Analysing the Frequency of Traffic Crashes in Riyadh City Using Statistical Models and Geographical Information Systems (GIS)

Dr.Saleh Abdulaziz Altwaijri

Contact Number: 07889388865
saltwaijri@yahoo.com
saltwaijri@hotmail.com

ABSTRACT

Traffic crashes in Riyadh city cause losses in the form of deaths, injuries and property damages, in addition to the pain and social tragedy affecting families of the victims. The primary objective of this paper is therefore to explore factors affecting the frequency of road crashes in Riyadh city using appropriate statistical models and GIS approach to integrate the datasets and to calculate the yearly crashes per spatial unit aiming to establish effective safety policies ready to be implemented to reduce the frequency of road crashes in Riyadh city. Crash data for Riyadh city were collected from the Higher Commission for the Development of Riyadh (HCDR) for a period of five years from 1425H to 1429H (2004-2008). A negative Binomial (NB) model was employed and the units of analysis were 168 HAIs (wards) in Riyadh city. The results from the frequency model suggest that population is positively significant with the frequency of fatal and serious injury crashes (at the 99% confidence level). Percentage of illiterate people and the income per capita found to be positively significant with the frequency of fatal crashes; and the increased residential, transport, and educational areas of land use is associated with the decreased level of fatal and serious crashes occurrences. Based on the findings, a range of countermeasures are proposed to reduce the frequency of traffic crashes in Riyadh city.

KEYWORDS: Traffic safety, Riyadh city, Crash frequency, GIS, Negative Binomial model.

1. Introduction

Riyadh, the capital of Kingdom of Saudi Arabia (KSA), is one of the fastest growing cities in the Middle East. Riyadh has experienced a very high rate of population growth as its population was 150,000 in the 1960s, over 4.5 million in 2005 and is expected to reach 10.5 million by 2020 (HCDR, 2008). This tremendous growth in population creates a high level of mobility and transport activities in the city. In 2005, there were about six million trips generated per day in Riyadh city. This is predicted to rise to about 15 million trips per day by 2020 (SMOT, 2007).

In 2005, there were a total of 47,341 injury traffic crashes in Riyadh (19% of the total KSA crashes). Previous studies have highlighted traffic safety as a serious issue for Riyadh and there is an urgent need to develop safety policies aimed at reducing both traffic crashes (Lee, 1986; Koushki and Al-Anazi, 1998; Al-Ghamdi, 1996a; Al-Ghamdi, 1996b; Al-Ghamdi, 1999).

Because of the negative impact of traffic crashes, it is important to carry out a careful investigation to understand the relationship between traffic crash frequency and their contributing factors aiming to establish effective safety policies ready to be implemented to reduce the frequency of road crashes in Riyadh city.

2. Data

Riyadh is selected because of the availability of crash data that can be geo-coded. According to the Riyadh Municipality, Riyadh contains 15 Municipalities and is divided into 130 districts. A lower-level administrative boundary called 'Hai' (ward) is however used as a unit of observation in the analysis as most of the data are available at Hai-level. The study area contains 169 spatial units (see Figure 1) and the average area of each unit is about 20 km^2.

It can be noted from Figure 1 which is a thematic map of the spatial distribution of crashes for Riyadh region and Riyadh city that HAIs with more roads have more crashes and these are concentrated on the Centre of Riyadh city.

However, studies from developing countries including KSA are not based on comprehensive crash datasets and do not employ appropriate statistical techniques or GIS approach. Therefore, factors affecting frequency were not conclusively identified and effective safety policies could not be formulated. This paper attempts to fill this gap by investigating the national crash dataset for KSA with the aid of GIS and applying appropriate statistical models.

In order to develop relationships between area-wide (i.e. Hai-level) crash frequency (per unit time) and area attributes such as socio-economic factors, road network, land-use patterns, crash data from 2004 to 2008 were obtained from (HCDR). These crashes are then geo-coded in GIS for the purpose of calculating yearly crashes per spatial unit suggesting that spatial-level crash counts are only possible for fatal crashes.

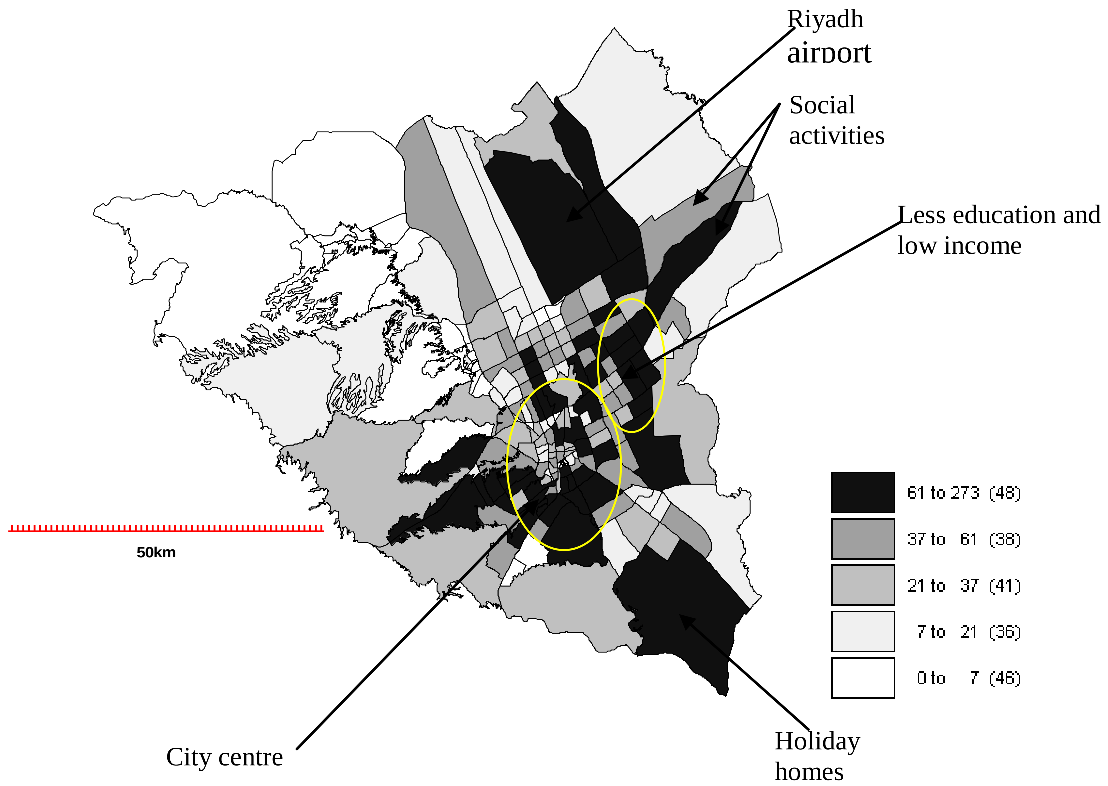

Figure 1. Spatial distribution of crashes for Riyadh region and Riyadh city.

Geographical information system (GIS) technology is becoming increasingly popular tool for visualisation and analyses of crash data in motorways. GIS has the ability to hold a vast amount of data that can be easily stored, shared, analysed and managed. It provides a platform for spatial data analyses and visualization to explore relationships between spatial and non-spatial data (Erdogan et al., 2007). When using a GPS device, one can obtain (x,y) coordinates at regular time steps. The trajectory is then automatically recorded, by joining the locations collected in time, and geo-referenced into a GIS. Different digital GIS maps can be plotted in different layers in the GIS software (e.g. roads, crashes), and GIS software (i.e. MapInfo in this case) can display them at the same time in one view, with each layer to be displayed in a desired way, such as making lines (roads) look wider. An example is shown in Figure 2, which represents a multi layered GIS map of Riyadh city, it integrates road networks as lines, road crashes as dots, land use and boundaries of HAI.

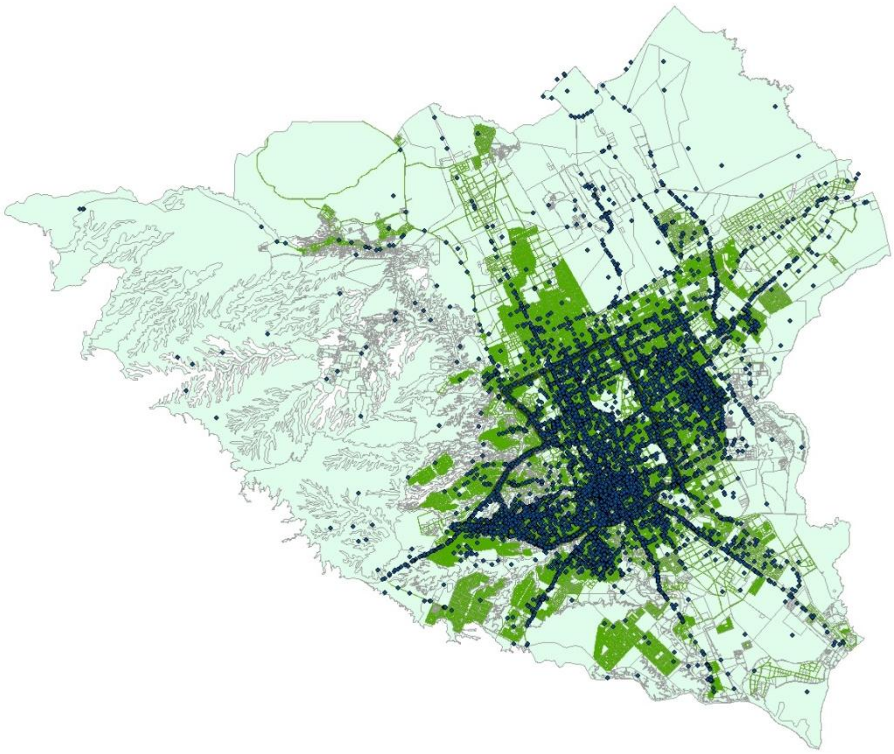

Figure 2 Example of GIS maps in multiple layers

Many researchers have used GIS to display crash locations on digital maps and perform various spatial analyses (including hot spot analysis) of crashes (Petch and Henson, 2000; Wang et al., 2009). In addition, GIS enables researchers to link crash data with travel information, land use, and socio-economic information to better capture the relationship between crash occurrence and contributing factors. There are several ways to locate crashes onto digital maps. Crashes can be directly added if the exact geographic references of crash locations are known, such as easting and northing coordinates which can be obtained by a GPS device. Address geocoding can also be conducted when the exact address (e.g., street name and number, city, state, post code) is available.

In terms of map construction MapInfo tool developed by ESRI was used for GIS modelling by cross mapping (overlaying) data. There are several ways to locate crashes onto digital maps; crashes can be directly added if the exact geographic references of crash locations, like coordinates, are available. First, base digital map was obtained from HCDR, which includes land use and road network. The next step was adding the geocoded data of the crash locations.

In this research, land-use and road network data were in a GIS format. The geo-coding of crash data allowed integration with other GIS datasets such as land use for frequency analysis. Figure 2 showed the spatial distribution of geo-coded crashes on the Riyadh region electronic map and Riyadh city marked on the map. Furthermore, road density data are calculated at spatial levels of HAI, the equivalent of a ward in England, and crash frequency per HAI are then determined and integrated with the HAI-level road density data using GIS.

The results and interpretations of NB models for fatal and serious injury crashes from this estimated model are presented below:

3. Models estimation results

Table 1 shows the estimation results of NB model for fatal crashes of the three models. Population is positively significant with the frequency of fatal injury crashes (at the 99% confidence level) in all the three models and percentage of illiterate people is also positively significant in model 1 (marginally at 90% confidence level). Percentage of residential and percentage of transport are negatively significant (at 99% confidence level) in all three models whereas percentage of educational areas is negatively significant (at 95% confidence level) in models 1 and 2 and (at 90% confidence level) in model 3. This means the more population and more illiterate people in the HAI; the more fatal crashes in that HAI and when the percentage of all types of land use (residential, transport, and educational land use) increase, fatal crashes will decrease.

Table 1 shows that model 1 and model 3 are very similar in the maximum likelihood, the values of R^2, and in AIC value which means that both models could be the preferred models as an NB models for fatal crashes. Model 2 could be also a preferred model due to that its values are close to those on model 1 and model 2.

Table 2 shows the estimation results of NB model for serious injury crashes of the three models. Population is positively significant with the frequency of serious crashes (at the 99% confidence level). Percentage of non-Saudi people and income per capita in model 1 are positively significant (at 85% confidence level), and percentage of illiterate people is positively significant (at 90% confidence level) whereas all types of land uses (residential, transport and educational areas of land use) are negatively significant (at 99, 95, marginally at the 85% confidence level respectively) in model 1. In model 2 percentage of non-Saudi is positively significant (at the 85% confidence level) whereas all types of land uses (residential, transport and educational areas of land use) are negatively significant (at 99%, 95%, marginally at the 85% confidence level respectively). Residential and transport utilities land uses are found negatively significant (at 99% and 90% confidence level respectively) in model 3 where as it is found that percentage of older people, income per adult, and percentage of low income insignificant. This means that when population, percentage of non-Saudi people, illiterate people and income per capita increase in the HAI, serious injury crashes will also increase in that HAI and when percentage of all types of land use (residential, transport, and educational land use) increase serious injury crashes will decrease.

Table 1. NB models for fatal crashes on Riyadh city

Variable	Model 1		Model 2		Model 3	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Log (population)	0.3244	4.81(2)	0.052436	4.72(2)	0.322312	4.80***
Percentage of non-Saudi	0.0006	0.17	0.04179	0.23	-0.00464	-0.82
Percentage of older people age 65+	0.0452	0.66	6.45E-05	0.77	0.055751	0.82
Percentage of illiterate people	0.0502	1.64	-0.02551	1.37	0.0356	1.40
Income per capita	0.0002	1.09	-	-	-	-
Income per adult	-	-	6.45E-05	0.61	-	-
Percentage of low income	-	-	-	-	0.009784	1.10
Percentage of residential	-0.0260	-3.57***	-0.02551	-3.5(3)	-0.02763	-3.71***
Percentage of transport utilities	-0.0759	-2.89(3)	-0.07304	-2.78(2)	-0.06562	-3.57(2)
Percentage of educational	-0.05597	-2.24(2)	-0.05655	-2.23(2)	-0.04583	-1.84(1)
Constant	-1.1662	-1.67(1)	-1.00972	-1.42	-1.00938	-1.57
Over-dispersion parameter	0.3159	5.05(3)	0.31895	5.07(3)	0.315754	5.04(2)

Models statistics:			
Log-likelihood	-346.2337	-346.6533	-346.2433
Pseudo R^2	0.0918	0.0907	0.0918
AIC value	712.4674	713.3067	712.4865
Observations (N)	131	131	131

(1) Statistically significantly (at 90% confidence level) critical t=1.65
(2) Statistically significantly (at 95% confidence level) critical t=1.96
(3) Statistically significantly (at 99% confidence level) critical t=2.58

Table 2. NB models for serious injury crashes on Riyadh city

Variable	Model 1		Model 2		Model 3	
	Coefficient	t-stat	Coefficient	t-stat	Coefficient	t-stat
Log (population)	0.4755	9.26***	0.4661	9.12***	0.4623	9.12***
% of non-Saudi	0.0040	1.52	0.0042	1.54	0.0012	0.31
% of older people age 65+	0.0420	0.72	0.0530	0.91	0.0595	1.01
% of illiterate people	0.0452	1.69*	0.0343	1.31	0.0215	1.00
Income per capita	0.0002	1.56	-	-	-	-
Income per adult	-	-	8.05E-05	0.94	-	-
% of low income	-	-	-	-	0.0049	0.81
% of residential	-0.0190	-3.6***	-0.0186	-3.48***	-0.0198	-3.6***
% of transport utilities	-0.0377	-2.17**	-0.0346	-1.99**	-0.0302	-1.76*
% of educational	-0.0229	-1.62	-0.0230	-1.61	-0.0181	-1.22
Constant	-1.5942	-2.91***	-1.4239	-2.54**	-1.2255	-2.56**
Over-dispersion parameter	0.2724	6.76***	0.2761	6.78***	0.2768	6.79***
Models statistics:					-1.05687	
Log-likelihood	-506.8352		-507.6501		-507.7723	
Pseudo R^2	0.0814		0.0799		0.0797	
AIC value	1033.67		1035.30		1035.545	
Observations (N)	131		131		131	

(1) Statistically significantly (at 90% confidence level) critical t=1.65
(2) Statistically significantly (at 95% confidence level) critical t=1.96
(3) Statistically significantly (at 99% confidence level) critical t=2.58

It can be noticed from Table 2 that model 1 has the highest values of maximum likelihood, values of R^2, and the lowest value of AIC value which means that model 1 could be the preferred models (NB models for serious injury crashes with income per capita).

4. Discussions and Conclusions

This paper has examined a range of factors affecting traffic crash frequency in Riyadh city using Negative Binomial (NB) models.

Broadly speaking, population variable revealed the positive sign suggesting that the increased population is associated with the increased level of fatal and serious injury crashes occurrences. NB models revealed the negative sign for all types of land use in all models suggesting that the increased residential, transport, and educational areas is associated with the decreased level of fatal and serious injury crashes occurrences. Percentage of non-Saudi to be statistically insignificant for fatal crashes

This study is the first attempt to examine spatial variation in crashes in Saudi Arabia. Findings from this study can be useful in formulating safety policies aimed at reducing the occurrence of traffic crashes.

In terms of future research, separate crash prediction models need to be developed for various road users such as pedestrians to reinforce some of the findings of this study.

5. Acknowledgement

The author would like to thank the High Commission for the Development of Riyadh (HCDR) and the Riyadh Traffic Department (RTD) for providing the crash data used in this study.

References

AL-GHAMDI A.S. (1996a) Road accidents in Saudi Arabia: a comparative and analytical study *Proceedings of the Presentation at the 75th Annual Meeting of the Transportation Research Board. Washington, DC.*

AL-GHAMDI, A.S (1999) Injury severity and duration of hospital stay for urban road accidents in Riyadh *Transportation Research Record No. 1665 TRB National Research Council, Washington DC* pp 125–132.

AL-GHAMDI, A.S. (1996b) Road accidents in Saudi Arabia: a comparative and analytical study *Proceedings of the Second International Conference on Urban Transport and the Environment for the 21st Century. Barcelona, Spain* pp 231–253.

ERDOGAN, S., YILMAZ, I., BAYBURA, T., & GULLU, M. (2007). Geographical information systems aided traffic accident analysis system case study: City of Afyonkarahisar. Accident Analysis and Prevention. Volume 40, Issue 1, January 2008, Pages 174-181.

HCDR (Higher Commission for the Development of Riyadh) (2008) *Phase III of ADA Strategic Traffic Safety Study Report* pp.40-42 Available at: *www.arriyadh.com/ADA-Emp/StudiesDep/Traffic-Safety-for-Arriyadhhase-I2.doc*

KOUSHKI, P & AL-ANAZI, S (1998) *Seat Belt Use Trend in Kuwait* Research Report CE-499 Department of Civil Engineering, Kuwait University Kuwait.

PETCH, R.O., AND HENSON, R.R. (2000) ‘Child road safety in the urban environment’, Journal of Transport Geography, 8, 197–211.

LEE K.W (1986) An analysis of automobile accidents in Riyadh *ITEJ Institute of Transportation Engineers Washington DC* pp 35–39.

SMOT (Saudi Ministry of Transport) (2007) *National Transport Strategy Report* Available at: http://www.mot.gov.sa/

WANG, C., QUDDUS, M.A., ISON, S. G., (2009) ‘Impact of traffic congestion on road accidents: A spatial analysis of the M25 motorway in England’. Accident Analysis and Prevention, 41(4), 798-808.