated for a while, the number of traffic accidents from 2007 to 2011 was likely to increase, with an average growth of 15.13% per year.Likewise, the number of vehicles involved the number of casualties and material losses also increased.Based on the above description, it can be concluded that the motorcycle involvement and its population are the largest factors in most traffic accidents.In addition, the number of traffic accidentstend to increase in terms of the parameters, for example: the frequency of incident; the vehicles involved; the number of victims; and the amount of material loss. Considering those parameters, the motorcycle would still be a research objectin the field of traffic safety.

Accident prediction modelis used to estimate the frequency of traffic accidents. In addition, it can be used to identify and determine the relationship between the factors that influence it, such as geometric, environmental and operational factors [5, 6]. The relationship between these factors has greatly influenced the variables used in the modeling of motorcycle accidents [7]. The development of prediction model on the road needs to consider the level of accidents, so the output can be used in planning and implementingthe action programs for the improvement of road safety [8, 9].

Previously, conventional linear regression approach is used for developing traffic accident prediction model. Through this approach, the traffic accident data are assumed to be normally distributed with a constant variance. But the assumption of this theoretical distributioncannot properly represent the data distribution of traffic accidents.The data characteristics of traffic accident cannot be represented by a normal distribution, especially the distribution of the time and place of occurrence. Finally, the conventional linear regression approach is not used in the modeling of traffic accidents [10].

Furthermore, the researcher used generalized linear models approach in the modeling, in which the traffic accident data is not assumed to be normally distributed anymore. Data distributionswhich are often used for modeling traffic accident in previous studies are the Poisson and the negative binomial distribution. The use of generalized linear models with the assumption of Poisson or negative binomial distributionand logarithmic link can describe random, discrete and non-negative events in orderto represent the characteristics of traffic accidents [10, 11]. Therefore, this paper will discuss the prediction modeldevelopment of motorcycle accidents occurred on arterial roads. Besides, it also discusses the effect of each explanatory variable on motorcycle accidents.

MATERIALS AND METHODS

The Data Collection: The location of this case study was on urban roads in Surabaya, Indonesia. Data collection comprises of: the number of motorcycle accidents, traffic volume, roadlength, number of lanes,number of access points, roadwidth, median facility, trafficspeed, number of directions and road shoulder. Road segments were selected based on the conditions between 2009 to 2012 with the criteriafollows: a. only on arterial roads, b. no significant change in the characteristics during the study period, c. do not have frontage road and d. geometric characteristics of each road segment is relatively the same or relatively homogeneous. Based on those criteria, this study selected 72 sections, from the total 82 arterial roads in Surabayawhich haveroad segments.

Motorcycle accident data, were obtained from the Traffic Accident Unit, Surabaya Police, Indonesian National Police. Based on the monthly police report, the accident frequencies were selected from the arterial roadschosen in this study. The traffic volume; vehicle speed and geometric arterial roaddata were acquired from Surabaya Government Agencies, including: the Transportation Department, Department of Highways, Planning and Urban Development. However, if the necessary data was not available or less valid, then data collection was conducted directly in the research locations (arterial roads). This field survey was to ensure that all data used in this study was true.

Model Development: In this study, the response or dependent variable used was the number of motorcycle accidents per year. While the explanatory variables, including: traffic volume, roadslength, number of lanes, number of access points, road width, the median facility, traffic speed, number of carriageways and average shoulder width, as shown in Table 2 below. The selection of explanatory variables was based on several previous studies, as well as on the justification of the researchers by considering the results of the field observations.In particular, the number of access points is a new explanatory variable that have not been used in previous research on modeling motorcycle crash.

Motorcycle accident prediction model used generalized linear models approach (GLM) with a Poisson distribution and logarithmic link function. Referringto the previous modeling studies, the theoretical model which contains all terms used in this study, such as the following equation:

$$MCA = k\ FLOW^{\alpha 1} LR^{\alpha 2}\ e(\beta_1 LN + \beta_2 AP + \beta_3 RW + \beta_4 RM + \beta_5 SPEED + \beta_6 CN + \beta_7 SHDW + e) \quad (1)$$

MCA is the number of motorcycle accidents per year, while the response variable, consisting of: FLOW, which represents the volume of vehicle (pcu / h); LR, which indicates arterial road length (meters), AP, which represents the number of access points per-mile and SPEED is the 85th percentile vehicle speed(km / hours). For more details, a description of the explanatory variables is presented in Table 2 below.

By using logarithmic transformation, the accident prediction model can be presented in a log-linear version, such as the following equation:

$$Ln(MCA) = Ln(k) + \alpha_1\ Ln(FLOW) + \alpha_2 Ln(LR) + \beta_1\ (LN) + \beta_2 (AP) + \beta_3\ (RW) + \beta_4 (RM) + \beta_5 (SPEED) + \beta_6\ (CN) + \beta_7\ (SHDW) + e \quad (2)$$

Table 3: Descriptions, variable type, factor levels, and coding system of the explanatory variables

No	Explanatory Variables	Description	Variable Type	Factor Levels	Coding System
1	FLOW	Traffic Flow (pcu/hour)	Continuous		Flow (pcu/hour)
2	LR	Length of Roads	Continuous		LR (m)
3	LN	Number of Lanes	Continuous		LN (lanes)
4	AP	Number of Access Points	Continuous		AP (access points/km)
5	RW	Road Width	Continuous		RW (m)
6	RM	Median Facility	Categorical	2	(0) Non-median
					(1) median
7	SPEED	Speed	Continuous		Speed (km/hour)
8	CN	Number of Carriageways	Categorical	2	(1) single carriageway
					(2) dual carriageway
9	SHDW	Average Shoulder Width	Continuous		SHDW (m)

"ln(x)" in equation (2) above represents a notation on a logarithmic-linear version. The "ln(x)" function is expressed as the inverse of the exponential function. In Equation (2) and (3), estimated coefficients will be represented by k, $\alpha 1$, $\alpha 2$, $\beta 1$, $\beta 2$, $\beta 3$, $\beta 4$, $\beta 5$, $\beta 6$ and $\beta 7$, while "e" is the error representing residual differences between actual and predicted models.

RESULTS AND DISCUSSION

In building the motorcycle accident prediction model, the researchersutilizedthe univariate and multivariate analysis using the auxiliary program, GENSTAT 15th Edition. The result of univariate analysis showed that all explanatory variables were used to qualify the significance, namely the t-probability value (tpr. <0.05), as shown in Table 3 below. It means that each of the explanatory variables affected the motorcycle accidents, when used as a partial analysis. Many studies have shown that the t-probability parameter estimates significant at the 95% confidence level, (t pr. <0.05) [11, 12]. If there were parameters of the explanatory variables which were not considered as significant, it was removed and not included, in the next stage of the analysis.

In order to determine the effect of explanatory variables simultaneously on the model of accident,multivariate analysis was conducted. The result of multivariate analysis indicated that only few explanatory variables that were considered as significant, which have a t-probability value (tpr. <0.05), as shown in Table 4 below. Thus, not all explanatory variables that significantly influence the accident model usingunivariate analysis, but also have a significant effect on the result of multivariate analysis.

Based on the results of multivariate analysis, explanatory variables were used, including: traffic volume, length of roads, number of access points and the speed of traffic. While other explanatory variables were excluded from the process of model building, because it couldnot be considered as significance, such as: variable number of lanes, roadwidth, median facility, number of directions and roadshoulder. It means that the variable of traffic volume, length of roads, number of access points and the speed of traffic have a significant effect on motorcycle accidents.

Table 4 shows that the estimated value of all parameters of the explanatory variablesare marked positive. It means that the increase in the value of traffic volume, roadlength, number of access points and traffic speed, will lead to the contributionof the increasing number of accidents. Furthermore, based on the multivariate analysis, the final result of the prediction model development of motorcycle accidents, it is obtained as follows:

$$MCA = 0{,}000008501\ FLOW^{0.567}\ LR^{0.7171} e(0.02085AP + 0.06958\ SPEED) \quad (3)$$

remark:

MCA = The number of motorcycle accidents per year
FLOW = The traffic volume (pcu / hour)
LR = Thelength of arterial roads (meters)
AP = The number of access points per kilometer
SPEED = The 85th percentile vehicle speed (km / hour)

The verification of the final model can be conducted by comparingthe estimated value of the model (fitted value) with the actual data (response), as shown in Figure 2 below. The result of the model estimation which is close to or suitable with the actual data of observation results shows that the resulting model represents the factual occurrence.

Interpretation: A final model has been developed to predict thelevel of motorcycle accident on arterial roads in Surabaya, Indonesia, as in equation 3. The model shows that the traffic volume, road length, number of access points and traffic speed significantly influence the motorcycle accidents. The result of this study supportsthe previous studies [11, 13, 14, 15, 20].

The influence of each explanatory variable on motorcycle accidentsis at least reflected in the value of the estimated coefficient, both marked positive or negative[11]. If estimated coefficientis marked positive, then the increasingvalue of these variables contributes to the increasing number of accidents. On the other hand, if estimated coefficient is marked negative, then the increasingvalue of these variables contributes to the reduced number of accidents. The analysis of the effect of changing variables on motorcycle accidents is briefly described in the following paragraphs.

The increasing number of traffic volume on arterial roads affects the increasing number of motorcycle accidents. For example, if the traffic volume has doubled up, then the model would predict an increasing number of motorcycle accidents by 32.5%. The effect of traffic volume in motorcycle accidentsis in accordance with the findings in previous studies [13, 14, 15, 16].

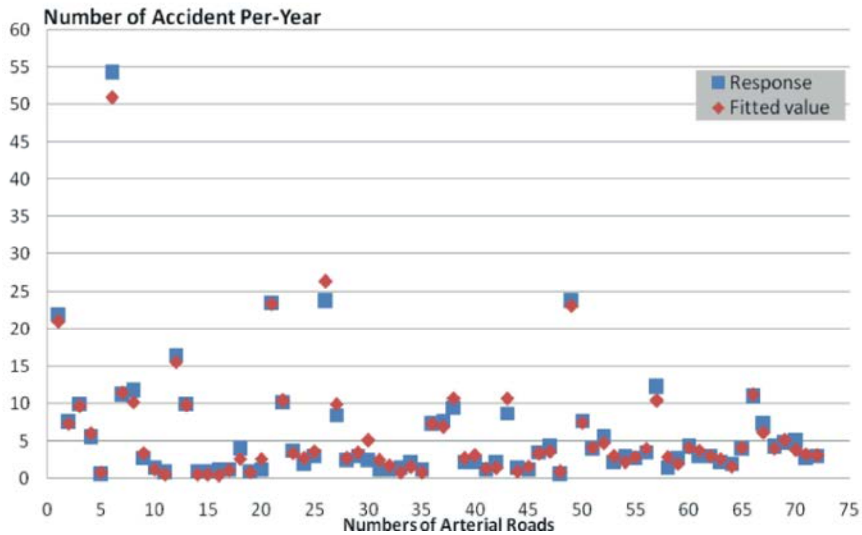


Fig. 2: The Comparison betweenModel Estimation with Actual Data

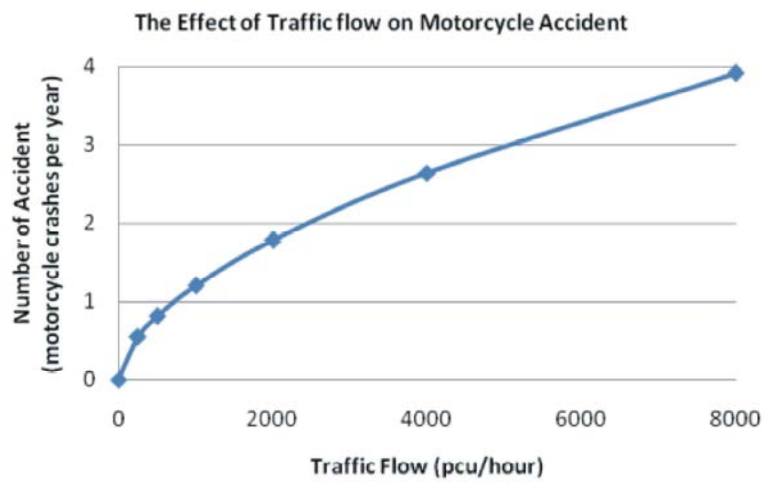


Fig. 3: The Effect of Traffic Flow on Motorcycle Accident

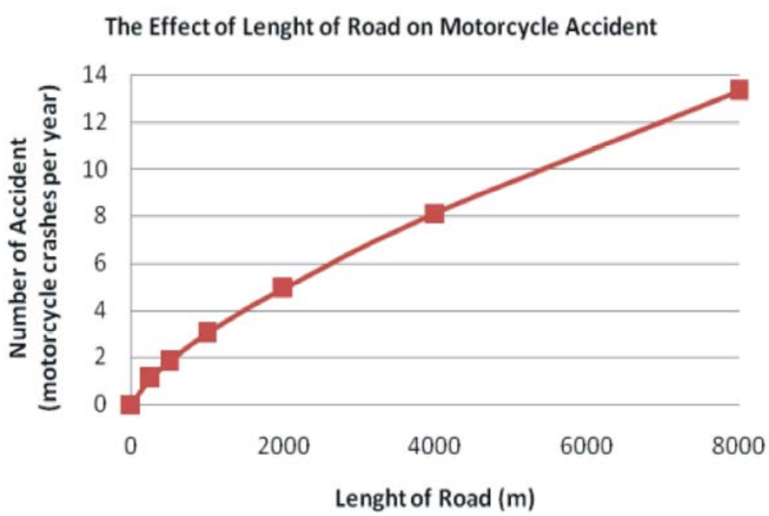


Fig. 4: The Effect of Length of Road on Motorcycle Accident

Table 4: Univariate Analysis of The Terms

No	Parameter	Estimate	Standard Errors	t (*)	t pr.	Antilog of Estimate	Significant at 0.05
1	Constant	-8.26	0.385	-21.43	<.001	0.0002587	Yes
	FLOW	1.3824	0.0454	30.42	<.001	3.984	Yes
2	Constant	-5.625	0.275	-20.46	<.001	0.003608	Yes
	LR	1.1965	0.0358	33.47	<.001	3.309	Yes
3	Constant	1.455	0.103	14.07	<.001	4.284	Yes
	LN	0.3692	0.0203	18.18	<.001	1.447	Yes
4	Constant	3.3153	0.0518	64.03	<.001	27.53	Yes
	AP	-0.01024	0.00355	-2.88	0.004	0.9898	Yes
5	Constant	1.949	0.104	18.69	<.001	7.018	Yes
	RW	0.07076	0.0056	12.63	<.001	1.073	Yes
6	Constant	2.8565	0.0479	59.67	<.001	17.4	Yes
	RM 1	0.4612	0.0553	8.34	<.001	1.586	Yes
7	Constant	-2.93	0.187	-15.64	<.001	0.05339	Yes
	SPEED	0.13167	0.00377	34.9	<.001	1.141	Yes
8	Constant	2.4277	0.0767	31.66	<.001	11.33	Yes
	CN 2	0.8829	0.0807	10.94	<.001	2.418	Yes
9	Constant	2.9763	0.0353	84.29	<.001	19.62	Yes
	SHDW	0.2061	0.0236	8.72	<.001	1.229	Yes

Note: Parameters for factors (RM1) are differences compared to the reference level (RM0),Parameters for factors (CN2) are differences compared to the reference level (CN1).

Table 5: Multivariate Analysis of the Terms

No	Parameter	Estimate	Standard Errors	t(*)	t pr.	Antilog of Estimate	Significant at 0.05
1	Constant	-11.675	0.857	-13.62	<.001	8.50E-06	Yes
2	FLOW	0.567	0.105	5.38	<.001	1.763	Yes
3	LR	0.7171	0.0972	7.38	<.001	2.048	Yes
4	AP	0.02085	0.00943	2.21	0.027	1.021	Yes
5	SPEED	0.06958	0.00996	6.99	<.001	1.072	Yes

Indeed, each road segment has a different length. In this model, the differences in the road length affect the number of motorcycle accidents. For example, if the difference in length between two sections of roads doubled up, then the model would predict an increasing number of motorcycle accidents by 39.2%. The influence of arterial roadlength differences on motorcycle accidentsis in accordance with the findings in previous studies [15].

The number of access points on every arterial road is certainly not the same. There are many roads that have access points, but there are also roads that have a slight number of access points. In this model, the differences in the number of access points per kilometer between each arterial road affect the number of motorcycle accidents. For example, in each additional 5 access points, per-kilometeron arterial roads, there will be an increasing number of motorcycle accidents by 9.9%. The influence of the number of access points on motorcycle accidents is in accordance with the findings in previous studies [16, 17, 18].

Finally, the addition of traffic speeds on arterial roads will have a significant effect on the increasing number of motorcycle accidents. For example, any increase of speed of 5 km / h on arterial roads, there will be an increasing number of motorcycle accidents by 29.38% [19, 20].

CONCLUSIONS

Based on the above discussion it can be concluded, as follows:

- The end result of the predictive modeldevelopment of motorcycle accidents occurred on arterial roads, as follows:
- Motorcycle accident prediction model which is generated in this study may explain the potential of motorcycle accidents on arterial roads, which are influenced by: traffic volume (FLOW), the roadlength (LR), the number of access points (APs) and traffic speed (SPEED).
- The increasing volume of traffic, lengthier road, number of access points per kilometer and increasingspeed of trafficwill lead to the increasing risk of motorcycle accidents on arterial roads.

The findings of this model development in this study provides information to assist the stakeholders in the field of traffic engineering, to overcome the problems related to motorcycle accidents on arterial roads in Surabaya and other places. However, the final model may only apply to the typical traffic conditions in developing countries such as Indonesia, where the percentage of motorcycles dominates the total vehicles by 60% to 80%.

ACKNOWLEDGMENTS

This paper reports some findings that are part of a research dissertation in the Doctoral Program in Civil Engineering, Faculty of Engineering, University of Brawijaya, Malang, Indonesia. We are grateful for the financial support from the Ministry of Education and CultureRepublic of Indonesia.

REFFERENCES

1. Nermin, M.M. and K.A. Ahmed, 2011. Landscape Impact on Roadside Improvement in Egypt Case Study of Salah Salem Road, Cairo, Egypt. World Applied Sciences Journal, 12(3): 266-278, 2011, ISSN 1818-4952, ©IDOSI Publications.
2. Direktorat Jenderal Perhubungan Darat, 2012. Perhubungan Darat Dalam Angka (PDDA). Jakarta.
3. Archer, J. and K. Vogel, 2000. “The Traffic Safety Problem in Urban Areas.”, Research Report CTR2000-04, Royal Institute of Technology, Stockholm, Sweden.
4. Widyastuti, H. and C. Mulley, 2005. Evaluation of Casualty Cost of Motorcyclist’s Slight Injury in Indonesia, Journal of the Eastern Asia Society for Transportation Studies, 6: 3497-3507.
5. Natalya, S.P., I.A Alexandr and M.U. Sergey, 2013. The Influence of Active Safety Characteristics of a Bus on Driver Working Conditions (By the Example of Vladivostok, Russia), World Applied Sciences Journal, 26(3): 318-321, 2013, ISSN 1818-4952, ©IDOSI Publications.
6. Keshava, Murthy G.N. and Zaved Ahmed Khan, 2013. Different Techniques to Quantify the Driver Alertness, World Applied Sciences Journal, 22(8): 1094-1098, 2013, ISSN 1818-4952, ©IDOSI Publications.
7. Nambuusi, *et al*., 2008. A review of accident prediction models for road intersections. Steunpunt Mobiliteit & Openbare Werken – Spoor Verkeersveiligheid, RA-MOW-2008-004
8. Olescenko, E. and V. Dobromirov , 2013. Model of the Projected Development of Road Safety in Megapolis (On the Example of Saint Petersburg), World Applied Sciences Journal 23(Problems of Architecture and Construction): 84-88, 2013, ISSN 1818-4952 ©IDOSI Publications.
9. Davati, A., Z. Rahmati and H.V. Arani, 2012. The Factors Related to Tehran Province Traffic Accident Mortalities in 2011, a Descriptive Study, World Applied Sciences Journal, 19(11): 1681-1686, 2012, ISSN 1818-4952, ©IDOSI Publications.
10. McCullagh, P. and J.A. Nelder. 1989. Generalized Linear Models, 2nd edition. London, England: Chapman and Hall.
11. Harnen, S., R.S. Radin Umar, S.V. Wong and W.I. Wan Hasim, 2006. Motorcycle Accident Prediction Models for Junctions on Urban Road in Malaysia. Journal of Advances in Transportation Studies, Section A, 8: 31-40.
12. Sawalha, Z. and T. Sayed, 2006. Traffic accident modeling: some statistical issues, Canadian Journal of Civil Engineering, 33: 1115-1124.
13. Xin Pei, S.C. Wong and N.N. Sze, 2011. A joint-probability approach to crash prediction models. Accident Analysis and Prevention, 43: 1160-1166.
14. Polus, A. and M. Cohen 2012. A new, non-canonical Poisson regression model for the prediction of crashes on low-volume rural roads. IATSS Research, 35(2): 98-103.
15. Chengye, P. and P. Ranjitkar, 2013. Modelling Motorway Accidents using Negative Binomial Regression, Proceedings of the Eastern Asia Society for Transportation Studies, pp: 9.
16. Tarko, A.P., S. Eranky, K.C. Sinha and R. Scinteie, 1999. An Attempt to Develop Crash Reduction Factors Using Regression Technique. The 78th Annual Meeting of Transportation Research Board, Washington D.C.
17. SLevinson, H.S., 2000. Access Spacing and Accidents, A Conceptual Analysis.In Transportation Research Circular E-C019: Urban Street Symposium Conference Proceedings, Dallas, TX, June 28-30, 1999. Washington, DC: TRB, NRC, December 2000.
18. Papayannoulis, V., J.S. Gluck, K. Feeney and H.S. Levinson. 2000. Access Spacing and traffic Safety. In Transportation Research Circular E-C019: Urban Street Symposium Conference Proceedings, Dallas, TX, June 28-30, 1999. Washington, DC: TRB, NRC, December 2000.

19. Ackaah, W. and M. Salifu, 2011. Crash prediction model for two-lane rural highways in the Ashanti region of Ghana. IATSS Research, 35: 34-40.
20. Taylor, M.C., A. Baruya and J.V. Kennedy, 2002. The Relationship Between Speed and Accidents on Rural Single-Carriageway Roads, Report TRL 511. Berkshire, England: TRL Limited.