
Time Series Analysis of Road Traffic Accidents in Bauchi State

By

**Ali Adamu and **Ogundipe Adeboye Elijah*

**Department of Mathematics
Federal College of Education (Tech),
Gombe, Gombe State.*

***Mathematical Sciences Department
Abubakar Tafawa Balewa University,
Bauchi, Bauchi State*

ABSTRACT

Road traffic accidents rate in the recent days is becoming alarming in the country given the human and technological advancement. Lot of lives are being wasted pathetically, premature deaths and destruction of scarce human and materials resources. Nigeria has a huge natural and human resources who are day by day wasted and destroyed by this controllable disaster. However, this work is based on the statistical analysis of road traffic accidents and ways of reducing it to its bearest minimum. The secondary data used in this research was collected from the Federal Road Safety Commission (FRSC) RS 12.1. Trend shows that road accidents are increasing in Bauchi and that it is highest between October and December. Time series (ARIMA) is the method used with model (1,1,1) having the least Akaike Information Criteria (AIC) selected as the best model in analyzing the collected data which clearly shows that there is an increase in road accidents over the course of the years in each quarter. Furthermore, the selected model was used to forecast four (4) years which also shows an increase. Various recommendations were made based on the findings of this work which include proper funding and provision of up to date equipment.

INTRODUCTION

Road traffic accident (RTA) is a serious problem in Bauchi state and Nigeria at large, which resulted in many deaths. RTA accounts for over a million deaths in the world yearly (Sanusi, Adebola and Adegoke, 2015). The World Health Organization (WHO) classified road traffic accident as one of the top ten causes of death in the world, using International Statistical Classification of Diseases and Related Health Problems (ICD) (Organization, 2013). In this paper, RTA is defined as accident which took place on the road between two or more objects, one

of which must be any kind of a moving vehicle (Jha *et al.*, 2014).

In the early days of road network development in Nigeria, vehicular traffic was sparingly distributed across the available road network. Also, all components that make a functional road traffic environment such as good road signs, markings, painting, cycle paths, pedestrian walkways and Zebra crossing, were not only adequately provided, but were also maintained regularly especially in the urban centers by the relevant authorities (Agyemang, Aledun & Senevoh, 2013). However, priority given to the road network maintenance gradually dropped

immediately after the political independence of Nigeria in 1960. Although, more roads were constructed, but without the components that make traffic environment safe. For instance, traffic signs, road markings and paintings are conspicuously absent from all roads, and whenever available, they are overshadowed by bushy road environment and often time pull down by run-off vehicle respectively. They are also defaced due to irregular maintenance activities. Based on that non-challant attitude of governments to pay attention to road constructions and to provide strong penalties to those that violate traffic rules by motors/motor vehicles users made RTAs to be on increase not only in Bauchi state but in the whole Nigeria. The rate at which RTAs are increasing is seriously disturbing and therefore deserved urgent attention of government and other stake holders.

Time series techniques will be used to analyze data on recorded RTAs in Bauchi. Time series analysis is a sequence taken at successive equally spaced points in time. Therefore, the method is used to describe the inherent nature of data over time, and using it to predict the future (Hosseini, Yavarifar, & Yavarifar, 2016).

It is highly impossible that any family exist in Nigeria today which has not experienced the anguish of the sudden loss of their beloved ones in road accidents that could have been avoided or adequately controlled.

It is obvious that the rate of road accidents in the recent days is becoming alarming in the country given the human and technological advancement. Many lives are being wasted pathetically, premature deaths and destruction of scarce human and materials resources. RTA results in the

deaths of 1.2m people worldwide each year and injuries about 4 times this number (WHO, 2004). According to Ohakwe and et al (2011), RTA are increasing with rapid pace and presently these are one of the leading causes of deaths in developing countries. It is therefore necessary at this point to identify certain causes of road traffic accidents, proffer solutions and make future forecast using time series techniques.

Aim and Objectives of the Study

The primary aim of this study is to make time series analysis of RTA using past records and to identify pattern if any and predict the future occurrences and establish whether or not the residuals are linked over time. Secondly, this study is meant to deduce whether the types of accidents depend on the periods of occurrences with the following objectives:

1. To plot the time series plot and fit ARIMA model to the data
2. To identify among the ARIMA models the one that best fit the data.
3. To forecast the future occurrence of RTA in Bauchi, Bauchi State.

MATERIALS AND METHODS

Data used for the analyses was collected from Federal Road Safety Corps (FRSC) Bauchi for a period of 10 years (2005-2014), FRSC is the government agency created in February, 1988 with statutory responsibilities for road safety administration in Nigeria. FRSC also provide works and devices designed to eliminate or minimize accidents on the highways and advising governments on localities where such works are required. Time series plays significant role in the analysis of socio-economic data, mostly data on finance,

insurance, population, etc. It is also very useful in forecasting because one of the objectives of time series is to forecast future values.

Box-Jenkins (ARIMA) method will be used to derive models for forecasting these data. This method is preferred because of its high accuracy in forecasting data, especially within a short to medium term period. Its model simplicity gives it an advantage of cost and response time, because high cost is required to run and set up complex models (Nihan & Holmesland, 2015).

Model Building

According to Box, Jenkins & Reinsel (1993), there are three primary stages in building a Box-Jenkins (ARIMA) time series model. These are Model Identification, Model Estimation and Model Validation. All the three were considered in building the models. Box-Jenkins was considered because of the following.

- i) Box-Jenkins models are quite flexible due to the inclusion of both autoregressive and moving average terms.
- ii) Based on the World decomposition theorem, a stationary process can be approximated by an ARIMA model. In practice, finding that approximation may not be easy.
- iii) Building good ARIMA models generally requires more experience than commonly used statistical methods such as regression.

Differencing to achieve stationarity

Box and Jenkins recommend the differencing approach to achieve stationarity. However, fitting a curve and subtracting the fitted values from the

original data can also be used in the context of Box-Jenkins models.

Autoregressive Integrated Moving Average Model (ARIMA)

An ARIMA model is a class of statistical model for analyzing and forecasting time series data. ARIMA is an acronym that stands for Auto Regressive Integrated Moving Average. It is a generalization of the simpler Auto Regressive Moving Average and adds the notion of integration. This acronym is descriptive, capturing the key aspects of the model itself. Briefly, they are:

- a) **AR: Autoregression.** A model that uses the dependent relationship between an observation and some number of lagged observations.
- b) **I: Integrated.** The use of differencing of raw observations (i.e. subtracting an observation from an observation at the previous time step) in order to make the time series stationary.
- c) **MA: Moving Average.** A model that uses the dependency between an observation and residual errors from a moving average model applied to lagged observations.

Each of these components are explicitly specified in the model as a parameter. A standard notation is used of ARIMA (p, d, q) where the parameters are substituted with integer values to quickly indicate the specific ARIMA model being used.

The parameters of the ARIMA model are defined as follows:

- a) **p :** The number of lag observations included in the model, also called the lag order.

- b) **d**: The number of times that the raw observations are differenced, also called the degree of differencing.
- c) **q**: The size of the moving average window, also called the order of moving average.

are very useful tools because of the similarity in their behaviours are competitors to identify time domain model, it is also convenient to make use of the ACF and PACF graph.

Autocorrelation Function (ACF) and Partial Auto Correlation Function (PACF)

Data Presentation and Analysis

The ACF and PACF are used to identify an appropriate ARIMA model and the required number of parameters. These

We shall discuss the analysis of data using time series. These methods will now be compared to see which one best fit the data by studying various values and components

Table 1: Road Traffic Accidents in Bauchi between 2005 and 2014

Month	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
January	15	14	13	24	26	16	27	26	27	46
February	12	13	14	14	21	17	22	37	31	39
March	14	13	14	15	25	14	23	27	32	38
April	14	13	16	15	21	15	14	34	28	23
May	14	13	15	13	19	28	21	31	47	44
June	26	14	14	17	19	24	34	18	34	25
July	14	13	15	25	29	22	35	28	42	30
August	13	14	15	24	19	18	33	17	37	27
September	14	13	14	26	28	16	34	25	34	42
October	12	16	15	27	19	17	35	34	37	37
November	12	10	14	16	22	23	39	19	46	25
December	14	19	14	18	39	34	48	35	49	26
Total	174	165	173	234	287	244	365	331	444	402

Model Fitting for Road Accident Data

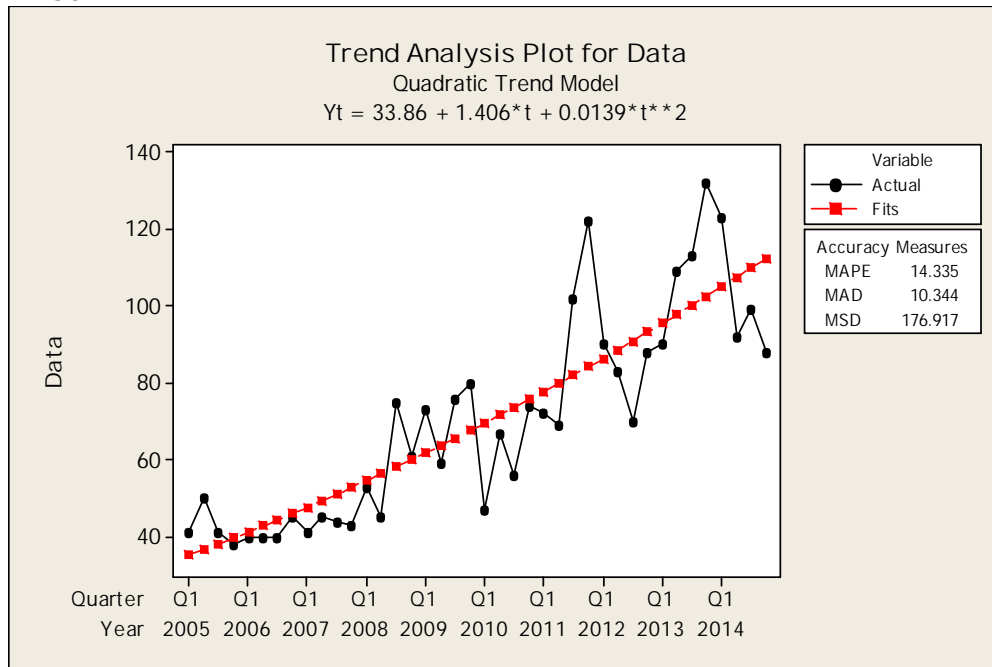


Fig.2: Quadratic Trend Model Plot of the Road Accidents

The graph above, quadratic trend model which has the least accuracy measure when compared with the linear model shows an increment in rate of road accidents in Bauchi State between 2005 and 2014.

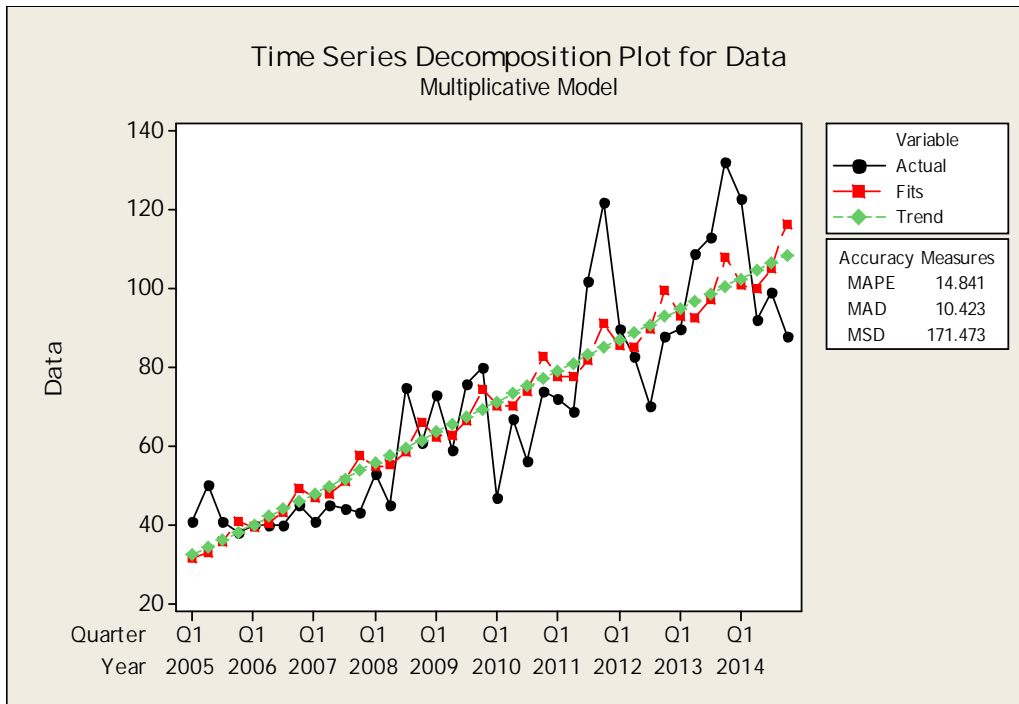


Fig.3: Multiplicative model plot of the road accident

Fitted Trend Equation: $Y_t = 30.29 + 1.95 \cdot t$

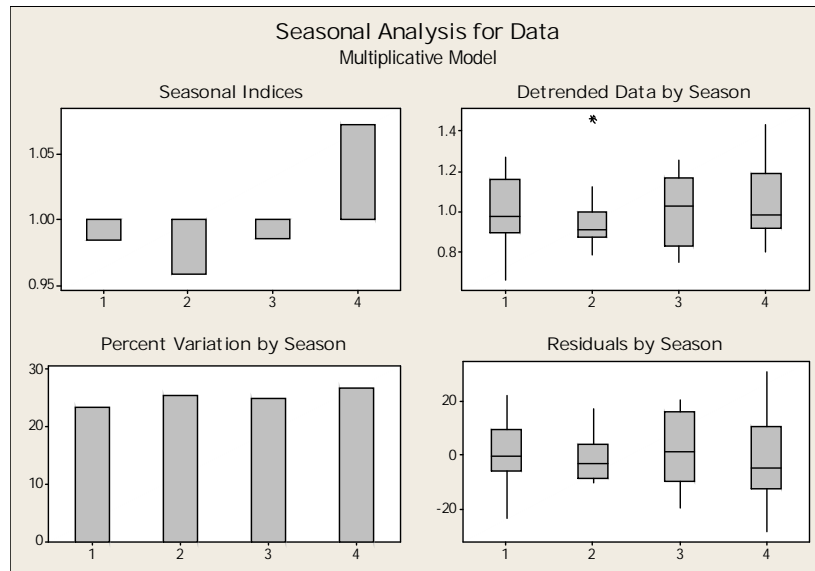


Fig.4: Seasonal plots for road traffic accidents

It can be seen from the seasonal indices graph that the accidents are high in lag 4 which signifies the 4th quarter of the year (October - December). Also, from the percent variation it is relatively high in 4th quarter of the year.

We therefore reject the null hypothesis H_0 and accept the alternative hypothesis H_1 and conclude that road accidents depend on the time of occurrence.

Autocorrelation Function (ACF) and Partial Autocarrelation Function (PACF)

The ACF graphs of road accident before differencing indicated the following

- It is no cyclical pattern
- That the observed highest lag is in lag 1

Whereas PACF shows

- It has a cyclical pattern
- The PACF were not in stationary

The ACF and PACF after first differencing.

The ACF shows:

- There is no correlation function after the first differencing
- Therefore, the series is stationary

The PACF shows:

- There is no correlation in the series after the first differencing
- The series indicates the presence of AR and MA process models

Time plot for quarterly road accident after first differencing is given on the fig. below

Unit Root Test for Road Accident

ADF Test for series: RTA

Sample range: [2005 Q4, 2014 Q4], T = 37

Lagged differences: 2

No intercept, no time trend

Asymptotic critical values

1% 5% 10%

-2.56 -1.94 -1.62

Value of test statistic: 0.1283

 $H_0: \alpha=0$ (the series contain unit root) $H_1: \alpha<0$ (the series is stationary)**Conclusion:** The test statistics value 0.1283 is greater than the asymptotic critical values therefore we reject the H_0 and conclude that the series is stationary

KPSS test for series: RTA

Sample range: [2005 Q1, 2014 Q4], T = 40

Number of lags: 2

KPSS test based on $y(t) = a + e(t)$ (level stationarity)

Asymptotic critical values:

10% 5% 1%

0.347 0.463 0.739

Value of test statistic: 1.2057

 H_0 : "Testing the null hypothesis of stationary" H_1 : "The series has presence of unit root"**Conclusion:** the test statistic is greater than the critical value, and then H_1 is accepted and concluded that the series has presence of unit root.**Model Specification for Road Accident**

The ARIMA identification and estimation was done using the time series modelling software J-multi.

Sample range: [2005 Q2, 2014 Q4], T = 39**Model:** ARIMA (0,1,0)

Final Results:

Iterations Until Convergence: 1

Log Likelihood: -161.529192 Number of Residuals: 39

AIC : 325.058385 Error Variance : 237.851551957

SBC : 326.721946 Standard Error : 15.422436641

Sample range: [2005 Q2, 2014 Q4], T = 39**Model:** ARIMA (0,1,1)

Final Results:

Iterations Until Convergence: 10

Log Likelihood: -159.603853 Number of Residuals: 39
 AIC : 323.207706 Error Variance : 220.438382573
 SBC : 326.534829 Standard Error : 14.847167493

Sample range: [2005 Q2, 2014 Q4], T = 39

Model: ARIMA (1,1,1)

Final Results:

Iterations Until Convergence: 6

Log Likelihood: -156.910261 Number of Residuals: 39
 AIC : 319.820522 Error Variance : 186.444419064
 SBC : 324.811207 Standard Error : 13.654465169

Sample range: [2005 Q2, 2014 Q4], T = 39

Model: ARIMA (2,1,0)

Final Results:

Iterations Until Convergence: 2

Log Likelihood: -159.818709 Number of Residuals: 39
 AIC : 325.637418 Error Variance : 229.463394227
 SBC : 330.628103 Standard Error : 15.148049189

Conclusion: After all possible combinations, the best model is the one with the smallest Akaike Information Criteria (1,1,1) with AIC=319.820522

Diagnostic Tests for Road Accidents

Residual Test

PORTMANTEAU TEST with 12 lags

Portmanteau: 14.4319
 p-Value (Chi²): 0.2740
 Ljung& Box: 18.9372
 p-Value (Chi²): 0.0901

H₀: none of the Autocorrelation coefficients up to lags are different from zero

H₁: at least one of the Autocorrelation coefficients up to lags is different from zero

Conclusion: the Ljung and box statistics is 18.9372 so at least one of the Autocorrelation coefficients is different from zero

JARQUE-BERA TEST:

Test statistic: 0.4415
 P-value (Chi²): 0.8019
 Skewness: -0.2415
 Kurtosis: 3.2305

H₀: the distribution of the series is symmetric and hence normal

H₁: the distribution of the series is asymmetric and hence non-normal

Conclusion: the Jarque–Bera test statistics 0.4415 and therefore the series is symmetric and hence normal, accepting H₀.

Forecasting for Future Road Accidents

An important application in the time series analysis is to forecast future values of the given time series data. A forecast is

characterized by its origin and lead time to access sample forecast ability of the model. A 4-year forecasted road accidents data and the graph are given below:

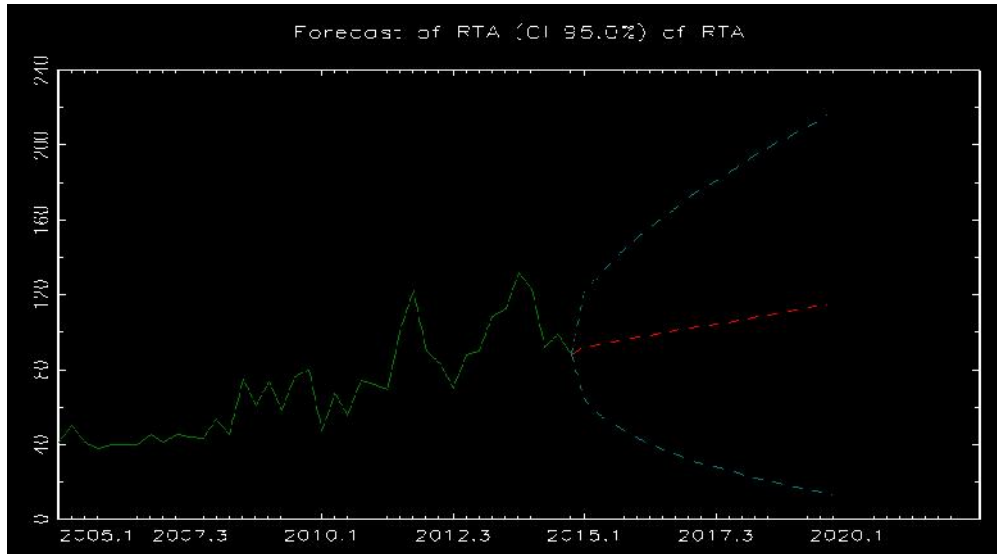


Fig.8: Graph of forecast for road accidents

Forecasted variable: RTA (in levels)

Forecast range: [2015 Q1, 2019 Q4], T = 20

Time	Lower CI	Forecast	Upper CI	Std. err
2015 Q1	63.0238	92.7135	122.4031	15.1480
2015 Q2	56.8701	93.1973	129.5246	18.5346
2015 Q3	51.6603	94.5752	137.4901	21.8958
2015 Q4	47.3383	95.7846	144.2309	24.7180
2016 Q1	43.5881	97.0235	150.4589	27.2635
2016 Q2	40.2649	98.2575	156.2501	29.5886
2016 Q3	37.2748	99.4923	161.7099	31.7442
2016 Q4	34.5538	100.7270	166.9002	33.7625
2017 Q1	32.0563	101.9617	171.8671	35.6667
2017 Q2	29.7482	103.1964	176.6446	37.4743
2017 Q3	27.6033	104.4311	181.2588	39.1986
2017 Q4	25.6009	105.6658	185.7306	40.8502
2018 Q1	23.7244	106.9004	190.0765	42.4375
2018 Q2	21.9602	108.1351	194.3101	43.9676
2018 Q3	20.2969	109.3698	198.4428	45.4462
2018 Q4	18.7249	110.6045	202.4841	46.8782
2019 Q1	17.2362	111.8392	206.4422	48.2677
2019 Q2	15.8237	113.0739	210.3241	49.6184
2019 Q3	14.4814	114.3086	214.1358	50.9332
2019 Q4	13.2039	115.5433	217.8826	52.2149

From the forecasted data above, road traffic accidents in Bauchi metropolis is likely to increase if necessary measures are not put in place.

CONCLUSION

From the analysis, trend is used to find out whether road accident is increasing or decreasing and it shows that there is an increase in road accidents over the course of the years in each quarter. Furthermore, there is no relationship between successive observed accidents from autocorrelation test.

From the forecast, it can be seen that road accident will be increasing by each quarter, which is very high in the last quarter of the years if the necessary combating efforts are not put in place. These increments, traceable to human factors are due to: over speeding, dangerous driving and over speeding, tyre burst, brake failure, fatigue, stationary vehicle /obstruction, night journey , drinking and driving, poor weather, mechanical deficiency etc.

RECOMMENDATIONS

Base on the findings of the study, the following recommendations are made:

- a) There should be effective electronic data management system to access road traffic accidents
- b) Effective utilization of speed limits across Nigerian roads especially during the festive season being the 4th quarter of the year and consequently throughout the year
- c) Provision of adequate fundamental functional road traffic environment including proper road / lane markings, street lights, safety barriers, and speed bumps etc.

- d) Adequate training for prospective road users from reputable government agencies and issuance of road worthiness certificate.
- e) Constant sensitization of critical stakeholders, transport associations, insurance companies, government agencies, haulage and other professional associations to fully cooperate with all tenets of the national road traffic regulations.
- f) Government should provide adequate funding to Federal Road Safety Corps and provision of up-to-date equipment.
- g) Finally, another research should be carried out on road accidents in other part of the country.

REFERENCES

- Agyemang, B., Abledu, G. K., & Semevoh, R. (2013). Regression Analysis of Road Traffic Accidents and Population Growth in Ghana. *International Journal of Business and Social research*, 3(10),41-47.
- Box, G. E. P., Jenkins G. M., Reinsel, G. C., (1993)- *Time Series Analysis, Forecasting and Control*. Third Edition, Prentice Hall, Englewood Cliffs, NJ.
- Hosseini, Y., Yavarifar, O. (2016). Rumble strips and pedestrian safety. *ITE journal* 63(8) 18-24.
- Jha, N., Srinivasa, D. K., Roy, G.& Jagdish, S. (2004). Epidemiological study of road traffic accident cases: *Indian Journal Community Med.*,29:20-24
- Nihan, P., Holmesland, G. (2015). Speed limitation and motorway casualties: a time series count data regression

- approach. *Accident Analysis and Prevention*. 28 (1) 73-87.
- Organization, W. H. (2014). Global Status report on Road Accident Safety. Injury Prevention, 318.
http://.who.int/violence_injury_prevention/road_safety_status/2015/en/index.html
- Ohakwe J. and et'al (2011). Analysis of Road Traffic Accident in Nigeria: A case study Obinze/Nekede/Iheagwa in Imo State, *Asian Journal of Applied Sciences* 4(2),166-175.
- Sanusi, D. Adebola, J. & Adegoke, S. (2015). Rumble strips and pedestrian safety. *ITE journal* 63(8) 18-24.
- World Health Organization, (2004). World report on Road Traffic Injury Prevention: Summary, World Health Organization, Geneva, Switzerland.