International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2020): 7.803

Modelling Auto - Crash Cases in Osun State, Nigeria

Salami, Taofeek A[1], Adeyemi, Adekunle M[2], Yusuf, Jamal. A[3]

Department of Mathematics/Statistics, Faculty of Pure and Applied Sciences, Osun State College of Technology, Esa - Oke, Osun State, Nigeria

Abstract: ***This paper examines the best Autoregressive Integrated Moving Average (ARIMA) model to study the trends of auto crash cases in Osun State and forecast beyond 2020 using thedata (2011 - 2020) obtained from Federal Road Safety Commission (FRSC), Osun State. The Time Plot showed that there are no patterns in auto crash cases withthe mean of 262 auto crash cases monthly within the study years. The Auto Correlation Function (ACF) of the study data provide good indication of a non - stationary series. Furthermore, the model fit indicated that ARIMA*** $(\mathbf{1},\mathbf{0},\mathbf{0})$ ***was the best applicable model selected for making the three years forecast ofabout 3100 in year 2021, 3129 in year 2022 and 3131 in year 2023of the expected auto crash casualties in Osun State which revealed an increase in deaths and injuriespredicted years of auto - crash cases.***

Keywords: Road Accident, ARIMA, BIC, R - Square, Auto Crash, Cases, Model, ACF

1. Introduction

The misadventure of transportation and its safety is of great significance. An analysis of the traffic crashes data documented from 2000 to 2006 reveals that 98, 494 cases of traffic crashes were recorded and out of which 28, 366 were fatal, 47, 092 were deadly and resulted into deaths (FRSC, 2009; Agbonkhese et al., 2013). 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: Katshina State (Suleiman, 2011), Ibadan (Charles et al., 2007). Casualties in road accidents were studied by Aderamo (2012). Various studies in traffic behavior had been done to determine the behavior of the traffic flow in expressways. In [Adorna, etal (2012)], it was shown that data signature - based density analysis and space mean speed can provide an efficient and effective representation of traffic behavior. In the results, outliers and potential outliers were determined by using data visualization (data images and Non - Metric Multidimensional Scaling (NMDS)) and confidence measures, providing a glimpse of interesting areas in some segments of the expressway. Close inspection and validation of their time - domain behavior pinpoint to those days of a week with abnormal traffic flow behavior (i.e. prolonged very low or high traffic volume, abrupt spikes or troughs brought about by various traffic disruptions).

The peak accident that occurs in the dry season is always due to impatience (Atubi, 2012). Ndefo reported that road failure is another cause of road accidents in Nigeria (Ndefo, 2012). According to (Eze, 2013), 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. Driver factors solely contributes to about 57 per cent of road traffic accidents and 93 per cent either alone or in combination with other factors. Driver's factors in road traffic accidents are all factors related to drivers and other road users. This may include driver's behaviour, visibility, auditory acuity, decision making ability and reaction speed. Drug and alcohol use while driving is an obvious predictor of road traffic accident, road traffic injury and death. Speeding, travelling too fast for prevailing conditions or above the speed limit, is also a driver factor that contributes to road traffic accidents. The risk of being injured increases exponentially with speed much faster than the average speed. The severity of injury depends on the vehicle speed change at impact and transfer of kinetic energy. Though vehicles travelling slower than average speed are also at increased risk of road traffic accidents, most involve speed too fast for the conditions. Many studies had been carried out on: causes and preventive measures of road accidents (Agbonkhese et al., 2013), effect of visual ability of commercial drivers (Oladehinde et al., 2007). Many agree that in the case of highways and road ways, crashes are not just a matter of luck or misfortune, but a combination of multiple conditions or actions. In this study the term "crash" or "collision" will be used in replace of the term "accident".

2. Literature Review

Typically, the objective of time series analysis is to forecast future values of X (such as X_{t+1}) based on present and past values of X and perhaps also on explanatory variables such as accidents. A model in which future values are forecast purely on the basis of past values of the time series is called an Autoregressive (AR) process. A model in which future values are forecast purely on the basis of past shocks (or noise or random disturbances) is called a Moving Average (MA) process. A model that uses both past values of the time series and past shocks is called an Autoregressive Moving Average (ARMA) process.

These models assume that the time series is stationary - that is the series fluctuates around a time invariant mean, and the variance and Auto - covariance i. e. covariance between X_t and X_{t-s} (for all values of s) do not vary with time. In practice, most time series need to be transformed to achieve stationary. To stabilize variance a logarithm transform is often used - appropriate where the variance of the series increases in proportion to the mean. To stabilize the mean,

differencing is usually employed, such as, first order differencing $Z_t = X_t - X_{t-1}$ which eliminates "drift" but it often needs to be applied twice to eliminate trend. Seasonal differencing is often necessary too. An ARMA model of a differenced series is called an ARIMA model, where the 'I' stand's for Integrated because the output needs to be anti - differenced or integrated; Balogun, *etal*., (2015).

Meyler et al (2010) drew a framework for ARIMA time series models for forecasting Irish inflation. In their research, they emphasized heavily on optimizing forecast performance while focusing more on minimizing out - of - sample forecast errors rather than maximizing in - sample 'goodness of fit'. Stergiou (2014) in his research used ARIMA model technique on a 17 years' time series data (from 1964 to 1980 and 204 observations) of monthly catches of pilchard (Sardinapilchardus) from Greek waters for forecasting up to 12 months ahead and forecasts were compared with actual data for 1981 which was not used in the estimation of the parameters. The research found mean error as 14% suggesting that ARIMA procedure was capable of forecasting the complex dynamics of the Greek pilchard fishery, which, otherwise, was difficult to predict because of the year - to - year changes in oceanographic and biological conditions. Contreras et al (2013) in their study, using ARIMA methodology, provided a method to predict next - day electricity prices both for spot markets and long - term contracts for mainland Spain and Californian markets. In fact a plethora of research studies is available to justify that a careful and precise selection of ARIMA model can be fitted to the time series data of single variable (with any kind of pattern in the series and with autocorrelations between the successive values in the time series) to forecast, with better accuracy, the future values in the series. This study is also an attempt to predict the future production values of sugarcane in India by fitting ARIMA technique on the time series data of past 62 years' production.

Aim: This paper is basically designed to model the Auto Crash cases in Osun State for the periods of the years 2021 to 2023.

Objectives: The objectives of this paper are to:

a) Draw the time plot and evaluate the trendsof auto crash cases in Osun State.
b) Fit various ARIMA models and use the best of them to predict.
c) Predict the auto crash cases for three years in Osun State.

3. Methods and Materials

The study area Osun State is an inland state in South - Western Nigeria which was carved out of old Oyo State on August 27, 1991 and its occupies 9, 251 square kilometres, shares borders withKwara State, Oyo State, Ogun State, Ondo and Ekiti States withan estimated population about 5, 521, 901. The major sub - ethnic groups are Ife, Ijesha, Oyo, Ibolo and Igbomina of the Yoruba people. This research paper adopted secondary data which of road crashes in Osun State and was collected from the record section of the Federal Road Safety Commission of Nigeria (FRSCN), Osun State Command, for a period of ten years (2011 to 2020).

Box - Jenkins (ARIMA) Model

ARIMA models are a class of models that have capabilities to represent stationary as well as non - stationary time series and to produce accurate forecasts based on a description of historical data of single variable. Since it does not assume any particular pattern in the historical data of the time series that is to be forecast, this model is very different from other models used for forecasting. The approach of Box - Jenkins methodology in order to build ARIMA models is based on the following steps: (1) Model Identification, (2) Parameter Estimation and Selection, (3) Diagnostic Checking (or Modal Validation); and (4) Model's use. Since we build an ARIMA model for forecasting of a variable requires to identify the model for fitness.

Model Identification

First stage of ARIMA model building is to identify whether the variable, which is being forecasted, is stationary in time series or not. By stationary we mean, the values of variable over time varies around a constant mean and variance. The ARIMA model cannot be built until we make this series stationary. We first have to difference the time series'd' times to obtain a stationary series in order to have an ARIMA (p, d, q) model with 'd' as the order of differencing used. The best idea is to start with differencing with lowest order (of first order, d=1) and test the data for unit root problems.

ACF Plot

ACF Plot is often used in time series analysis and forecasting. The plot graphically summarize the strength of a relationship between present time series values and its lagged values. Auto - correlation Function (ACF) is a complete auto - correlation function which gives us values of auto - correlation of any series with its lagged values. ACF considers all these components while finding correlations hence its' a 'complete auto - correlation plot' and finding correlations of present with lags like ACF, it finds correlation of the residuals with the next lag value.

Autoregressive Integrated Moving Average (ARIMA)

ARIMA models are, in theory, the most general class of models for forecasting a time series which can be made to be "stationary" by differencing (if necessary), perhaps in conjunction with nonlinear transformations such as logging or deflating (if necessary). A random variable that is a time series is stationary if its statistical properties are all constant over time. A stationary series has no trend, its variations around its mean have a constant amplitude, and it wiggles in a consistent fashion, i. e., its short - term random time patterns always look the same in a statistical sense. The latter condition means that its autocorrelations remain constant over time, or equivalently, that its *power spectrum* remains constant over time. Autoregressive Integrated Moving Average (ARIMA) model of order p, q is given as:

$$\hat{y} = a + \emptyset y_{t-1} + \cdots + \emptyset_p y_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \cdots + \theta_q \omega_{t-q} + \sum_{m=1}^{M} \beta_m X_{m,t}$$

4. Discussion of Results

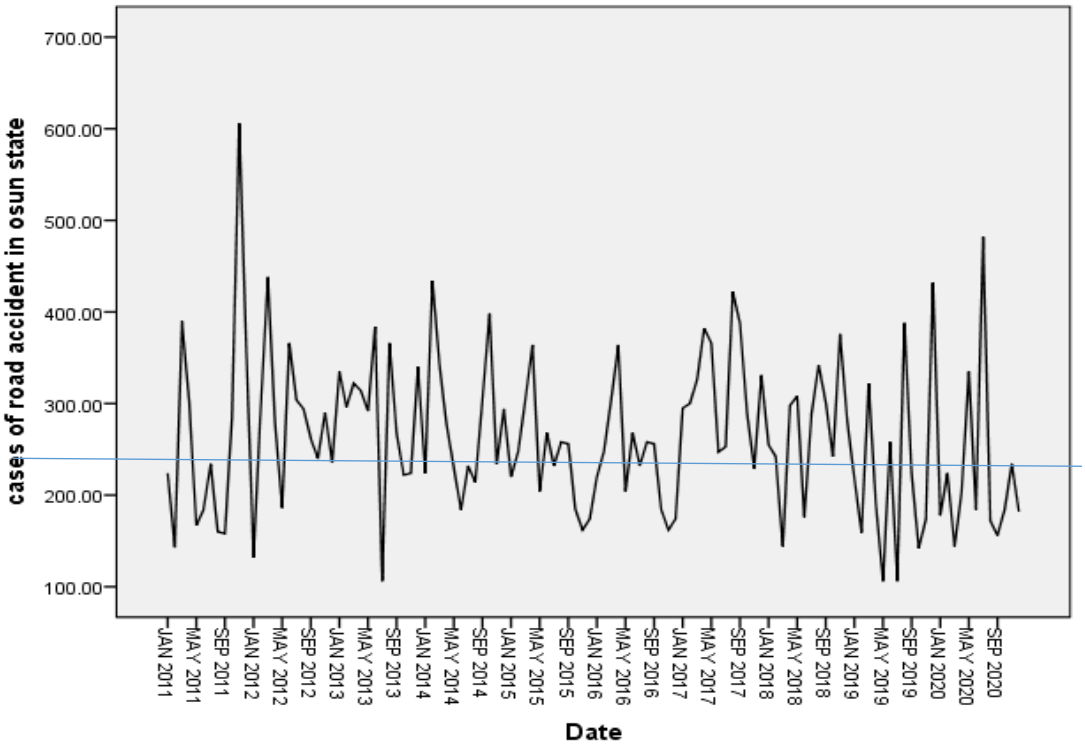

Figure 1: Showing the Time Plot of Auto - Crash Cases in Osun State from original series

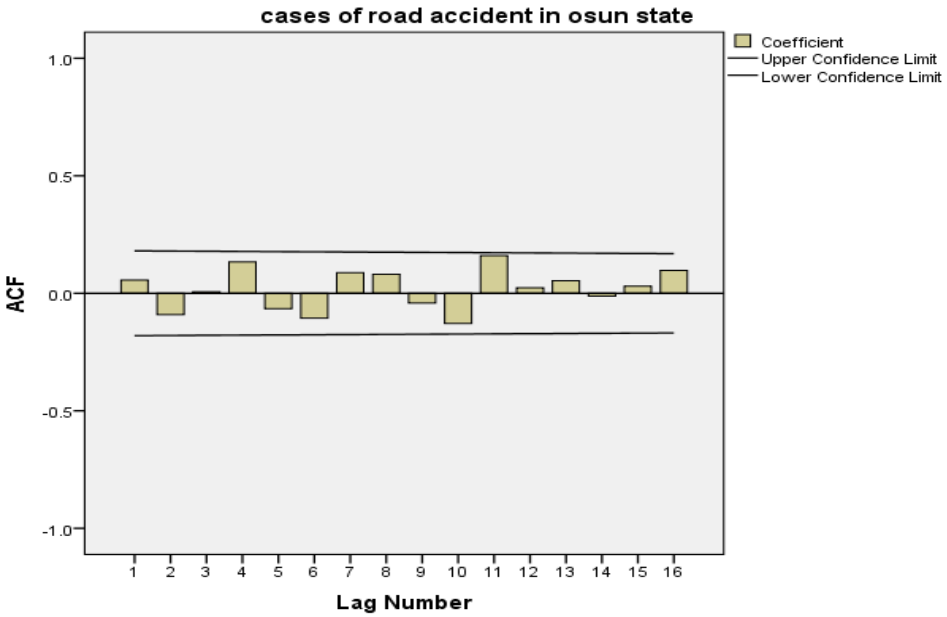

Figure 2: The graph of Autocorrelations (ACF) of the original series

The Fig 1 above showed that there are no patterns or cycles in the auto crash cases data plotted (that is, random variation which means that the auto crash can occur at anywhere anytime) and there is no consistent trend over the entire time span. The series appears to be slowly wander up and down alone the mean of 262 auto crash cases monthly in the study State.

Likewise in Fig 2 above, Auto Correlation Function (ACF) of the study data provide good indication of a non - stationary series. Also if a data series is stationary then the variance of any major subset of the series will differ from the variance of any major subset only by chance.

Table 1: Showing theNormalized Bayesian Information Criteria (BIC) and Stationary R - Squared of fitted ARIMA models

Models	Normalized BIC	Stationary R Squared
1, 0, 0	9.043	0.003
1, 1, 0	9.827	0.425
2, 0, 0	9.494	0.170
0, 2, 0	11.051	0.314
2, 1, 0	9.614	0.487
1, 2, 0	10.614	0.579
0, 0, 1	9.441	0.170
3, 0, 0	9.546	0.171
3, 1, 0	9.727	0.532

From Table 1 we can clearly observed that the lowest BIC and Stationary R - squared values was from the ARIMA $1,0,0$ model with $p = 1, d = 0, q = 0$. Hence this model gives the best model for making forecasts for future values of the Auto Crash Cases data in Osun State.

Forecasting Using Selected Arima Model

Here the best selected model ARIMA 1, 0, 0 will now be used to forecast for the three years (future values) of the Auto Crash Cases data in Osun State.

Table 2: Showing the forecast of Auto - Crase Cases

Model		2021	2022	2023
Total Cases of Auto Crash In Osun State Model 1	Forecast	3099.95	3128.77	3130.74
	UCL	4069.29	4100.37	4102.36
	LCL	2130.61	2157.16	2159.12

5. Conclusion and Future Work

In this study, the Time Plot showed that there are no patterns or cycles in the auto crash cases data plotted (that is, random variation which means that the auto crash can occur at anywhere anytime) and there is no consistent trend over the entire time span. The series appears to be slowly wander up and down alone the mean of 262 auto crash cases monthly in the study State. Likewise the Auto Correlation Function (ACF) of the study data provide good indication of a non - stationary series. Furthermore, the model fit indicated that the lowest BIC (9.043) and Stationary R - squared (0.003) wereobserved form ARIMA $(1,0,0)$ whichwas the best applicable model selected for making forecasts in this study of auto crash cases in Osun State. However, the three years predicted values of the auto crash cases indicated that there will be an increase in deaths and injuries in the coming years as a result of auto - crash cases. Meanwhile the next research will compare the effectiveness between ARIMA and SARIMA modeling.

6. Recommendations

For the purpose of the study and in line with Balogun, (2015), and others, we hereby recommend that (i) attentionshould be drawn to the importance of implementing key road safety measures in order to change the increasing pattern of road accident in Osun State; (ii) an improvement and better policies of Federal Road Safety Commission should be introduced with much emphasis on publication and education to ensure maximum reduction in auto crashes in the study State.

7. Acknowledgements

The research of Balogun, (2015), all the referees used in this paper, the paper reviewer of this research, were all acknowledged. Members of academic and Non Academic staff of Osun State College of Technology, Esa - Oke, and the Federal Road Safety Commission Osun State command were all acknowledged.

References

[1] Aderamo, A. J., (2012). Spatial Pattern of Road Traffic Accident Casualties in Nigeria. Mediterranean Journal of Social Sciences, 3 (2): 61 – 72.

[2] Agbonkhese, O, Yisa, G. L., Agbonkhese, E. G, Akanbi, D. O, Aka, E. O, and Mondigba, E. B., (2013). Road Traffic Accidents in Nigeria: Causes and Preventive Measure. Civil and Environmental Research, 3 (13): 90 - 99.

[3] 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.

[4] Balogun, (2015): Application of ARIMA American Journal of Research Communication Vol 3 (5) 145.

[5] 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.

[6] Contreras, J., Espinola, R., Nogales, F. J., Conejo, A. J., (2013), ARIMA Models to Predict Next - day Electricity Prices, IEEE Transactions on Power Systems, Vol.18, No.3, pp.1014 - 1020.

[7] Eze, B. (2012). Road Traffic Accidents in Nigeria: A Public Health Problem. Afrimedic Journal 3 (2): 35 - 36. FRSC, (2009): Federal Road Safety Commission Annual Report.

[8] FRSC, (2009). Federal Road Safety Commission Annual Report.

[9] Myers, N. and Kent, J. (2010) New Consumers: the influence of affluence on the environment. Proc. Natl Acad. Sci. USA 100, 4963 - 4968.

[10] Ndefo O., (2012). Causes of Highway Failure in Nigeria. International Journal of Engineering and Technology, 4 (11): 4695 - 4704.

[11] Oladehinde, M. K, Adeoye, A. O., Adegbehingbe, B. O and Onakoya, A. O., (2007). Visual functions of commercial drivers in relation to road accidents in Nigeria. India Journal of Occupational and Environmental Medicine, 11 (2): 71 - 75.

[12] Stergiou, K. I. (2014), Modeling and forecasting the fishery for pilchard (Sardinapilchardus) in Greek waters using ARIMA time - series models, ICES Journal of Marine Science, Volume 46, No.1, Pp.16 - 23.

[13] Suleiman, I. L., (2011). Road Development and the Development of Katsina State, Katsina, (JORIND), 9 (1), 401 – 410.