ISSN 2320-5407 International Journal of Advanced Research (2014), Volume 2, Issue 5, 1045-1059

ISSN NO. 2320-5407

Journal homepage: http://www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH

RESEARCH ARTICLE

USE OF TIME SERIES ANALYSIS OF ROAD ACCIDENT DATA IN LAGOS STATE

[1]Balogun .O.S.*, [2]O.O.Awoeyo and [3]Dawodu .O.O.

1.Department of Statistics and Operations Research, Modibbo Adama University of Technology, P.M.B. 2076, Yola, Adamawa State, Nigeria.

2.Department of Statistics,University of Ilorin, P.M.B.1515, Ilorin, Kwara State. Nigeria.

3. Department of Physical and Computer Sciences(Statistics Unit), McPherson University, Seriki Sotayo, Ogun State.

Manuscript Info

Manuscript History:

Received: 23 March 2014
Final Accepted: 23 April 2014
Published Online: May 2014

Key words:
Autocorrelation, Autovariance, AIC, SIC and Dubin Watson.

Corresponding Author

Balogun .O.S.

Abstract

We analyze a data set collected from FRSC Ojodu command on road traffic accident using time series approach. The data collected spanned the period between 1989 to 2008. Our analysis revealed a positive cubic trend of number of accidents on time. The best model was identified through the criteria for AIC & SIC. These two model selection methods agreed on AR (1) as being the best model for the data. Various results obtained from this study indicated an increasing trend (though not linear) spate in road accidents in Nigeria.

1. Introduction

Having being disturbed by the unpleasant trend in the nation's road traffic system, the Federal Government of Nigeria initiated and established the Federal Road Safety Commission (*FRSC*) to check the alarming increase in the number of road traffic in Nigeria. The FRSC was established by Decree 45 of 1988, as the lead agency in Nigeria on road safety administration and management. The vision of the commission is to eradicate road traffic crashes and create safe motoring environment in Nigeria. Missions includes regulate, enforce and coordinate all road traffic and safety management activities through:

- Sustained public enlightenment.
- Effective patrol operations.
- Prompt rescue services.
- Improved vehicle administration.
- Robust data management.
- Promotion of stakeholder cooperation.

Within the provision of its enabling Act, the functions of the FRSC are as follows:

(a) Preventing and minimizing road traffic accidents.
(b) Clearing obstructions on the highways.
(c) Educating drivers, motorists and other members of the public on the proper use of the highways.

(d) Providing prompt attention and care to victims of road traffic accidents.
(e) Conducting researches into causes of road traffic accidents and researches.
(f) Determining and enforcing speed limits for all categories of roads and vehicles.
(g) Co-operating with bodies, agencies and group engaged in road safety activities or the prevention of highway accident.

In particular the commission is charged with the responsibilities for:

1. Preventing or minimizing accidents on the highway;
2. Clearing obstructions on any part of the highways;
3. Educating drivers, motorists and other members of the public generally on the proper use of the highways;
4. Designing and producing the driver's license to be used by various categories of vehicle operators;
5. Determining, from time to time, the requirements to be satisfied by an applicant for a driver's license;
6. Designing and producing vehicle number plates
7. The standardization of highway traffic codes;
8. Giving prompt attention and care to victims of accidents
9. Conducting researches into causes of motor accidents and methods of preventing them and putting into use the result of such researches;
10. Determining and enforcing speed limits for all categories of roads and vehicles and controlling the use of speed limiting devices;
11. Cooperating with bodies or agencies or groups in road safety activities or in prevention of accidents on the highways;
12. Making regulations in pursuance of any of the functions assigned to the Corps by or under this Act.
13. Regulating the use of sirens, flashers and beacon lights on vehicles other than ambulances and vehicles belonging to the Armed Forces, Nigeria Police, Fire Service and other Para-military agencies;
14. Providing roadside and mobile clinics for the treatment of accident victims free of charge;
15. Regulating the use of mobile phones by motorists;
16. Regulating the use of seat belts and other safety devices;
17. Regulating the use of motorcycles on the highway;
18. Maintaining the validity period for drivers' licenses which shall be three years subject to renewal at the expiration of the validity period; and

In exercise of the functions, members of the Commission shall have power to arrest and prosecute persons reasonably suspected of having committed any traffic offence.

The aims of this research is to study the past behavior of road accident in Lagos State, to predict the future behavior of road accident in Lagos State, to offer useful suggestions to FRSC, government, motorist and the populace and to identify the presence of likely component and estimate the identified component in the data.

The scope of the research is to use Time series analysis on the annually number of road accident for 20 years in Lagos state. Accident has led to loss of many lives in this country, many people sustained permanent injuries or some forms of disabilities. This sad event occurs mostly to Nigerian Youths in their productive age which is diminishing the work force of the country.

Time series was originated in 1807 by French Mathematician name FOURIER, who claimed that any Series could be approximated as the sum of the Sine and cosine terms. In 1960 Schituster used Fourier's idea to estimate the length periodicities and utilized peridogram analysis in his research.

According to Murray R. Spiegel (1981) he defines time series as a set of observations taken at a specified time usually at equal interval.

According to Frankowne and Ronjones (1975) define time series as a statistical series which tell us how data has been behaving in the past.

According to Chatfield (1987), he defines time series as a collection of observation segmental in time at regular intervals.

The usage of time series models is in twofold:

(1) To obtain an understanding of the underlying forces and structure that produced the observed data, and

(2) To fit a model and proceed to forecasting, monitoring or even feedback and feed forward control.

Time Series Analysis's includes: Economic Forecasting, Sales Forecasting, Budgetary Analysis, Stock Market Analysis, Yield Projections, Process and Quality Control, Inventory Studies, Workload Projections, Utility Studies, Census Analysis, and many, many more...

The component of time series otherwise refers to as characteristics of time series can be classified into 4 main types:

(a) Secular trend (Tt)
(b) Seasonal variation (St)
(c) Cyclical variation (Ct)
(d) Irregular variation (It)

The objectives of Time series include: description, explanation, prediction and control.The time series model refers to how the four components of time series are related. If the individual influence of each component of the series is known, we can forecast base on the effect of each of them. The relationship existing between the components are

(1) Additive model
(2) Multiplicative model

(1) Additive Model: In the additive model, the four components are related as

$Y_t = T_t + S_t + C_t + I_t$,

whereY_t is the sum of the variables,

T stand for Secular, S is the Seasonal, C stand for Cyclical and I stand for Irregular.

(2) Multiplicative Model: The four component are related as

$Y_t = T_t * S_t * C_t * I_t$,

whereY_t is the product of the variables,

T stand for Secular, S is the Seasonal, C stand for Cyclical and I stand for Irregular.

1.1 Time Series Analysis

Time series analysis consists of description of the component movements present in the data. The graphical representation i.e. the time plot reveals quantitatively the presence of long term trend, cyclical, seasonal and irregular variations. Time series analysis amounts to investigating the factors T, C, S & I and is refers to as a decomposition of time series into its basic component movements.

The indication from the time plot enable us to construct the long trend term curve and obtain appropriate trend value by the use of:

(1)The method of freehand,
(2)Least squares method.
(3)Moving average method
(4)Semi-average method.

In this research work, the moving average and the least square will be employed.

1.2 Durbin-Watson Test

The Durbin–Watson statistic is a test statistic used to detect the presence of autocorrelation in the residuals from a regression analysis. It is named after James Durbin and Geoffrey Watson.

However, the small sample distribution of this ratio was derived in a path-breaking article by John von Neumann (von Neumann, 1941). Durbin and Watson (1950, 1951) applied this statistic to the residuals from least squares regressions, and developed bounds tests for the null hypothesis that the errors are serially independent against the alternative that they follow a first order autoregressive process.

Later, John Denis Sargan and Alok Bhargava developed several von Neumann-Durbin-Watson type test statistics for the null hypothesis that the errors on a regression model follow a process with a unit root against the alternative hypothesis that the errors follow a stationary first order autoregression(Sargan and Bhargava, 1983).

If e_t is the residual associated with the observation at time t, then the test statistic is

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}.$$

Since *d* is approximately equal to 2(1-*r*), where *r* is the sample autocorrelation of the residuals, *d* = 2 indicates no autocorrelation. The value of *d* always lies between 0 and 4.

If the Durbin–Watson statistic is substantially less than 2, there is evidence of positive serial correlation. As a rough rule of thumb, if Durbin–Watson is less than 1.0, there may be cause for alarm. Small values of *d* indicate successive error terms are, on average, close in value to one another, or positively correlated. If $d > 2$ successive error terms are on average, much different in value to one another, i.e., negatively correlated. In regressions, this can imply an underestimation of the level of statistical significance.

To test for **positive autocorrelation** at significance α, the test statistic *d* is compared to lower and upper critical values ($d_{L,\alpha}$ and $d_{U,\alpha}$):

- If $d<d_{L,\alpha}$, there is statistical evidence that the error terms are positively autocorrelated.
- If $d>d_{U,\alpha}$, there is statistical evidence that the error terms are **not** positively auto correlated.
- If $d_{L,\alpha}<d<d_{U,\alpha}$, the test is inconclusive.

To test for **negative autocorrelation** at significance α, the test statistic $(4 - d)$ is compared to lower and upper critical values ($d_{L,\alpha}$ and $d_{U,\alpha}$):

- If $(4 - d) <d_{L,\alpha}$, there is statistical evidence that the error terms are negatively auto correlated.
- If $(4 - d) >d_{U,\alpha}$, there is statistical evidence that the error terms are **not** negatively auto correlated.
- If $d_{L,\alpha}< (4 - d) <d_{U,\alpha}$, the test is inconclusive.

The critical values, $d_{L,\alpha}$ and $d_{U,\alpha}$, vary by level of significance (α), the number of observations, and the number of predictors in the regression equation.

1.3 Time Plot

The time plot is the graphical representation of data. We plot the observed variable against time. The time plot reveals the presence of the likely component in the data.

1.4 Correlogram

In the analysis of time series, a **correlogram**also known as an autocorrelation plot is a plot of the sample autocorrelations. If cross-correlation is used, it is called a cross**-correlogram.**

The **correlogram** is a commonly-used tool for checking randomness in a data set. This randomness is ascertained by computing autocorrelations for data values at varying time lags. If random, such autocorrelations should be near zero for any and all time-lag separations. If non-random, then one or more of the autocorrelations will be significantly non-zero.

In addition, correlograms are used in the model identification stage for Box-Jenkins autoregressive moving average time series models. Autocorrelations should be near-zero for randomness; if the analyst does not check for randomness.

2. Research Methodology

2.1 Methods of Estimating Trend

The model used to estimate the trend and for prediction for 4 years is

$\hat{Y}_t = \alpha + \beta t_i^3 + e_t$ (1)

Where Y_t is the observed data

t is the time of the variable

α and β are the intercept and the slope

e_t is the random error term assumed to be

NIID$(0,\delta^2)$

$\alpha = \sum^{Y_i}/_N - \beta\, t_i^3$ (2)

$$and\ \beta = \frac{n\sum y_i t^3{}_i - \sum y_i \sum t^3{}_i}{n\sum y_i t^3{}_i - \sum y_i \left(\sum t_i\right)^3} \quad (3)$$

2.2 Durbin-Watson Test.

The Hypothesis is

H_0= First serial correlation equals to zero

H_1= First serial correlation not equals to zero

The test statistic is

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}.$$

2.3 Autocovariance and Autocorrelation Function

The autocovariance of lag k is denoted as

$Y_k = e_k = \frac{1}{N-K}\sum(X_K - \bar{X}_K)\,(X_{K+1} - \bar{X}_K)$

Where k = 0, 1, 2,……

The autocorrelation lag k is obtain by dividing the autocovariance function of lag k by that of lag 0

,i.e., $e_k = \frac{ek}{eo}$

2.4 Fitting of Autoregressive Model

A stationary process Y_t is said to follow an Autoregressive process of order p if

$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \ldots\ldots\ldots\ldots.. + \Phi_p Y_{t-p} + e_t$

Where e_t = white noise process

$\Phi_1, \Phi_2, \text{-----------} \Phi_p$ = finite set weight parameter.

2.4 Model Selection Method

(1) Akaike Information Criteria (AIC)

This is used for chosen the best model using Autoregressive model

It is written as $AIC\ (p) = \ln\delta^2_p + \frac{2P}{T}$

Where T = Number of observation,

P = Number of order

$\ln\delta^2_p$ = the maximum likelihood estimate of residual variance after fitting Autoregressive of order p

Note: the minimum AIC (p) is selected as the best order.

When a model involving *q* parameters is fitted to data, the criterion is defined as $-2L_q + 2q$, where L_q is the maximized log likelihood. Akaike suggested maximizing the numbers of parameters. It was originally proposed for time-series models, but has also been used in regression. Marriott (1990), A Dictionary of Statistical Terms

(2)SCHWARTZ INFORMATION CRITERIA (SIC)

This is another model selection criterion, which is written as

$SIC = \ln\delta^2_p + \frac{Pln\ T}{T}$

Where T = Number of observation

P = number of order

$\ln\delta^2_p$ = the maximum likelihood estimate of residual variance after fitting Autoregressive of order p

note: the minimum SIC (p) is selected as the best order.

2.6 Levison Durbin Algorithm

Levison Durbin Algorithm is used to estimate the parameters needed in the estimation of Akaike Information Criteria and Schwartz Information Criteria.

In Engineering, The Levinson-Durbin algorithm uses the autocorrelation method to estimate the linear prediction parameters for a segment of a random signal. Suppose we have an autoregressive process for successive order of P_k, we fit AR (p) as follows

$\Phi_{k+1} = a_{k+1,k+1} = 1/\, \delta_k^2\ (¥_{k+1} - \sum a_{k+1,}¥_{k+1-i})$

For j = 1,2,3,..,k

Note, $\delta_0 = ¥_0 = Var(Y_t)$

$\Phi_{k+1,i} = a_{k+1,k+1} = 1/\, \delta_k^2\ \{¥_{k+1} - \sum a_{k,i,}\ ¥_{k+1-i}\}$

$\delta_{k+1}^2 = \delta_k\ (1-\Phi_{k+1}^2)$

for k =0

$\Phi_1 = a_{11} = ¥k/\, \delta_0^2 = e_1$

$\delta_1^2 = \delta_0^2\ (1 - \Phi_1^2)$

for k =1

$\Phi_2 = a_{22} = 1/\, \delta_1^2\ \{¥_2 - ¥_1\, a_{11}\}$

$a_{21} = a_{11} - \Phi_2 a_{11}$

$\delta_2^2 = \delta_1^2\ (1 - \Phi_2^2)$

for k =2

$\Phi_3 = a_{33} = 1/\, \delta_2^2\ \{¥_3 - a_{21}¥_2 - ¥_1\, a_{22}\}$

$= ¥_0/\delta_2^2(e_3 - a_{21}e_2 - a_{22}e_1)$

$a_{31} = a_{21} - \Phi_3 a_{22}$

$a_{32} = a_{22} - \Phi_3 a_{21}$

$\delta_3^2 = \delta_2^2(1 - \Phi_3^2)$

2.7 Estimation of Forecast Error Accuracy

The standard error is necessary so as to know which method will give a more reliable prediction, by choosing the estimate with lowest standard error.

Standard error is computed using

$$SE = \sqrt{\frac{\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)}{N-K}}$$

Where

Y_t = Original data

$\hat{Y}_t$ = predicted data

K = Number of predictor

N = Number of parameters

3. Data Presentation

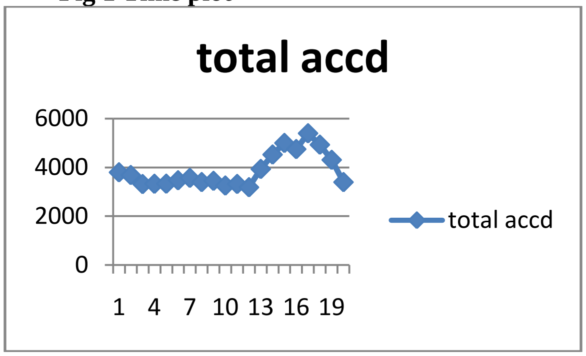
Table 1: No of Accident

YEAR	$\mathbf{Y_t}$
1989	**3798**
1990	**3686**
1991	**3315**
1992	**3327**
1993	**3331**
1994	**3470**
1995	**3569**
1996	**3397**
1997	**3447**
1998	**3255**
1999	**3319**
2000	**3186**
2001	**3934**
2002	**4527**
2003	**5001**
2004	**4750**

2005	**5398**
2006	**4931**
2007	**4309**
2008	**3396**

4. Analysis

Fig 1 Time plot



4.1 Estimation of Trend by Least Square Method

To complete the equation of trend T_t with the data using the Least Squared Method (LSM), the required equation model is

$\hat{Y} = \alpha + \beta\, t_i^3$

Where $\alpha = \sum {}^{Y_i}/_N - \beta\, t_i^3 = 3515.714$

$$\beta = \frac{n\sum y_i t^3_i - \sum y_i \sum t^3_i}{n\sum y_i t^3_i - \sum y_i \left(\sum t_i\right)^3} = 0.159$$

$$\textit{The trend equation is } \hat{Y} = 3515.714 + 0.159 t^3_i$$

4.2 Estimation of Standard Error

S.E = $\sqrt{\dfrac{\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)}{N-K}}$

Where

Y_t = Original data

$\hat{Y}$ = predicted data

K = Number of predictor

N = number of parameters

S.E = $\sqrt{\frac{(3396\ -\ 4988)2\ +\ ......+(5398\ -\ 5714)2}{20-1}}$

= $\sqrt{\frac{3713681}{19}}$

S.E = 442.105

4.3 Durbin-Watson Analysis

H_0= First Serial Correlation equals to Zero

H_1= First Serial Correlation not equals to Zero

Test Statistic is

$$d = \frac{\sum_{t=2}^{T}(e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}.$$

d=0.618

Decision rule

Reject H_0 if d=0.952 otherwise accept

Decision and Conclusion

Since d≠0.952 accepts H_0 and Conclude that there is Positive Autocorrelation i.e. the Variance of the error are dependent.

4.4 Estimation of Autocorrelation

$Y_k = e_k = \frac{1}{N-K}\sum(X_K - \bar{X}_K)(X_{K+1} - \bar{X}_K)$

Where k = 0,1,2,……

for k = 0

$e_0 = \frac{1}{20}\sum(X_t - \bar{X}_t)^2$

For k=1

$e_1 = \frac{1}{N-1}\sum(X_t - \bar{X}_t)(X_{t+1} - \bar{X}_t)$

$e_k = \dfrac{e_k}{e_o}$

when k=0 , $e_0 = \dfrac{e_0}{e_0}$

when k=1 , $e_1 = \dfrac{e_1}{e_0}$

Autocorrelation at lag16 is given in the table

Table2 Autocorrelation at lag 16

Lag	Autocorrelation
1	0.805
2	0.558
3	0.267
4	0.078
5	-0.152
6	-0.215
7	-0.206
8	-0.160
9	-0.194
10	-0.225
11	-0.258
12	-0.262
13	-0.239
14	-0.188
15	-0.100
16	-0.032

Correlogram

The Correlogram of lag (0 -16)

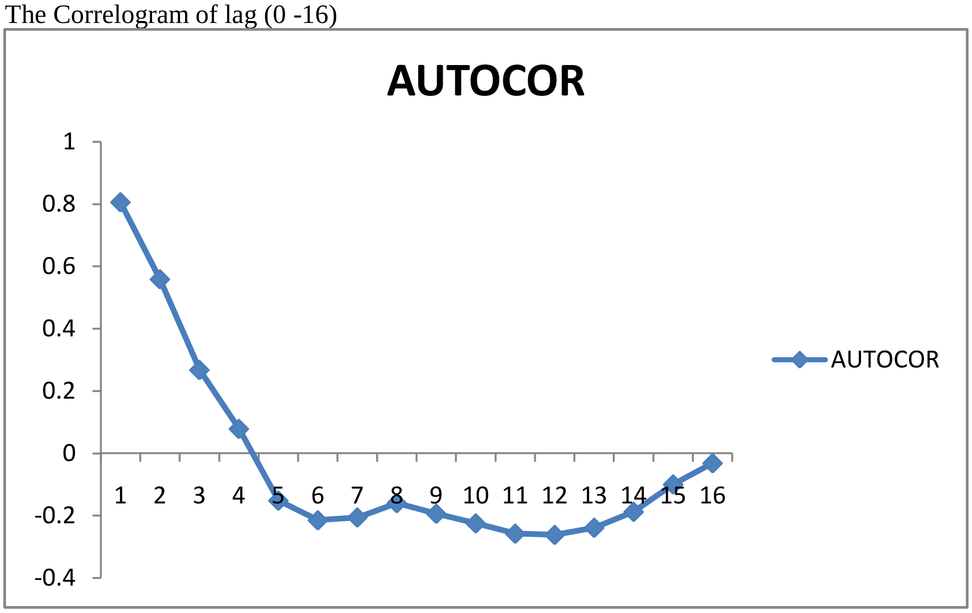


Fig 2

4.5 Estimation of Levison Durbin Algorithm

Using Levison Durbin Algorithm for

$$for\ k = 0,1,2,3,4,5$$

$$\Phi_{k+1} = a_{k+1,k+1} = \frac{1}{\delta_k^{\ 2}}(¥_{k+1} - \sum a_{k+1}, ¥_{k+1,i})$$

$$for\ \ j = 1,2,3,...,k$$

$$\Phi_{k+1,j} = a_{k+1,j} = a_{k+1,j} - \Phi_{k+1}a_{k+1,j}$$

$$\delta_{k+1}^{\ \ 2} = \delta_k(1 - \Phi_{k+1}^{\ \ 2})$$

$$Note, \delta_0 = ¥_0 = Var(Y_t) = 479075.69$$

$$for\ k = 0$$

$$\Phi_1 = a_{11} = \frac{¥_0}{\delta_0^{\ 2}} = 479075.69$$

$$\delta_1^{\ 2} = \delta_0^{\ 2}(1 - \Phi_1^{\ 2})$$

$$= 479075.69 - (1 - 0.805^2)$$

$$= 168622.67$$

$$for\ k = 1$$

$$\Phi_2 = a_{22} = \frac{e_2 - e_1^{\ 2}}{1 - e_1^{\ 2}}$$

$$= \frac{0.558 - 0.805^2}{1 - 0.805^2}$$

$$= -0.256$$

$$a_{21} = a_{11} - \Phi_2 a_{11}$$

$$= 0.805 - (-0.256 \times 0.805)$$

$$= 1.011$$

$$\delta_2^{\ 2} = \delta_1^{\ 2}(1 - \Phi_2^{\ 2})$$

$$= 168622.67 - (1 - (-0.256^2))$$

$$= 157571.81$$

$$for\ k = 2$$

$$\Phi_3 = a_{33} = \frac{¥_0}{\delta_2^{\ 2}}(e_3 - a_{21}e_2 - a_{22}e_1)$$

$$= \frac{479075.69}{157571.81}(0.267 - (1.011 \times 0.558) - (-0.256 \times 0.805))$$

$$= -0.277$$

$$a_{31} = a_{21} - \Phi_3 a_{32}$$
$$= 1.011 - (-0.277 \times -0.256)$$
$$= 0.940$$
$$a_{32} = a_{22} - \Phi_3 a_{31}$$
$$= -0.256 - (-0.277 \times 1.011)$$
$$= 0.024$$
$$\delta_3^{\ 2} = \delta_2^{\ 2}(1 - \Phi_3^{\ 2})$$
$$= 157571.81 - (1 - (-0.277^2))$$
$$= 14581.48$$

for $k = 3$

$$\Phi_4 = a_{44} = \frac{¥_0}{\delta_3^{\ 2}}(e_4 - a_{31}e_3 - a_{32}e_2 - a_{33}e_1)$$
$$= \frac{479075.69}{145481.48}(0.078 - (0.940 \times 0.267) - (0.024 \times 0.558) - (-0.277 \times 0.805))$$
$$= 0.121$$
$$a_{41} = a_{31} - \Phi_4 a_{33}$$
$$= 0.940 - (0.121 \times -0.277)$$
$$= 0.974$$
$$a_{42} = a_{32} - \Phi_4 a_{32}$$
$$= 0.024 - (0.121 \times 0.024)$$
$$= 0.021$$
$$a_{43} = a_{33} - \Phi_4 a_{31}$$
$$= -0.277 - (0.121 \times 0.940)$$
$$= -0.391$$
$$\delta_4^{\ 2} = \delta_3^{\ 2}(1 - \Phi_4^{\ 2})$$
$$= 145181.81(1 - 0.121^2)$$
$$= 143351.49$$

for $k = 4$

$$\Phi_5 = \frac{¥_0}{\delta_4^{\ 2}}(e_5 - a_{41}e_4 - a_{42}e_3 - a_{43}e_2 - a_{44}e_1)$$

$$= \frac{479075.69}{143351.49}(-0.152-(0.974\times 0.078)-(0.021\times 0.267)-(-0.391\times 0.558)-(0.121\times 0.805))$$
$$= -0.377$$

$$a_{51} = a_{41} - \Phi_5 a_{44}$$
$$= 0.974 - (-0.377\times 0.121)$$
$$= 1.020$$

$$a_{52} = a_{42} - \Phi_5 a_{43}$$
$$= 0.021 - (-0.377\times 0.391)$$
$$= -0.126$$

$$a_{53} = a_{43} - \Phi_5 a_{42}$$
$$= -0.391 - (-0.377\times 0.021)$$
$$= -0.383$$

$$a_{54} = a_{44} - \Phi_5 a_{41}$$
$$= 0.121 - (-0.377\times 0.974)$$
$$= 0.488$$

$$\delta_5^{\,2} = \delta_4^{\,2}(1-\Phi_5^{\,2})$$
$$= 143351.49(1-(-0.377^2))$$
$$= 122977.09$$

Table 3 Estimations of AIC and SIC

AR(p)	AIC	SIC
1	**12.135**	**12.185**
2	**12.168**	**12.267**
3	**12.188**	**12.337**
4	**12.273**	**12.472**
5	**12.220**	**12.469**

Table 4 Summary for the Forecast Values

YEAR	T	X_t		AR(1)	LSM
2009	**21**	**3396**		**4131**	**4988**
2010	**22**	**4309**		**4810**	**5209**
2011	**23**	**4931**		**5307**	**5450**
2012	**24**	**5398**		**4876**	**5714**

5. Discussion of Result and Conclusion

From the time plot of the raw data, it could be seen that the highest total number of accident occurred in 2005 and the lowest total number of accident occurred in the year 2000 (see fig1). The Time plot in fig 1 generally indicated a non-stationary series and the stationary series were obtained by taking the first difference of the original accident data. The least square trend of the total number of accident on the Cube of the time was fitted to the data, and the results indicated significant dependence of the number of accident on time. The trend equation is

$\hat{Y} = 3515.714 + 0.159t_i^3$

Thus, the average number of Road accidents in Lagos state per year is about 3,516 which increase at 0.159 rate of cubic time. According to the AIC and SIC model selection procedure adopted, an autoregressive of order 1 (AR (1)) was selected as the best model for the data. The AR (1) model is given as

$Y_t = 0.805Y_{t-1} + e_t$

The fitted trend was used to forecast the number of roads accident for the next four years being from 2009 to 2012 that were not covered by the data. Hence, 4131, 4810, 5307 and 4876 accidents were prediction for the coming years which are 2009, 2010, 2011 and 2012.

6. Recommendations

From the analysis conducted in this research work and outcome of our findings, we decide to offer these responsive recommendations for the stakeholders in Lagos state:

- The Federal Road Safety Corp should upgrade their effort in term of sensitization of the road users on the rules guiding driving and provide Severe punishment for road law offenders
- Due to apparent increasing trend in the outcome of road accident on our road, the government should look into the poor state of the country's road being a major cause of road accident.
- More efforts should be concentrated on the maintenance of our road as is being championed by FERMA.

References

Chatfield .C. (1987). The Analysis of Time Series. An Introduction. London Chapman and Hall (Third Edition)

Durbin J., and Watson, G. S. (1950) "Testing for Serial Correlation in Least Squares Regression, I. Biometrika 37, 409-428.

Durbin J., and Watson, G. S. (1951) "Testing for Serial Correlation in Least Squares Regression, II. Biometrika 38, 159-178.

Frank Owen and Ron Jones (1982). "Statistics" 2nd Edition, Pitman, Chapter 14 "Sample Design" and Chapter 15, "Planning a Sample Survey"

Marriot, F.H.C. (1990). A dictionary of Statistical Terms, Longman Group United Kingdom, 5th Edition

Murray .R. Spiegel (1960). Theory and Problems on Statistics (First Edition). McGraw-Hill Company.

Sargan, J.D. and Alok Bhargava (1983). "Testing Residuals from Least Squares Regression for being generated by Gaussian Random walk". Econometrics, 51, 153-174.

Von Neumann John (1941). "Distribution of the ratio of the mean square successive difference to the variance". Annals of Mathematical Statistics, 12, 367-395.