
Statistical Analysis of Pattern on Monthly Reported Road Accidents in Nigeria

Oyenuga Iyabode Favour¹, Ayoola Femi Joshua², Shittu Olanrewaju Ismail²

¹Department of Mathematics and Statistics, The Polytechnic, Ibadan, Oyo State, Nigeria

²Department of Statistics, University of Ibadan, Ibadan, Oyo State, Nigeria

Email address:

iyaboadebayo2001@yahoo.com (I. F. Oyenuga), ayoolafemi@yahoo.com (F. J. Ayoola), oishittu@gmail.com (O. I. Shittu)

To cite this article:

Oyenuga Iyabode Favour, Ayoola Femi Joshua, Shittu Olanrewaju Ismail. Statistical Analysis of Pattern on Monthly Reported Road Accidents in Nigeria. *Science Journal of Applied Mathematics and Statistics*. Vol. 4, No. 4, 2016, pp. 119-128.

doi: 10.11648/j.sjams.20160404.11

Received: June 7, 2016; **Accepted:** June 15, 2016; **Published:** June 28, 2016

Abstract: Road accidents in Nigeria have always been in the increase. Efforts made by Federal Road Safety Commission (FRSC) in tackling the menace have not yielded much result. This paper aims to find a suitable time series model to forecast the future characteristics of the road accident data on Oyo-Ibadan express road. The data used for this paper was monthly data collected for a period of Eleven years between 2004-2014. In achieving this, the additive model approach was adopted in the analysis. It includes the estimation of trend, seasonal variation and random variation using moving average method. Autoregressive Moving average model were also fitted to the data and the best order was choosing using Akaike Information Criteria (AIC). The order = $c(1, 1, 2)$, seasonal = $c(1, 1, 2)$ gives the best description of the data with minimum (AIC). A forecast based on the model obtained was made by the use of m-step predictor. The time plot plotted shows that the graphs maintain a constant movement from 2004-2008 but increases abnormally in 2010 and later drop again maintaining appreciable downward movement as the year progresses. Judging from the result, accidents and deaths are higher during the festive period months because of the various festivities lined up during this period, which involve much more traveling than usual.

Keywords: Autoregressive Moving Average, Akaike Information Criteria (AIC), Road Accident, Trend

1. Introduction

Road traffic fatalities are forecast to increase over the next ten years from a current level of more than 1.3 million to more than 1.9 million by 2020. The Commission for Global Road Safety believes that the urgent priority is to halt this appalling and avoidable rise in road injury and then begin to achieve year on year reductions. The world could prevent 5 million deaths and 50 million serious injuries by 2020 by dramatically scaling up investment in road safety, at global, regional and national levels. Each year nearly 1.3 million people die as a result of a road traffic collision, more than 3000 deaths each day and more than half of these people are not travelling in a car. According to John Cohen in his book, causes and prevention of road accidents "says if we had the will, we should find ways, for we cannot assume to the problem of road safety are beyond the wit of man to solve once they are identified. We do not have the will because we

are not sufficiently moved by disaster on the road". [1] Define accident as a chance occurrence, which produce unexpected and unpleasant consequences resulting from unforeseen and often a disastrous event.

According to the UN Secretary General, Ban Ki-Moon, lives will be saved through this decade of action. Following the declaration by the UN in 2011, the Federal Road Safety Commission (FRSC) in Nigeria set out to adopt and domesticate the UN action plan by developing a number of programmes suitable for every road user in the country. Despite integrated efforts towards reducing fatal road accidents, Nigeria still remains one of the worst hit countries. With a human population of about 167 million, a high level of vehicular population estimated at over 7.6 million, a total road length of about 194,000 kilometres (comprising 34,120 km of Federal, 30,500 km of State and 129,580 km of local roads), the country has suffered severe losses to fatal car accidents. Its population density varies in rural and urban areas at about 51. Undoubtedly, this

immense pressure contributes to the high road traffic accidents in the country [2]. Nigeria is ranked second-highest in the rate of road accidents among 193 countries of the world. Thus, Nigeria's annual 8,000 to 10,000 traffic accident deaths between 1980 and 2003 were a major personal and traffic safety problems as well as a terrible waste of human resources for the country. In terms of the personal safety problem, Nigeria is a high risk region with an average of 32 traffic deaths per 1,000 people [3]. This is very high compared with the United State's 1.6 traffic deaths per 1,000 populations and with the United Kingdom's 1.4 deaths per 1,000 people [4]. In terms of traffic safety, there are on average 23 accidents per 1,000 vehicles in Nigeria (i.e. 230 per 10,000 vehicles) far in excess of the accident rate in the USA (2.7 accident per 10,000 vehicles) and the UK (3.2 accidents per 10,000 vehicles).

According to [5], between 1970 and 2001, Nigeria recorded a total of 726,383 road traffic accidents resulting in the death of 208,665 persons and 596,425 injuries. In that period, each succeeding year recorded more accidents, deaths and injuries. Indeed, the Nigeria accident pattern seems to suggest that the better the road, the higher the accident and fatality rate as well as the severity and non-survival indices because of driver noncompliance with speed limits [3, 6, 7]. Road accidents claim the largest toll of human life and tend to be the most serious problem all over the world [8]. Worldwide, the number of people killed in RTA each year is estimated at almost 1.2 million while the number of people injured could be as high as 50 million (WHO, 2004). Currently, motor vehicle accidents rank 9th in order of disease burden and are projected to be ranked 3rd in the year 2020. Nearly three-quarters of deaths resulting from motor vehicle crashes occur in developing countries [9] and this problem appears to be increasing rapidly in these countries [10]. Apart from humanitarian aspect of the problem, traffic accidents and injuries in these countries incur an annual loss of \$65 billion to \$100 billion annually. These costs include both loss of income and the burden placed on families to care for their injured relatives. The Americans bear 11% of the burden of road traffic injury mortality (WHO, 2002). The socio-economic costs of RTA in Nigeria are immense and the direct cost of traffic casualties can perhaps, at best be understood in terms of the labour lost to the nation's economy [11]. It has been estimated that persons injured in accidents on Nigerian highways and streets no longer participate in the economic mainstream and this amounts to a loss of labour of millions of person's years to the nation [12].

It is also evident that Nigeria is worse than most other countries in terms of traffic accidents, in spite of her relatively good road network. As a 2004 World Bank report asserts "from the view – point of road development, Nigeria would no longer be regarded as a developing country [13]. But unlike in most countries where improved road development and vehicle ownership (as barometers of economic advancement) is accompanied by better traffic management, higher road safety awareness, and a relative

decrease in the number of motor accidents, the opposite is true of Nigeria.

The problem of transportation and its safety is of great importance. An analysis of the traffic crashes data recorded over a seven-year period of 2000 - 2006 shows that 98,494 cases of traffic crashes were recorded out of which 28,366 were fatal and resulted into 47,092 deaths [14]. This revealing statistics show that Nigeria is placed among the fore front nations (especially the third world nations) experiencing the highest rate of road tragedies in the world. Many researches had been conducted on road accident in different states in Nigeria: Kogi State, Katsina State, Ibadan [15, 16]. Casualties in road accidents were studied by [17].

Armed robbers attempts and attacks are discovered to be one of the major causes of auto-crash on Nigeria roads [18]. A conducted survey revealed that auto-crash is a function of traffic volume as shown by [19]. The research showed that peak rainy season (June, July and September) which may be due to slippery road and poor visibility of the drivers, and the September, October, November and December in Lagos state Nigeria.

The peak accident that occurs in the dry season is always due to impatience [19, 20]. [21] reported that road failure is another cause of road accidents in Nigeria. According to [22], the causes of road traffic accidents are multi-factorial. These factors can be divided broadly into driver factors, vehicle factors and roadway factors. Accidents can be caused by a combination of these factors.

2. Methodology

A predictable change or pattern in a time series that recurs or repeats over a one-year period can be said to be seasonal. Seasonality can be seen in many time series, and it's more common than you might think. For example, if you live in a climate with cold winters and warm summers, your home's heating costs probably rise in the winter and fall in the summer. You would reasonably expect the seasonality of your heating costs to recur every year. Similarly, a company that sells sunscreen and tanning products would see sales jump up in the summer, but drop in the winter. It is important to remember the effects of seasonality when analysing stocks from a fundamental point of view. The statistical technique used is the class Autoregressive Integrated moving average model (ARIMA). The statistical package used was R-Language.

A time series is a collection of observations of well-defined data items obtained through repeated measurements over time. For example, measuring the value of retail sales each month of the year would comprise a time series. This is because sales revenue is well defined, and consistently measured at equally spaced intervals. Data collected irregularly or only once are not time series. An observed time series can be decomposed into three components: the trend (long term direction), the seasonal (systematic, calendar related movements) and the irregular (unsystematic, short term fluctuations).

2.1. Unit Root Tests

Unit Root tests are usually performed on variables to determine if they stationary (i.e. zero mean and constant variance) and if otherwise, to determine their order of integration (i.e. number of times they are to be differenced to achieve stationarity). The time series characteristics of the variables using the Augmented Dickey-Fuller (ADF) and Phillips-Perron (P-P) tests were examined. Basically, the idea is to ascertain the order of integration of the variables as to whether they are stationary I (0) or non-stationary; and, therefore, the number of times each variable has to be differenced to arrive at stationarity. The standard DF test is carried out by estimating the following;

$$y_t = \rho y_{t-1} + x_t' \delta + \varepsilon_t \tag{1}$$

After subtracting y_{t-1} from both sides of the equation:

$$\Delta y_t = \alpha y_{t-1} + x_t' \delta + \varepsilon_t \tag{2}$$

Where $\alpha = \rho - 1$

The simple Dickey-Fuller unit root test described above is valid only if the series is an AR (1) process. If the series is correlated at higher order lags, the assumption of white noise disturbances ε_t is violated. The Augmented Dickey-Fuller (ADF) test constructs a parametric correction for higher-order correlation by assuming that the y series follows an AR (P) process and adding P lagged difference terms of the dependent variable y to the right-hand side of the test regression:

$$\Delta y_t = \alpha y_{t-1} + x_t' \delta + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \dots + \beta_p \Delta y_{t-p} + v_t \tag{3}$$

The usual practice is to include a number of lags sufficient to remove serial correlation in the residuals and for this; the Akaike Information Criterion is employed. Therefore, the ADF test given in equation (3) above is first used and then the Phillips Perron test described below.

Phillips and Perron propose a non-parametric alternative

- a. finite mean $E(x_t) = E(x_{t-1}) = \dots = \mu$
- b. finite variance $E[(x_t - \mu)^2] = E[(x_{t-1} - \mu)^2] = \dots = \sigma_x^2 = Var_x$
- c. finite autocovariance $E[(x_t - \mu)(x_{t-s} - \mu)] = E[(x_{t-j} - \mu)(x_{t-j-s} - \mu)] = \dots = \gamma_s = Cov(x_{t-j}, x_{t-j-s})$

autocovariance between $x_t, x_{t-s} = \sigma_s = \frac{\gamma_s}{\gamma_0}$

But white noise process does not explain macro variables characterized by persistence so we need Autoregressive (AR) and Moving Average (MA) features.

AR (1): $x_t = \rho x_{t-1} + e_t$ $e_t \sim i.i.d(0, \sigma^2)$, (random walk: $\rho = 1$)

MA (1): $x_t = e_t + \theta e_{t-1}$

More generally:

method of controlling for serial correlation when testing for a unit root. The P-P method estimates the non-augmented DF test equation (2), and modifies the t-ratio of the α coefficient so that serial correlation does not affect the asymptotic distribution of the test statistic. The PP test is based on the statistic:

$$t_\alpha = t_\alpha \left(\frac{\gamma_0}{f_0} \right)^{\frac{1}{2}} - \frac{T(f_0 - \gamma_0)(se(\hat{\alpha}))}{2f_0^{\frac{1}{2}}s} \tag{4}$$

Where $\hat{\alpha}$ is the estimate, and t_α the t-ratio of α , $se(\hat{\alpha})$ is the coefficient standard error, and s is the standard error of the test regression. In addition, γ_0 is a consistent estimate of the error variance in equation (2) (calculated as $(T - K) s^2$ where k is the number of regressors). The remaining term, f_0 , is an estimator of the residual spectrum at frequency zero. Therefore, both equation (3) and (4) are used to test for the stationarity of the variables.

2.2. Stationary Time Series Models

White noise process, covariance stationary process, AR (p), MA (p) and ARIMA processes, stationarity conditions, diagnostic checks.

White noise process $x_t \sim i.i.d N(0, \sigma^2)$

A sequence $\{x_t\}$ is a white noise process if each value in the sequence has

- 1. zero-mean $E(x_t) = E(x_{t-1}) = \dots = 0$
- 2. constant conditional variance $E(x_t^2) = E(x_{t-1}^2) = \dots = \sigma^2 = Var(x_{t-1})$
- 3. is uncorrelated with all other realizations $E(x_t x_{t-s}) = E(x_{t-j} x_{t-j-s}) = \dots = 0 = Cov(x_{t-j}, x_{t-j-s})$

2.2.1. Covariance Stationarity (Weakly Stationarity)

A sequence $\{x_t\}$ is covariance stationary if the mean, variance and auto covariance do not grow over time, i.e. it has

AR (p): $x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \dots + \rho_p x_{t-p} + e_t$

MA (q): $x_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q}$

ARMA (p, q):

$x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \dots + \rho_p x_{t-p} + e_t + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q}$

Using the lag operator:

AR (1): $(1 - \rho L)x_t = e_t$

$$\text{MA (1): } x_t = (1 + \theta L)e_t$$

$$\text{AR (p): } (1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_p L^p)x_t = \rho(L)x_t = e_t$$

$$\text{MA (q): } x_t = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q)e_t = \theta(L)e_t$$

$$\text{ARMA (p, q): } a(L)x_t = b(L)e_t$$

Stationarity Conditions for an AR (1) process

$$(1 - \rho L)x_t = e_t \text{ with } \rho(L) = 1 - \rho L \text{ and substituting for } L:$$

$$\rho(z) = 1 - \rho z$$

The process is stable if $\rho(z) \neq 0$ for all numbers satisfying $|z| \leq 1$. Then we can write

$$x_t = (1 - \rho L)^{-1} e_t = \sum_{i=0}^{\infty} \rho^i e_{t-i} \quad (5)$$

If x is stable, it is covariance stationary:

$$1. E(x_t) = \mu \text{ or } 0 - \text{finite}$$

$$2. Var_x = E[x_t^2] = E\left(\sum_{i=0}^{\infty} \rho^i e_{t-i}\right)^2 = \sigma_e^2 \sum_{i=0}^{\infty} \rho^{2i} = \frac{\sigma_e^2}{1 - \rho^2} = \gamma_0$$

-- finite

$$3. \text{covariances}$$

$$\gamma_1 = E(x_t x_{t-1}) = E[(\rho x_{t-1} + e_t)x_{t-1}] = \rho \sigma_x^2$$

$$\gamma_2 = E(x_t x_{t-2}) = E[(\rho x_{t-2} + e_{t-1} + e_t)x_{t-2}] = \rho^2 E(x_{t-2}^2) = \rho^2 \sigma_x^2$$

$$\gamma_s = E(x_t x_{t-s}) = \rho^s \sigma_x^2 = \rho^s Var(x)$$

Autocorrelations between x_t, x_{t-s} :

$$r_s = \frac{\gamma_s}{\gamma_0} = \rho^s \quad (6)$$

Plot of r_s over time = Autocorrelation function (ACF) or correlogram.

For stationary series, ACF should converge to 0:

$$\lim_{s \rightarrow \infty} r_s = 0 \text{ if } |\rho| < 1$$

$\rho > 0 \rightarrow$ direct convergence

$\rho < 0 \rightarrow$ dampened oscillatory path around 0.

2.2.2. Partial Autocorrelation (PAC)

In AR (p) processes all x 's are correlated even if they do not appear in the regression equation.

AR (1)

$$\gamma_s = \rho^s Var(x_t)$$

$$r_1 = \frac{\gamma_1}{\gamma_0} = \rho ; r_2 = \frac{\gamma_2}{\gamma_0} = \rho^2 = r_1 \rho = r_1^2 ;$$

$$r_3 = \frac{\gamma_3}{\gamma_0} = \rho^3 = r_2 r_1$$

We want to see the direct autocorrelation between x_{t-j} and x_t by controlling for all x 's between the two. For this, construct the demeaned series and form regressions to get the PAC from the ACs.

1st PAC:

$$x_t^* = \phi_{11} x_{t-1}^* + e_t \Rightarrow \phi_{11} = r_1$$

2nd PAC:

$$x_t^* = \phi_{21} x_{t-1}^* + \phi_{22} x_{t-2}^* + e_t \Rightarrow \phi_{22} = \frac{r_2 - r_1^2}{1 - r_1^2}$$

In general, for $s \geq 3$, sth PAC:

$$\phi_{ss} = \frac{r_s - \sum_{j=1}^{s-1} \phi_{s-1,j} r_{s-j}}{1 - \sum_{j=1}^{s-1} \phi_{s-1,j} r_j}, \phi_{sj} = \phi_{s-1,j} - \phi_{ss} \phi_{s-1,s-j} \quad (7)$$

$$\text{Ex: for } s=3, \phi_{33} = \frac{r_3 - (\phi_{21} r_2 + \phi_{22} r_1)}{1 - (\phi_{21} r_1 + \phi_{22} r_2)} \quad (8)$$

Identification for an AR (p) process

$$\text{PACF for } s > p: \phi_{ss} = 0$$

$$\text{Hence AR (1): } \phi_{22} = \frac{r_2 - \phi_{11} r_1}{1 - \phi_{11} r_1} = \frac{r_2 - r_1^2}{1 - r_1^2} = \frac{r_1^2 - r_1^2}{1 - r_1^2} = 0$$

$$\phi_{33} = \frac{r_3 - (\phi_{21} r_2 + \phi_{22} r_1)}{1 - (\phi_{21} r_1 + \phi_{22} r_2)}$$

To evaluate it, use the relation $\phi_{sj} = \phi_{s-1,j} - \phi_{ss} \phi_{s-1,s-j}$:

$$\phi_{21} = \phi_{1,1} - \phi_{22} \phi_{11} = \phi_{11} = r_1, \text{ substitute it to get:}$$

$$\phi_{33} = \frac{r_3 - r_1 r_2}{1 - r_1^2} = \frac{r_1^3 - r_1^3}{1 - r_1^2} = 0 \quad (9)$$

2.3. Stability Condition for an AR (P) Process

$$(1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_p L^p)x_t = \rho(L)x_t = e_t$$

The process is stable if $\rho(z) \neq 0$ for all z satisfying $|z| \leq 1$, or if the roots of the characteristic polynomial lie outside the unit circle. Then, we can write:

$$x_t = \rho(L)^{-1} e_t = a(L)e_t = \sum_{j=0}^{\infty} a_j e_{t-j}$$

Then we have the usual moment conditions:

$$a. E(x_t) = \mu \text{ or } 0 - \text{finite}$$

$$b. Var_x = E[x_t^2] = E\left(\sum_{i=0}^{\infty} a_i e_{t-i}\right)^2, a_0 = 1, e_{t-i} e_{t-j} = 0$$

$$\Rightarrow Var_x = \sigma_e^2 \sum_{i=0}^{\infty} a_i^2 = \frac{\sigma_e^2}{1 - a^2} = \gamma_0 \text{ -- finite variance,}$$

hence time independent.

c. covariances

$$\gamma_s = E(x_t x_{t-s}) = E[(e_t + a_1 e_{t-1} + a_2 e_{t-2} + \dots)(e_{t-s} + a_1 e_{t-s-1} + \dots)] = \rho^2 E(x_{t-2}^2) = \rho^2$$

$$= [(a_s + a_1 a_{1+s} + a_2 a_{2+s} + \dots)] \sigma_e^2 = \sigma_e^2 \sum_{i=0}^{\infty} a_i a_{i+s}$$

finite and time independent.

$$r_1 = \frac{\gamma_1}{\gamma_0} = \frac{a_1 a_2 + a_2 a_3 + a_3 a_4 + \dots}{(a_1 + a_2 + \dots)^2} \quad (10)$$

$$r_s = \frac{\gamma_s}{\gamma_0} = \frac{\sum a_i a_{i+s}}{\sum a_i^2} \quad (11)$$

d. MA process

$x_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q}$, $e = 0$ mean white noise error term.

$$= (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) e_t = \theta(L) e_t$$

If $\theta(z) \neq 0$ for $|z| \leq 1$, the process is invertible, and has an $AR(\infty)$ representation:

$$\begin{aligned} \theta(L)^{-1} x_t &= \alpha(L) x_t = \sum_{j=0}^{\infty} \alpha_j x_{t-j} = e_t \\ \Rightarrow x_t - \sum_{j=1}^{\infty} \alpha_j x_{t-j} &= e_t \\ \Rightarrow x_t &= \sum_{j=1}^{\infty} \alpha_j x_{t-j} + e_t \end{aligned} \quad (12)$$

2.4. Stability Condition for MA (1) Process

$$x_t = e_t + \theta_1 e_{t-1}$$

Invertibility requires $|\theta_1| < 1$

Then the AR representation would be:

$$\begin{aligned} x_t &= (1 + \theta_1 L) e_t \Rightarrow (1 + \theta_1 L)^{-1} x_t = e_t \\ \Rightarrow x_t &= \sum_{i=1}^{\infty} \theta_1^i x_{t-i} + e_t \end{aligned}$$

- $E(x_t) = 0$ finite
- $Var(x_t) = \gamma(0) = \text{var}(e_t + \theta_1 e_{t-1}) = (1 + \theta_1^2) \sigma_e^2$ finite.
- $\gamma(1) = \theta_1 E e_{t-1}^2 = \theta_1 \sigma_e^2 \Rightarrow r_1 = \frac{\gamma(1)}{\gamma(0)} = \frac{\theta_1}{1 + \theta_1^2}$

$\gamma(2) = \gamma(3) = 0 \Rightarrow r_2 = r_3 = 0$, hence autocorrelations' cut off point = lag 1

More generally: AC for MA (q)=0 for lag q.

- PAC:

$$\phi_{11} = r_1 = \frac{\theta_1}{1 + \theta_1^2} \quad (13)$$

$$\phi_{22} = \frac{r_2 - r_1^2}{1 - r_1^2} = \frac{-\theta_1^2}{(1 + \theta_1^2)(1 + \theta_1^2 - \theta_1)} = -\phi_{11} \left(\frac{\phi_{11}}{1 - \phi_{11}^2} \right)$$

$$\begin{aligned} \phi_{33} &= \frac{-\phi_{22} r_1}{1 - \phi_{21} r_1} = \frac{-\phi_{22} \phi_{11}}{1 - \phi_{21} \phi_{11}} = \frac{-\phi_{22} \phi_{11}}{1 - (\phi_{11}(1 - \phi_{22})) \phi_{11}} \\ &= -\phi_{11} \frac{\phi_{22}}{1 - \phi_{22}} \end{aligned} \quad (14)$$

For AR:

AC depends on the AC coefficient (rho), thus tapers off

PAC depends on r_s or ρ^s , cuts off 0 at s (AR (1): cut off

at L=1)

For MA:

AC depends on var of error terms: abrupt cut off

PAC depends on the MA coefficient θ , thus tapers off.

ARMA process

$$\text{ARMA (p, q): } \rho(L) x_t = \theta(L) e_t$$

$$\begin{aligned} (1 - \rho_1 L - \rho_2 L^2 - \dots - \rho_p L^p) x_t &= \\ (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) e_t \end{aligned}$$

If q=0 \rightarrow pure AR (p) process

If p=0 \rightarrow pure MA (q) process

If all characteristics roots of $x_t = \sum_{i=1}^p \rho_i x_{t-i} + \sum_{i=0}^q \theta_i e_{t-i}$ are within the unit circle, then this is an ARMA (p, q) process. If one or more roots lie outside the unit circle, then this is an integrated ARIMA (p, d, q) process.

Stability condition for ARM (1, 1) process

$$x_t = c_1 x_{t-1} + e_t + a_1 e_{t-1} \quad (15)$$

$$(1 - c_1) x_t = (1 + a_1 L) e_t \Rightarrow x_t = \left(\frac{1 + a_1 L}{1 - c_1 L} \right) e_t \quad (16)$$

If $|c_1| < 1$ then we can write

$$\begin{aligned} x_t &= (1 + a_1 L)(1 + c_1 L + (c_1 L)^2 + \dots) e_t \\ &= (1 + c_1 L + (c_1 L)^2 + \dots + a_1 L + a_1 c_1 L^2 + \dots) e_t \end{aligned}$$

$= (1 + (a_1 + c_1)L + c_1(c_1 + a_1)L^2 + c_1^2(c_1 + a_1)L^3 + \dots) e_t$
 \rightarrow an $MA(\infty)$ representation.

$E(x_t) = 0$ finite

$$Var(x_t) = \gamma(0) = \left(1 + \frac{(a_1 + c_1)^2}{1 - c_1^2} \right) \sigma_e^2 \text{ finite}$$

Covariances --finite

$$\gamma(1) = c_1 Var(x_t) + a_1 \sigma_e^2$$

$$\gamma(2) = c_1(c_1 Var(x_t) + a_1 \sigma_e^2) = c_1 \gamma(1)$$

$$\Rightarrow \gamma(j) = c_1 \gamma(j-1), \quad j \geq 2$$

Autocovariance function:

$$r_1 = \frac{\gamma(1)}{\gamma(0)}$$

$$r_2 = c_1 r_1$$

$$r_3 = c_1 r_2 = c_1^2 r_1$$

Any stationary time series can be represented with an ARMA model:

$$AR(p) \Rightarrow MA(\infty)$$

$$MA(p) \Rightarrow AR(\infty)$$

3. Data Analysis

It is organized into two broad sections, namely; descriptive analysis and empirical analysis. Each of these sections is further broken down appropriately. To illustrate this point, we are considering the monthly reported road accidents Monthly Cases of Road Accidents Oyo-Ibadan express way, Oyo State (2004-2014).

Table 1. Monthly reported road accidents.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Total
2004	15	4	7	12	6	8	5	6	5	4	7	11	90
2005	10	6	15	13	5	7	11	9	4	6	4	14	104
2006	8	4	10	5	6	4	10	7	5	9	19	16	103
2007	8	7	16	9	4	6	12	10	7	11	18	20	128
2008	7	16	12	13	5	7	8	12	9	11	19	21	140
2009	7	6	10	8	8	14	16	12	18	10	18	21	148
2010	8	12	18	12	22	11	20	17	19	13	27	20	199
2011	11	12	19	18	15	19	23	25	22	26	13	23	226
2012	26	26	14	16	23	26	24	25	24	13	22	24	263
2013	28	25	25	20	23	22	30	29	19	31	25	26	303
2014	25	21	27	19	22	24	28	25	18	23	18	29	279
Total	153	139	173	145	139	148	187	177	150	157	190	225	1983

Source: FRSC Oyo State

This indicates that 225 cases of accident were recorded in the month of December; 173 in March; and 187 in July. This result shows that fatal accidents happened more during festive and seasonal period.

Represented in the above Fig. 1 is the time plot of all the variables used in this analysis. A critical look at the plots show that the series exhibited trend and seasonal effect. Since the plots cannot provide sufficient evidence to render the series not stationary, it is essential to use standard tests of stationarity.

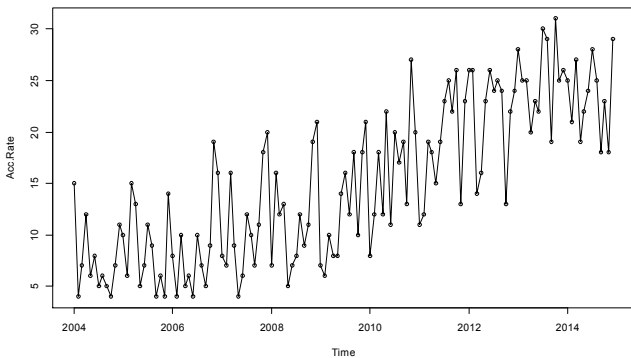


Figure 1. Time Plot of Monthly Reported Road Accidents.

3.1. Stationarity Test

The time series behaviour of each of the series is presented in Tables 2 below, using the ADF and PP tests at both levels.

H₀ = There is unit root

H₁ = There is no unit root

Table 2. ADF and PP Test on Original Data.

	ADF-value	Prob (p-value)	PP-value	Prob (p-value)	
ADF value	-4.7613	0.01	PP-test	-112.99	0.01
Critical-value at 1%	-3.3991				
At 5%	-3.4258				
At 10%	-3.1361				

Since the absolute value of Augmented Dickey-Fuller (ADF) and Philip Perron (PP) test are 4.7613 and 112.99 which is greater than absolute critical value at 1%, 5% and 10% and also the p-value is less than all the critical value. Therefore, this leads to the rejection of the null hypothesis indicating that there is no unit root. This implies that the data is stationary.

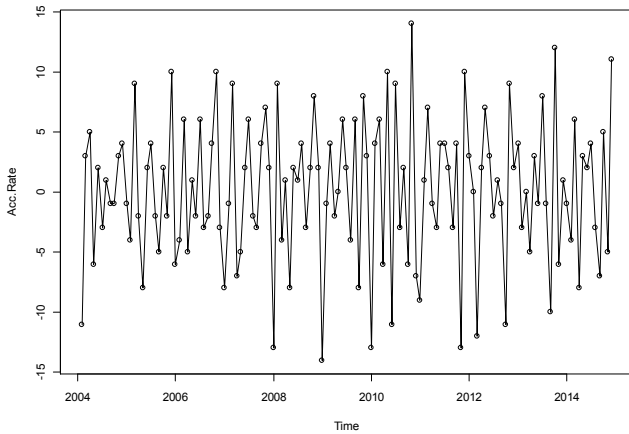


Figure 2. Time Plot of the Data after Applying the First Difference.

Although the ADF and PP test shows that our data is weakly stationary but taking a critical look at the time plot above, we could notice a slight trend of non stationarity. To clear the element of doubt, we can remove this threat of non stationarity by differencing the series and then run the unit root test on the first difference.

This shows that the slight trend of non stationarity has disappeared which shows the real stationarity picture. The following is the first difference ADF and PP test to confirm the claim.

ADF and PP Test of the First Difference of the Data

H_0 = There is unit root

H_1 = There is no unit root

Table 3. ADF and PP Test on First Difference of the Data.

	ADF-value	Prob (p-value)	PP-value	Prob (p-value)
1 st diff ADF value	-7.3095	0.01	PP-test -144.07	0.01
Critical-value at 1%	-3.3991			
At 5%	-3.4258			
At 10%	-3.1361			

This also shows that the absolute value of ADF (7.3095) and PP (144.07) are still greater than all the critical values of 1%, 5% and 10% and also p-value is less than all the critical values, with this, H_0 can be rejected and this implies that the data is stationary.

3.2. Model Fitting, Selection and Diagnostics

We considered both the deterministic and stochastic approach. Under the deterministic, least square trend was fitted for the variables and the estimated trend equation was recorded. The stochastic part include ACF and PACF inspection of the variables, fitting of the appropriate model as subjected by the ACF and PACF, selection of adequate model using Akaike information criterion (AIC) and model diagnostic.

3.2.1. Least Square Trend

The trend equations of the variables are shown below alongside the time plot. This helps us to visualize the movement of the series along the trend line. The trend equation will also help us to determine the growth rate of the series as

well as the direction of the growth with changing in time. Eventually the following trend equation surfaced for the series.

$$Y_t = a + bX_t \text{ Least square model} \quad (17)$$

$$T = a + bt \text{ Trend model by least square.} \quad (18)$$

Table 4. Estimation of the Trend Values By least square.

Variables	Estimate	Std. Error	t-value	Pr (> t)
(Intercept)	-3806.15	250.57	-15.19	<2e-16 ***
Time (Accident)	1.9016	0.1247	15.25	<2e-16 ***

Method: Ordinary Least Square. Period of study: 2004 – 2014, Included Observations: 132

---Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 4.549 on 130 degrees of freedom; Multiple R-squared: 0.6414, Adjusted R-squared: 0.6387; F-statistic: 232.6 on 1 and 130 DF, p-value: < 2.2e-16

Source: R-Language

The value of the intercept which -3806.15 shows that the accident rate at Oyo-Ibadan express way will experience -3806.15 decreases when time is constant. And also there will be 1.9016 increase with unit increase in time.

$$T = -3806.15 + 1.9016t \text{ The Fitted Model } R^2 = 0.6414$$

Road Accident

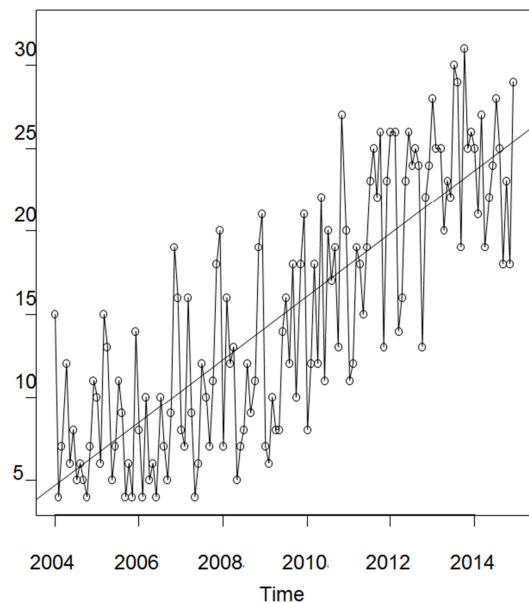


Figure 3. Time Plot of Least Square.

The graph above shows accident rate increasing by $1.9016 \cong 2$ every month.

3.2.2. ACF and PACF of the First Difference of Accident Rate

Examining fig. 3, one could find out that ARMA model should be used for this $data U_t$ but due to the seasonality that was discovered in the time plot fig. 1, we shall use the seasonal ARMA (i.e. SARIMA) model for the data. Here are the feasible tentative models with their Akaike's information criterion (AIC)

Table 5. The Following Are The List Of Possible Models With Their AIC.

Seasonal ARIMA model	AIC	Log-likelihood
order=c (1, 1, 1), seasonal=c (1, 1, 1)	820.97	-408.48
order=c (2, 1, 1), seasonal=c (1, 1, 1)	805.91	-399.91
order=c (1, 1, 1), seasonal=c (1, 1, 2)	790.58	-391.29
order=c (1, 1, 2), seasonal=c (1, 1, 2)	778.53	-386.26
order=c (2, 1, 2), seasonal=c (1, 1, 2)	778.86	-386.26
order=c (3, 1, 2), seasonal=c (1, 1, 2)	780.83	-385.41
order=c (2, 1, 1), seasonal=c (2, 1, 2)	805.91	-399.95
order=c (3, 1, 1), seasonal=c (1, 1, 1)	790.58	-391.29
order=c (3, 1, 3), seasonal=c (1, 1, 2)	780.73	-384.36
order=c (1, 1, 0), seasonal=c (1, 1, 1)	911.62	-454.81
order=c (2, 1, 0), seasonal=c (1, 1, 2)	885.22	-440.61

Source: R-Language

Table 5 above shows that the best model is SARMA= c (1, 1, 2), seasonal=c (1, 1, 2) since it has the smallest AIC = (778.53) and log likelihood = -386.26

Table 6. Parameter Estimation.

	AR (1)	MA (1)	MA (2)
Coefficient	0.1007	-1.9298	0.9405
S. E.	0.0946	0.0381	0.0405
log likelihood = -386.26, AIC = 778.53			

Source: R-Language

3.3. Diagnostics Checking

Checking whether the model capture the data well enough, we carry out the following diagnostics test for residual of the model.

From figure 4 below, it is observed that the residual plot suggested is rectangular scattered and irregular round a zero horizontal level with no trends whatsoever and also the point are within the tolerance line in the p-value, with all this, we can say the model capture the data well enough.

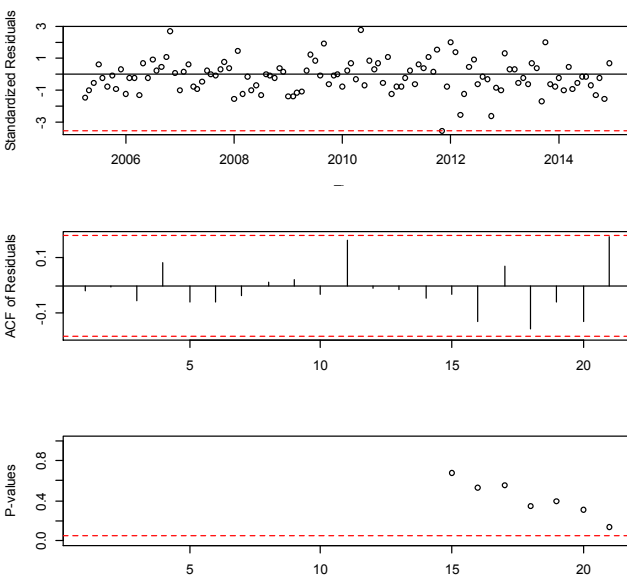


Figure 4. Residual plots.

3.4. Predicted Values with Their Corresponding Standard Error

Table 7. Forecast of the Data values to the nearest whole numbers.

Year/Month	Predicted values	S. E
2015/Jan	-5	5.583937
Feb	-2	6.048035
Mar	2	6.049604
Apr	-5	6.04979
May	2	6.050053
Jun	1	6.05032
Jul	4	6.050587
Aug	-2	6.050854
Sep	-5	6.051121
Oct	2	6.051388
Nov	-1	6.051655
Dec	4	6.051922
2016/Jan	-4	6.225174
Feb	-2	6.330321
Mar	1	6.330338
Apr	-5	6.33074
May	2	6.331174
Jun	0	6.331609
Jul	4	6.332044
Aug	-2	6.33248
Sep	-5	6.332915
Oct	2	6.33335
Nov	-1	6.333785
Dec	4	6.334219
2017/Jan	-4	6.50275
Feb	-2	6.599745
Mar	1	6.599824
Apr	-5	6.600419
May	2	6.601052
Jun	0	6.601687
Jul	4	6.602322
Aug	-2	6.602957
Sep	-5	6.603592
Oct	2	6.604227
Nov	-1	6.604862
Dec	4	6.605495

Source: R-Language

4. Conclusion and Recommendation

Having carried out the statistical analysis with the aid of time series, we obtained some results. In the time series analysis, we used the method of moving averages, autocorrelation and partial autocorrelation to analyse the series on the cases of road accidents along Ibadan-Oyo road, Oyo State. From the summary, results were observed that Christians celebrate Christmas in December and Easter in March or April. Also, there is Eid-el-kabir and Eid-el-tri June-September, within these periods, people travel to celebrate with their loved ones. Accidents and deaths are higher during the festive periods and these EMBER months because of the various festivities lined up during this period, which involve much more traveling than usual. It is a period when commercial drivers make more money through overloading and excessive speeding, among other factors.

Recommendation

Based on the summary of the findings enumerated above on the number of road accidents, the number of cases will be reduced to the beeriest minimum by the government if the following recommendations are considered.

1. Creating awareness about Federal Road Safety: The government will have to increase efforts to promote awareness about the road safety issues and their social economic implications. The strategy to implementing this policy is by rising existing awareness, among stakeholders for planning and promoting road safety and their roles and responsibility.
2. Providing enabling legal, institutional and financial environment for road safety: Many government departments as well as various public and private agencies, share the responsibility of the various safety information data base: Detailed analysis of road accidents is essential if the causes of the accident are to be fully understood. At present time, the policy prepares a report for the accident that they are aware of. Accident report requires a precise location of the accident and condition at the time of the accident.
3. The government should improve data collection details at the scene of accident, improve the storage and accessibility of all data relevant to an accident such as vehicles involved, road, environment and drivers detail, etc.

References

- [1] Onakomaiya, S. O. (1990) Trends in Nigeria Road Safety and Accident Situation: A Paper Presented at a Conference on Accident Control.
- [2] Fedearal Road Safety Corp (2010). Report on Road Traffic Crashes (RTC) Involving Buses on Nigerian Roads (2007-2010)
- [3] Filani, M. O. and Gbadamosi, K. T. (2007) Spatial and Temporal Pattern of Road Traffic Accident Occurrences in Nigeria: 1970-1995. *Nigerian Geographical Journal*, Vol. 5, No. 1, Pp. 55-70.
- [4] Trinca, G. W. et al (1988) *Reducing Traffic Injury; A Global Challenge* A. H. Massina and Co. Melbourne.
- [5] Fedearal Road Safety Corp (2007). An article on FRSC Establishment Act. www.Frsc.gov.ng
- [6] Onakomaiya, S. O. (1988) Unsafe at any Speed: Toward Road Transportation for Survival; *Inaugural Lecture*, University of Ilorin, Ilorin.
- [7] Gbadamosi, K. T. (1994) "Spatial Analysis of Road Accidents in Nigeria, 1970-1990" Unpublished M. Sc. Thesis, Ogun State University, Ago-Iwoye, Nigeria.
- [8] Kual A, Sinha VS, Pathak YK, Singh A, Kopoor AK, Sharma S, Singh S (2005). Fatal Road Traffic Accidents, Study of Distribution, Nature and Type of Injury, *JIAFM*. 27 (2): 71-78.
- [9] Odero W (1998). Alcohol-related road traffic injuries in Eldoret, Kenya, *East Afr. Med. J.* 75: 70 Obinna, C. (2007) Road Traffic Crashes Kill 0.4 m Youths every year. *The Vanguard*. April, 24, P. 35.
- [10] Jacobs, G. D.; Aaron – Thomas, A; Astrop, A. (2000) Estimating Global Road Facilities. Crow Thorne, U. K. *Transport Research Laboratory Report 445*. Pp. 1-35.
- [11] Adekunle JA (2010). Road Traffic accident deaths and socio-economic development in Nigeria. *Int. Rev. Bus. Soc. Sci.* 1 (5): 47-60.
- [12] Pratte D (1998). Road to Ruin: Road Traffic Accidents in the Developing World. *NEXUS*, 13: 46-62.
- [13] World Bank (2004): Development Practice: Sustainable Transport Priorities for Policy Reform Chapter 3, Washington D. C.
- [14] Agbonkhese, O, Yisa, G. L., Agbonkhese, E. G, Akanbi, D. O, Aka, E. O, and Mondigba, E (2013). Road Traffic Accidents in Nigeria: Causes and Preventive Measure. *Civil and Environmental Research*, 3 (13): 90-99.
- [15] Suleiman, I. L., (2011). Road Development and the Development of Katsina State, Katsina, (*JORIND*), 9 (1), 401 – 410.
- [16] Charles, O. B., Olafunmilayo, I. F., Afolabi, E. B., Lola, V. A, Ronke, A and Aderonke, M. B., (2007). Risk factors for road traffic accidents among drivers of public institutions in Ibadan, Nigeria. *African Journal of Health Sciences*, 14 (3-4): 136-142.
- [17] Aderamo, A. J., (2012). Spatial Pattern of Road Traffic Accident Casualties in Nigeria. *Mediterranean Journal of Social Sciences*, 3 (2): 61 – 72.
- [18] Omidiji, A. A and Ibitoye, S. A, (2010). Crime and road crashes prevention in public transportation system in Nigeria: the case study of Kwara, Kogi and Ekiti states. *24th ARRB Conference-Building on 50 Years of Road and Transport Research*, Melbourne Australia.
- [19] Atubi, A. O, (2012). A Monthly Analysis of road Traffic Accident in Selected Local Government Areas of Lagos state, Nigeria. *Mediterranean Journal of Social Sciences*, 3 (11): 47-62
- [20] Salako R. J, Adegoke B. O and Akanmu T. A (2014). Time Series Analysis for modeling detecting Seasonality pattern of autocrash cases recorded at FRSC, Osogbo, Osun State. *International Journal of Engineering and Adv. Technology Studies*. Vol. 2 (4), pp 25-34, December 2014.
- [21] Ndefo O., (2012). Causes of Highway Failure in Nigeria. *International Journal of Engineering and Technology*. 4 (11): 4695-4704.
- [22] Eze, B.(2012). Road Traffic Accidents in Nigeria. A Public Health Problem. *Afrimedical Journal*, 3 (2), 35-36
- [23] Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* AC 19: 716-723.
- [24] Akaike, H. 1978. A Bayesian analysis of the minimum AIC procedure. *Ann. Inst. Statist. Math* 30 A: 9-14.
- [25] Spiegel M. R, Stephens L. J (1999), *Theory and Problems of Statistics* (Third edition)
- [26] Box George and Jenkins, Gwilyn (1976), *Times Series Analysis, Forecasting and Control*, Holden Day, San Francisco.

- [27] Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. 1994: Time Series Analysis, Forecasting and Control, 3rd ed., Prentice Hall, Englewood Cliffs.
- [28] Brockwell, P. J. and Davis, R. A. 2002: Introduction to Time Series and Forecasting, 2nd ed., Springer, Verlag.
- [29] Chatfield (1987), The Analysis of Time Series: An Introduction. London Chapman and Hall (Third Edition)
- [30] McCleary, R. and Hay, R. A., Jr (1980). Applied Time Series Analysis for the Social Sciences, Beverly Hills, CA: Sage.