Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.4, No.13, 2014 www.iiste.org

Modeling Road Traffic Fatality Cases in Ghana.

Adu-Poku, K.A. [1*], Avuglah, R.K. [2], and Harris, E.[3]

[1]Department of Mathematics, Kwame Nkrumah University of Science and Technology Kumasi, Ghana
[2]Department of Mathematics, Kwame Nkrumah University of Science and Technology Kumasi, Ghana
[3]Department of Mathematics, Kwame Nkrumah University of Science and Technology Kumasi, Ghana
[*]Corresponding Author's Email: adupoku2009@yahoo.com

Abstract

Road traffic fatalities are of major concern in Ghana. According to the annual report of National Road Safety Commission (NRSC), 2,047 road traffic fatality cases were recorded in the year 2007 which is an increase of 29.5% over that of the year 2000 (NRSC, 2007 Report). Relative to the year 2001(NRSC, 2001) the 2010 figures for fatal crashes and fatalities also recorded corresponding increases of 34.1% and 19.6% respectively. Using annual road fatality data for the period 1991-2011 from the National Road Safety Commission and Building and Road Research Institute (BRRI) in Ghana, the Autoregressive Integrated Moving Average (ARIMA) model has been utilized to study the trend and pattern of road fatalities in Ghana. The results showed that road traffic fatalities in the country increased over the period of study. Moreover, a forecasting approach indicates that road traffic fatalities would continue to increase over the next five years.

Keywords: Accident, Autoregressive (AR), Moving average (MA) and ARIMA.

1.0 Introduction

It is estimated that exactly 3,000 people die as a result of road traffic accidents all over the world on a given day. (Penden et al., 2004). Over 90% of the world fatalities on roads occur in low-income and middle-income countries (WHO, 2009). However, middle income countries have the highest annual fatality rates at 20.1 per 100,000 while high income countries have the lowest fatality rate at 8.7 per 100,000 (WHO, 2013). In 1999, it was estimated that between 750,000 and 880,000 people died from road crashes and that the majority of these deaths occurred in developing and transitional nations. (G Jacobs et al., 2000). Statistical projections have shown that the highest fatality rates (death per 10,000 motor vehicles) worldwide occur in African countries particularly Ethiopia, Uganda and Malawi. (G Jacobs et al., 2000).

According to Traffic Statistics, road traffic fatalities in the member states of the European Union annually claim about 43,000 lives and leave more than 1.8 million people injured representing an estimated cost of 160 billion euros (European Road Safety Observatory, 2008). In an effort of halving the number of road traffic fatalities by 2010, the European Commission published a white paper on transport policy on 12th September 2001 (European Commission , 2001). "Time to decide" was the subtitle of the white paper and this subtitle was chosen because the commission had the view that the time had come as far as road safety was concerned (European commission, 2001 memo).

Pedestrian death remains the leading effect of fatalities among urban road users in Ghana (James Damsere et al., 2010). At least 1,800 deaths are recorded every year and 6 people die every day as a result of road traffic crashes in Ghana (NRSC, 2013). Despite increased Road Safety Campaigns, the rate at which accidents occur on our roads is very alarming. Annual report from the National Road Safety Commission indicates that the total number of road traffic crashes in 2007 was 12,038 representing an increase of 3.2% over the 11,668 recorded in 2006 and 5.9% over the 2005 figure of 11,328 (NRSC, 2007 report). Similarly, road traffic fatalities have increased accordingly over the years. Relative to the year 2001, the 2010 fatality figure of 1,986 deaths represented an increase of 19.6% indicating an upward trend (BRRI, NRSC, 2010 report).

Road traffic fatalities occur as a result of several factors such as over speeding, poor vehicle maintenance, human errors and pedestrian activity (Aworemi et al., 2010, Shinar 1978, Afuukar et al., 2003). Studies have indicated that road side activities such as jaywalking, nighttime walking and street walking are also associated with fatality cases in Ghana (James Damsere et al., 2010).

Several researchers have used different statistical models in analyzing road traffic fatalities. For instance, Chang and Mannering (1998) used nested logit model to analyze road fatalities and concluded that fatality is worsened if the accident has a truck involved. Multinomial logistic regression model was used by (Bedard et al., 2002) to analyze road fatalities and the results showed that female drivers with blood alcohol concentration greater than 0.3 were associated with high fatality rates. Similarly, the logistic regression model was also used by (Ballesteros et al., 2004, Plurad et al., 2006, Macleod et al., 2010 and Mouden et al., 2009) to analyze road fatalities.

Correspondingly, the studies of accident variables including injuries and fatalities have extensively been considered within the framework of time series modeling. Using ARIMA time series model, (Yuan Yuan pack et al., 2013) indicated that China's accident cases would increase from 107,579 to 401,536 over a period of one year. (Rohayu Sarani et al., 2012) used ARIMA model to forecast Malaysia road fatalities for the period 2015-2020 and the results showed that the predicted fatalities for the year 2015 will be 8,760 and 10,716 in 2020. Similarly, time series analysis has also shown that road traffic fatalities are on the decrease with the exception of Lagos Island local area (Atubi et al., 2013). The aim of the present study is to study the trends, patterns and forecasts of road traffic fatalities in Ghana. The data set used is mainly road traffic fatalities for the period 1991-2011 and was obtained from the National Road Safety Commission and Building and Road Research Institute in Ghana.

2.0 Time Series

Time series is a sequence of observations ordered by a time parameter. Mostly these observations are collected at equally spaced discrete time intervals.

2.1 Components of time series

- Trend: When there is a long term increase or decrease in a data.
- Seasonal: The periodic fluctuations in a time series within a certain time period.
- Cyclical: Long departure from the trend due to factors other than seasonality.
- Irregular: The movement or component left after accounting for trend, seasonal and cyclical movement in a time series.

2.2 Stationary process

A stochastic process $\{X_t\}$ is said to be strictly stationary if the joint distribution of $X_{t1}, X_{t2}, X_{t3}, \ldots\ldots X_{tn}$ is the same as the joint distribution of $X_{t1-k}, X_{t2-k}, X_{t3-k}, \ldots\ldots X_{tn-k}$ for all $t_1, t_2, t_3, \ldots . t_n$.

Mathematically,

- $E(X_t) = E(X_{t-k})$
- $Var(X_t) = Var(X_{t-k})$ for all t and k.
- $Cov(X_t, X_s) = Cov(X_{t-k}, X_{s-k})$ for all t, s and k.

2.3 Differencing

Consider a time series Y_t, first order differencing is defined as $y^1{}_t = y_t - y_{t-1}$

We can use backshift operator to express differencing as follows $y^1{}_t = y_t - y_{t-1} = y_t - By_t = (1-B)y_t$ and second order differencing is also expressed as $y^{11}{}_t = y_t - 2y_{t-1} + y_{t-2}$. In terms of backshift operator, the equation becomes $y^{11}{}_t = y_t - 2y_{t-1} + y_{t-2} = y_t - 2By_t + B^2 y_t = (1 - 2B + B^2)y_t$

For example, $BY_t = Y_{t-1}$ or $Be_t = e_{t-1}$ and $B^2 e_{t=} e_{t-2}$.

ARIMA (1,0,0) can be expressed in terms of backshift operator as $y_t = \emptyset_1 y_{t-1} + e_t + u$ or $y_t - \emptyset_1 y_{t-1} = e_t$ but $By_t = y_{t-1}$ put into the equation, we get $y_t - \emptyset_1 By_t = e_t$ or $y_t(1 - \emptyset_1 B) = e_t$

Similarly, ARIMA (2,0,0) can be written as $(1 - \emptyset_1 B - \emptyset_2 B^2)Y_t = e_t$
$y_t - y_{t-1} = \emptyset_1(y_{t-1} - y_{t-2}) + \emptyset_2(y_{t-2} - y_{t-3}) + e_t = (1 - B)(1 - \emptyset_1 B - \emptyset_2 B^2)Y_t = e_t$.

2.4 Integrated ARMA (or ARIMA models)

Autoregressive Integrated Moving Average (ARIMA) model is a generalization of an Autoregressive Moving Average (ARMA) model. The model is generally referred to as an ARIMA (p, d, q) model where p, d and q are integers greater than or equal to zero. The first parameter (p) refers to the number of autoregressive lags, the second parameter (d) refers to the order of integration and the third parameter (q) gives the number of moving average lags.

- Suppose a series { Y_t } is non-stationary, if the dth difference of the series is a stationary ARMA(p , q) process, then { Y_t } is said to be an ARIMA(p, d, q) process. The general Autoregressive integrated moving average (ARIMA) is of the form:
- $W_t = \sum_{i=1}^{p} \phi_i W_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \mu + \varepsilon_t$
  $where \quad W_t = \Delta^d Y_t = (1 - B)^d Y_t$

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \ldots\ldots\ldots\ldots ARIMA(1,0,0)$$
$$y_t = \phi_1 y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \ldots\ldots\ldots\ldots ARIMA(1,0,1)$$
$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1} \ldots\ldots\ldots\ldots ARIMA(2,0,1)$$
$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \ldots\ldots\ldots\ldots ARIMA(2,0,2)$$
$$y_t = y_{t-1} + \phi(y_{t-1} + y_{t-2}) + \varepsilon_t \ldots\ldots\ldots\ldots ARIMA(1,1,0)$$

2.5 Box Jenkins Method

This method proposes three stages namely; identification, estimation and diagnostics checking.

The first stage in building the model is the identification of the appropriate ARIMA models through the study of the autocorrelation and partial autocorrelation function. The sample autocorrelation plot and sample partial autocorrelation plot are compared to the theoretical behavior of these plots when the order is known.

Other tools for model identification are:

- Akaike Information Criterion (AIC)
- Akaike Information Criterion Corrected (AIC_C)
- Bayesian Information Criterion. (BIC).

The BIC imposes a greater penalty for the number of estimated model parameters than AIC. The use of minimum BIC for model selection results in a chosen model whose number of parameters is less than that chosen under AIC.

The next stage is to estimate the parameters of the ARIMA model chosen. The nonlinear least squares and maximum likelihood method is normally used.

The third stage is the diagnostics checking of the model. The Q statistics is used for the model adequacy check.

The fourth stage is where the analyst uses the model chosen to forecast and the process ends.

Mathematical Theory and Modeling
ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online)
Vol.4, No.13, 2014

3.0 Analysis and results

3.1 Preliminary Analysis

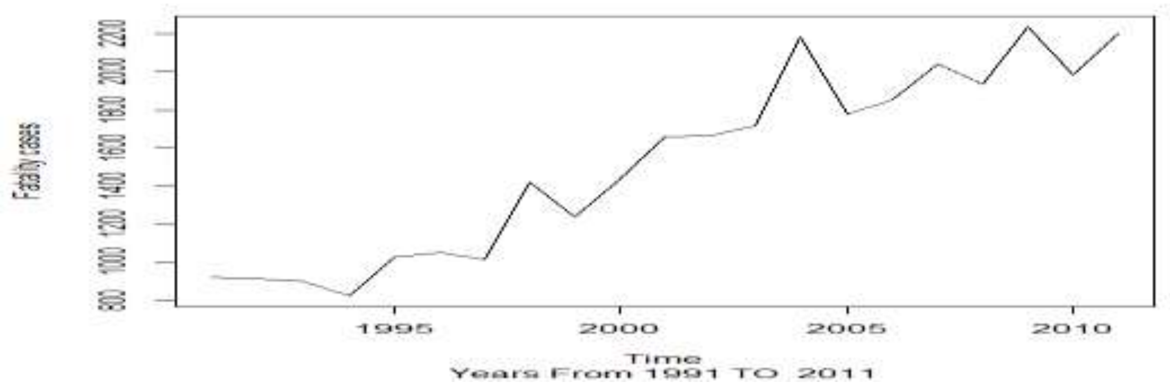

Fig 1: Time plot of fatality cases in Ghana from 1991 to 2011.

Fatalities in Ghana decreased from 1991 to 1994 and slightly increased from 1995 to 1996. It then declined from 1997 to 1998 but rose sharply from 1999 to 2004. An irregular pattern was then observed from 2005 to 2011. In general, the trend of fatality cases in Ghana was on a rise but not consistent. The fatality time plot in fig 1 was not stationary due to the trend component.

3.2 First differencing of fatality cases

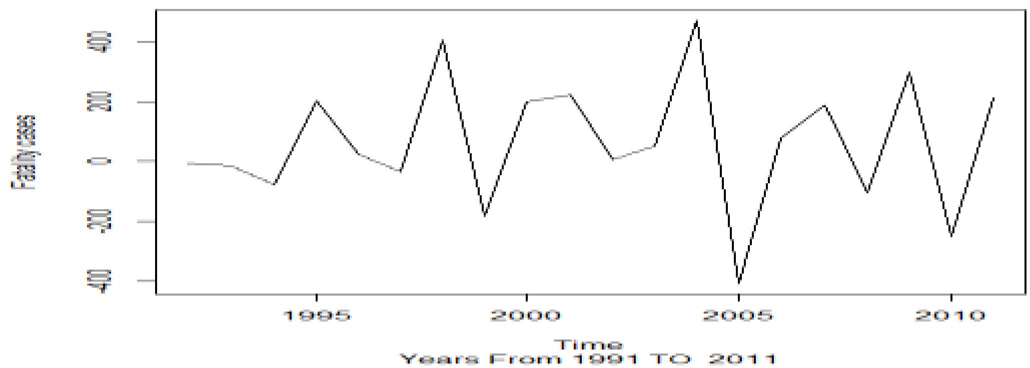

Fig 2: First differencing of fatality cases in Ghana.

First differencing was performed to remove trend component in the original data

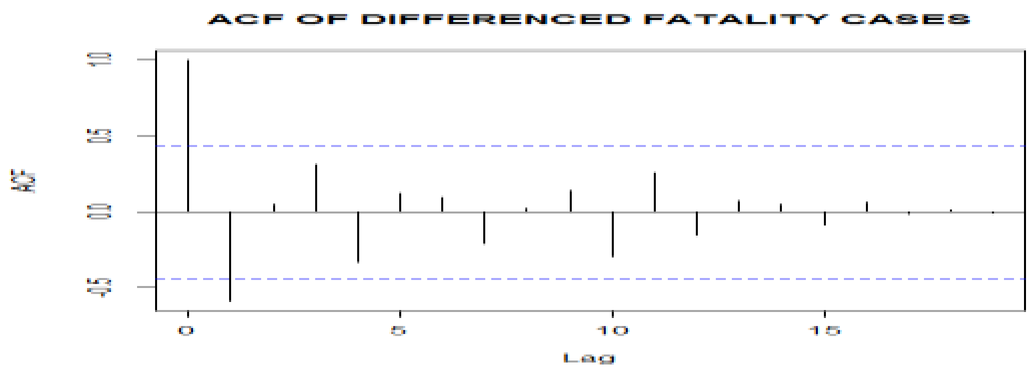

Fig 3: ACF plots of the differenced fatality cases in Ghana.

From the correlogram, the sample ACF tails off to zero after the first significant lag. That is lag 1 exceeded its significant bound and there after dies off exponentially representing a moving average of 1 (MA 1).

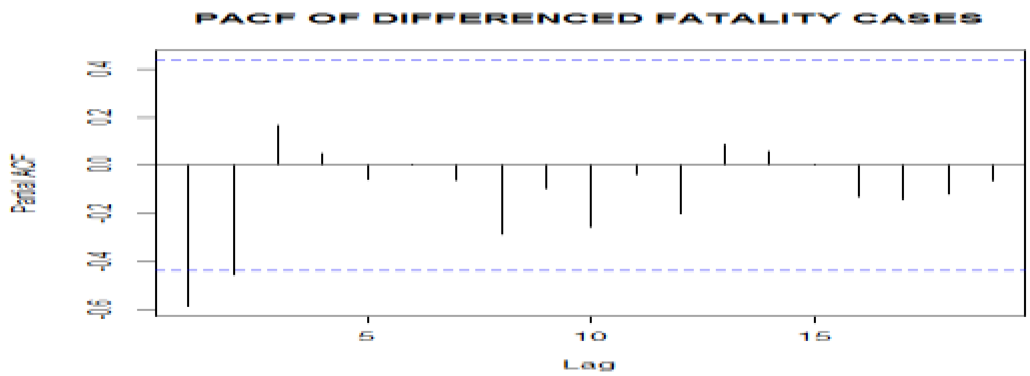

Fig 4: PACF plots of the differenced fatality cases

The partial Autocorrelation function tails off to zero after the first two significant lags. Comparing with the error limits, we could see that lag 2 was significant, meaning an Autoregressive model of 2 (AR 2). From the foregoing analysis, the following ARIMA models are therefore possible for the data series.

ARIMA(2,1,1)

ARIMA(2,1,0)

ARIMA(0,1,1)

PARAMETER ESTIMATION

TABLE I : ARIMA (2,1,1) MODEL

COEFFICIENT	ESTIMATE	STD. ERROR	T-VALUE	AIC	BIC
AR 1	1.0545	0.3405	3.09	11.37	10.57
AR 2	0.5820	0.2292	2.54		
MA 1	0.2749	0.4242	0.64		

The t-value for MA 1 is not statistically significant while that of AR 1 and AR 2 are statistically significant since the t-value is greater than 2 in absolute terms. (Table I).

TABLE II : ARIMA (2,1,0) MODEL

COEFFICIENT	ESTIMATE	STD. ERROR	T-VALUE	AIC	BIC
AR 1	0.8483	0.1942	4.37	11.29	10.45
AR 2	0.4604	0.1980	2.33		

From table II, the parameter based on the t-value test is statistically significant.

TABLE III: ARIMA (0,1,1) MODEL

COEFFICIENT	ESTIMATE	STD. ERROR	T-VALUE	AIC	BIC
MA 1	1.000	0.2025	4.94	11.20	10.30

The parameter based on the t-value test is statistically significant in table III.

3.3 Model Diagnostics

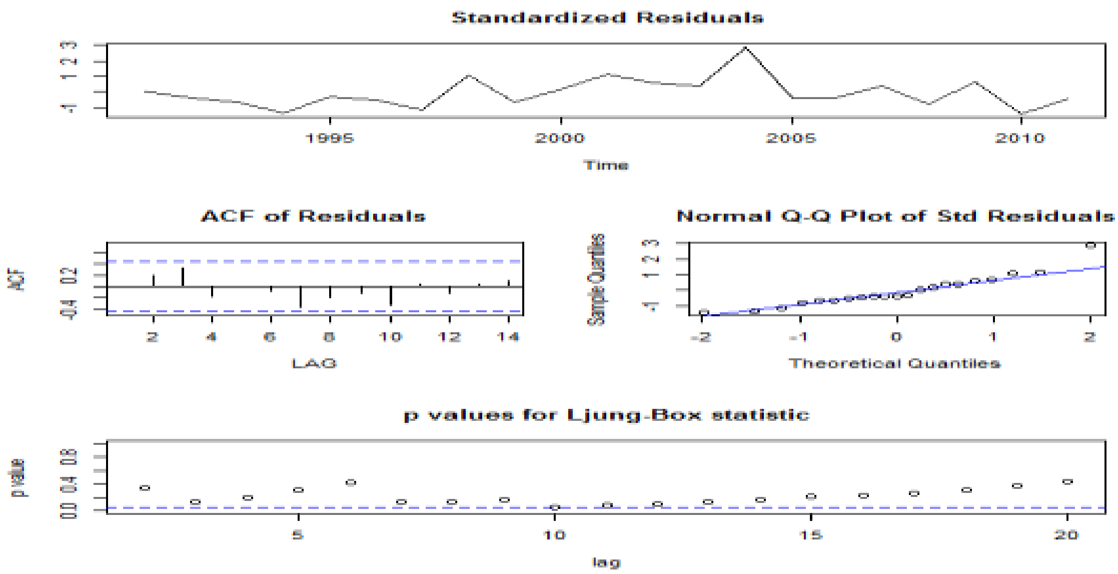

Fig 5: Diagnostics of the residuals from ARIMA (0, 1, 1)

From figure 5, it is evident that there was no obvious trend and pattern for the standardized residuals plot. The ACF of the residual plot shows that none of the lags exceeded its significant bound. Most of the residuals do not deviate significantly from the line of best fit and its distribution looks approximately linear. Lastly, the Ljung-Box statistics plot was not significant at any positive lag. In conclusion, the model was adequate and fits well. It was also observed that ARIMA (2,1,1) and (2,1,0) models exhibited similar diagnostic characteristics as ARIMA (0,1,1).

3.4 Selection of best model for forecasting fatality cases in Ghana.

The standardized residual plots of all the models were independently and identically distributed with a mean of 0 and some few outliers. There was no evidence of significance in the autocorrelation functions of the residuals of all the models and the residuals appear to be normally distributed in all the models. The Ljung – Box statistics was not significant at any positive lag for all the models. The AR and MA parameters in ARIMA (2,1,1) were not significant at 5% level of significance which could have a negative effect on the forecast if used for prediction. The parameters in ARIMA (2,1,0) and (0,1,1) were however significant.

The AIC, AIC_C, and BIC were good for all the models but they favored ARIMA (0, 1, 1) model since it had the minimum AIC value.

From the discussion above, it is evident that ARIMA (0, 1, 1) model was the best model for forecasting fatality cases in Ghana.

The fitted ARIMA (0, 1, 1) model for forecasting fatality cases from 1991-2011 was given by

$$\hat{y}_t = y_{t-1} + \varepsilon_t + 71.8104 + \varepsilon_{(t-1)}$$

TABLE IV: FORECASTING FATALITY CASES IN GHANA.

YEAR	FORECASTED FATALITY CASES
2012	2347.057
2013	2421.868
2014	2496.678
2015	2571.488
2016	2646.299

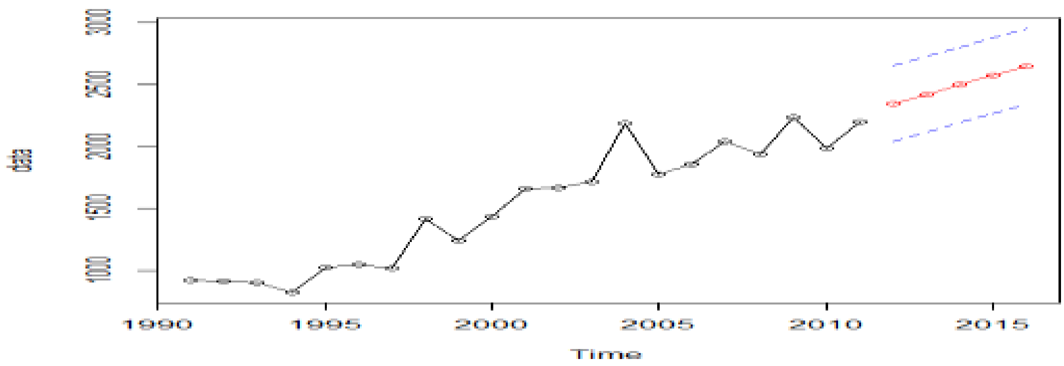

Fig 6 : Graph of the fatality cases, its forecasts and confidence intervals

The Figure gives the visual representation of the original fatality cases data (black line), its forecasts (red line) and confidence interval (blue short dashes lines).

From the prediction values and the graph above, it can be observed that, fatality cases in Ghana would continue to increase in the next 5 years.

4.0 Conclusions

Time series analysis of data from the years 1991 – 2011 showed that patterns of Road Traffic fatality cases were increasing in Ghana.

ARIMA (0, 1, 1) was identified to be a suitable model for forecasting into the future of the fatality cases in Ghana.

The study also revealed that road traffic fatality cases in Ghana would continue to increase over the next 5 years. The findings of this study draw attention to the importance of implementing key road safety measures in order to change the increasing pattern of road fatalities in the country. Future studies are therefore needed to understand the mechanisms underlying road fatalities in Ghana.

References

1. Afukaar, F.K., Antwi, P., and Ofosu-Amaah, S. (2003). Pattern of road traffic injuries in Ghana: implications for control. Injury control and safety promotion. vol. 10 (No. 1-2):69-76. [PubMed: 12772488].
2. Atubi, A.O. (2013). Time series and trend analysis of fatalities from road traffic accident in Lagos State, Nigeria. Mediterranean journal of social sciences, vol.4 (1) January 2013. Doi: 10.5901/mjss.2013.V4n1p251. ISSN:2039-9340.
3. Aworemi, J.R., Abdul Azeez, I.A., and Olabode, S.O. (2010). Analytical studies of the causal factors of road traffic crashes in South Western Nigeria. International research journals. Educational research Vol. 1(4) pp. 118-124. www.interesjournals.org/ER.

4. Ballesteros, M.F., Dischinger, P.C., and Langenberg, P. (2004). Pedestrian Injuries and Vehicle Type in Maryland, 1995–1999. *Accident Analysis and Prevention,* Vol. 36, No. 1 (2004): 73–81.
5. Bedard, M., Guyatt, G.H., Stones, M.J., and Hirdes, J. P. (2002). The independent contribution of driver, crash, and vehicle characteristics to driver fatalities. *Accident Analysis and Prevention Vol 34 No 6*, pp.717-727.
6. BRRI/NRSC (2010). Road traffic crashes in Ghana (statistics summary).
7. Chang, L.Y., and Mannering, F. (1999). Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. Accident Analysis and Prevention, **31**, 579.
8. Damsere-Dery, J., Beth, E., Ebel, C., Mock, N., Afukaar, F.K., and Donkor, P. (2010). pedestrians injury pattern in Ghana. Accident Anal prev 2010 July, 42(4):1080-1088.
9. European Commission, Directorate-General for energy and transport. (2001). Memo-Halving the number of road accident victims in the EU by 2010. A shared responsibility, European road safety action programme.
10. European Road Safety Observatory. (2008). Annual statistical report. www.transport-research.info/web/projects/project_details.cfm?id=35438.
11. Jacobs, G., Aeron-Thomas, A., and Astrop, A. (2000). Estimating global road fatalities. Transport research laboratory. ISSBN 0968-4107.
12. Khayesi, M. (1997). Livable streets for pedestrians in Nairobi: The challenge of road traffic accident. World transport policy and practice, 1997; 3(1):4-7.
13. MacLeod, K.E., Julia, B., Griswold, L.S., and Arnold, D.R. (2011). Factors Associated with Hit-and-Run Pedestrian Fatalities and Driver Identification. Accident Analysis and Prevention 45(2012) 366-372. www.elsevier.com/locate/aap.
14. Moudon, A.V. (2011). The Risk of Pedestrian Injury and Fatality in Collisions with Motor Vehicles, A Social Ecological Study of State Routes and City Streets in King County, Washington. Accident Analysis and Prevention, Vol. 43, No. 1 (2011): 11–24.
15. National Road Safety Commission (NRSC, 2007 Report). Performance Audit Report on Road Safety, Ghana).
16. National Road Safety Commission (NRSC, 2009). Ghana, newsletter 2009.
17. National Road Safety Commission (NRSC, 2013). "4th National Road Safety Awards". (www.nrsc.gov.gh/site/news/34).
18. Penden, M., McGee K, K.E. (2000). Injury-a leading cause of the global burden of disease 2000. Geneva, WHO, www.who.int/violenceinjuryprevention/index.html.
19. Penden, M., Scurfield, R., Sleet, D., and Dinesh, M. (2004). World Health Organization report on road traffic injury prevention, Geneva.
20. Plurad, D. (2006). Pedestrian Injuries: The Association of Alcohol Consumption with the Type and Severity of Injuries and Outcomes. *Journal of the American College of Surgery,* Vol. 202, No. 6 (2006): 919–927.
21. Rohayu, S., Sharifah Allyana, S.M.R., Jamilah, M.M., and Wong, S.V. (2012). Predicting Malaysian road fatalities for year 2020. Kuala Lumpur: Malaysian Institute of Road Safety Research 2020, MRR(6)2012.
22. Shinar, D. (1978). Psychology on the road: The human factor in traffic safety, New York, John Wiley and Sons. ISBN: 9780471039976.
23. World Health Organisation (2009). Global status report on road safety: time for action, Geneva; WHO.
24. World Health Organization (2013). Global status report on Road Strategy. Supporting a decade of action (Official-Report), Geneva-Switzerland. ISBN 9789241564564.
25. World Health Organization (2001). A 5-year Health Organization Strategy for Road Traffic Injury Prevention. Geneva Switzerland.
26. Yuan, Y.P., Xu-jun, Z., Zhi-bin, T., Meng-jing, C., and Yue,G. (2013). Autoregressive Integrated Moving Average Model in predicting road traffic injury in China. Zhonghua Liu Xing Bing XueZaZhi 2013 Jul; 34(7):736-9.