Accident Analysis and Prevention 112 (2018) 21–29

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Road traffic accidents prediction modelling: An analysis of Anambra State, Nigeria

Chukwutoo C. Ihueze, Uchendu O. Onwurah*

Department of Industrial and Production Engineering, Nnamdi Azikiwe University, Awka, Nigeria

ARTICLE INFO

Keywords:
Road traffic crashes
Time series analysis
ARIMA model
ARIMAX model
Forecasting
Anambra State

ABSTRACT

One of the major problems in the world today is the rate of road traffic crashes and deaths on our roads. Majority of these deaths occur in low-and-middle income countries including Nigeria. This study analyzed road traffic crashes in Anambra State, Nigeria with the intention of developing accurate predictive models for forecasting crash frequency in the State using autoregressive integrated moving average (ARIMA) and autoregressive integrated moving average with explanatory variables (ARIMAX) modelling techniques. The result showed that ARIMAX model outperformed the ARIMA (1,1,1) model generated when their performances were compared using the lower Bayesian information criterion, mean absolute percentage error, root mean square error; and higher coefficient of determination (R-Squared) as accuracy measures. The findings of this study reveal that incorporating human, vehicle and environmental related factors in time series analysis of crash dataset produces a more robust predictive model than solely using aggregated crash count. This study contributes to the body of knowledge on road traffic safety and provides an approach to forecasting using many human, vehicle and environmental factors. The recommendations made in this study if applied will help in reducing the number of road traffic crashes in Nigeria.

1. Introduction

One of the major problems in the world today is the rate of crashes and deaths on our roads. Each year, an estimated 1.24 million people are killed in road crashes and up to 20–50 million others injured, costing over US $500 billion worldwide (WHO, 2013). The cost of road traffic injuries is estimated to be between 1–2% gross national product in low-and-middle-income countries, which is over US $100 billion a year (Jacobs et al., 2000). The current trends show that if urgent action is not taken, road traffic injuries could be the seventh leading cause of death by the year 2030, and ninety percent of these deaths occur in low and middle-income countries (WHO, 2015). Not to mention the emotional trauma of losing loved ones, psychological impact on crash victims and permanent disability because of road traffic crashes.

The causes of road traffic crashes (RTC) are multi-factorial and involve interaction of a number of pre-crash factors that include human, vehicles and road environment (Haddon, 1980). Many studies have been conducted to investigate and understand the factors that are contributing to RTC in order to provide countermeasures. For instance, Ojo (2014) examined the factors contributing to road traffic crashes in Ekiti State, Nigeria using linear regression analysis and found that over speeding (speed violation), drivers distraction and dangerous overtaking contributed significantly to road crashes in the State. Ogunmodede et al. (2012) found that over speeding, wrong overtaking, bad roads, sudden mechanical defects, alcoholic intake, tyre burst and heavy rainfall contributed to increasing rate of motorcycle road crashes in Oyo State, Nigeria. Olawole (2016) studied the impact of weather (rainfall and temperature) on road traffic crashes in Ondo State, Nigeria between 2005 and 2012, and found that the correlations between road traffic crashes and elements of weather were generally low and never exceeded 0.41.

Haadi (2014) analyzed the factors that contributed to road crash severity in Ghana's Northern Region using binary logistic regression. The study found that overloading and obstruction were the two most significant factors contributing to road crash severity in Ghana. Also, brake failure has been found by previous studies to be one of the factors that has contributed to road traffic crashes in the developing countries (Oduro, 2012; Oluwole et al., 2015).

The results of most of the road traffic crashes studies in the developed countries have been effective in developing countermeasures that have helped to reduce road traffic crashes in the developed countries (Abdel-Aty, 2003; Noland and Quddus, 2004; Aguero-Valverde and Jovanis, 2006). But, the case has not been the same in some developing countries, more especially in Nigeria where road traffic crashes'

* Corresponding author.
E-mail addresses: cc.ihueze@unizik.edu.ng (C.C. Ihueze), debest2006@yahoo.com (U.O. Onwurah).

https://doi.org/10.1016/j.aap.2017.12.016
Received 26 July 2017; Received in revised form 21 November 2017; Accepted 21 December 2017

statistics still reveal a serious and growing problem with absolute fatality rate rising rapidly (Atubi and Gbadamosi, 2015). The factors that might account for the situation include differences in road users' behaviour, traffic mix, road quality design, unworthy road vehicles, inefficient enforcement of road safety laws, and most importantly, inadequate road crash predictive models that addressed crash contributing factors. This calls for more extensive research on road crashes in order to develop countermeasures and policies that will help to reduce road crashes in Nigeria and position her towards attainment of United Nations decade (2011–2020) of action on road safety.

In order to come up with effective countermeasures or address the contributing factors in Nigeria, there is a need to understand how traffic crashes will change and grow over time. One of the most effective methods of forecasting future occurrences in order to verify significance of certain variables is time series analysis (Mcleod and Vingilis, 2008). The time series analysis method mostly used in road safety research is autoregressive integrated moving average (ARIMA) models proposed by Box and Jenkins (1976). According to Quddus (2008), ARIMA model is the best crash predictive model for aggregated time series count data. Most of the past studies using ARIMA models used only the aggregated crash count without considering the explanatory variables that influenced the crash occurrences (Adu-poku et al., 2014; Avuglah et al., 2014; Balogun et al., 2015; Salifu, 2016; Sanusi et al., 2016).

However, some past studies have incorporated exogenous variables observed over the same period in investigating the relationships between the crash frequency or the severity and the contributing factors (Law et al., 2005; Quddus, 2008; Li and Chen, 2009; Bergel-Hayat et al., 2013; Theofilatos and Yannis, 2014). Although, some variables have been studied, few of these studies in the developed countries examined a comprehensive driver factors that were affecting road crashes. Therefore, more studies are still needed in incorporating broad spectrum of human, vehicle and environmental factors in ARIMA model for analysis and forecasting of RTC.

Hence, this study aims at developing an effective time series crash predictive model incorporating many human, vehicle and environmental (road) factors in comparison to solely using aggregated crash count. This study compares the performances of the univariate time series model (ARIMA model) and the multiple time series model (ARIMAX model) using Bayesian information criteria (BIC), mean absolute percentage error (MAPE), root mean square error (RMSE) and coefficient of determination (R-Squared) as accuracy measures. The model that performs better will be used for future crash analysis and forecasting in Anambra State, Nigeria.

The rest of the paper is organized as follows. Section 2 describes the data sources used for the analysis. Section 3 covers the methods used in data analysis. Section 4 presents the results and discussion of the findings. Section 5 presents the conclusion, recommendation and limitation of this study.

2. Data sources

In Nigeria, the Federal Road Safety Commission (FRSC) is the government agency with the statutory responsibilities for road safety administration. Among the various roles of FRSC are giving prompt attention and care to the victims of crashes, carrying out thorough investigation on the remote and immediate contributing factors to road crashes and filing their reports. They gather the crash information through on the spot assessment of crash scenes, vehicle, environmental conditions, and thorough interviews of the crash victims (drivers and passengers or pedestrians) and the onlookers. The records contain information on the types of road crashes, crash severity (fatal, serious and minor), categories of road users involved, vehicle type, number of vehicles involved, number injured, number killed, causes of crashes, date and time of occurrence, location among other things. The reports record crashes that accounted for at least one minor injury.

For this study, FRSC Anambra State Sector Command supplied data about road traffic crashes in the State for the period 1st January 2005 to 31st December 2015 (a total of 132 months). During this period, 1675 crashes were recorded; 18.84% involved minor injury crashes, 57.30% involved serious injury crashes and 23.86% involved fatal crashes. The 2005 and the 2006 crash data were not used because of some missing values in some months. In this study, an aggregated monthly count dataset of minor, serious and fatal crashes were used in the analysis. Only data from 2007 to 2015 were used. Fig. 1 shows the time series plot of the aggregated monthly count data of the number of crashes.

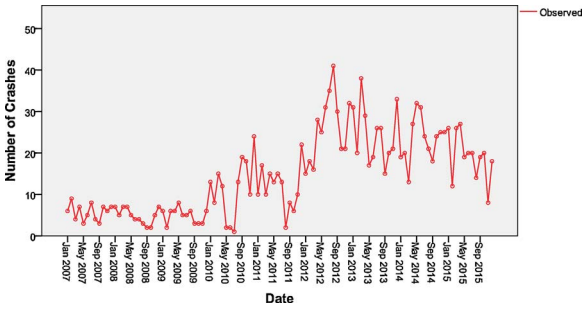

Fig. 1. Time Series Plot of Number of Crashes in Anambra State, Nigeria.

Each crash observation comprises a number of attributes (explanatory variables) relating to the victims of the crash, vehicle, and roadway and environmental conditions as judged by the investigating officers. The crash contributing factors observed over the same period and used in this study are tyre failure/burst, brake failure, over speeding (speed violation), loss of control, route violation, sign light violation, obstruction, wrong and dangerous overtaking, dangerous/reckless driving and weather condition (moderate and heavy rainfall only). The weather condition considered here is road crash that occurred under moderate and heavy rainfall. The descriptive statistics of the response variable and the explanatory variables are presented in Table 1. The crash prediction model that will be developed in this study using the dataset will also investigate the impact of each of the contributing factors on the monthly number of crashes.

Table 1
Descriptive Statistics of Road Traffic Crashes Data.

Variable	Number of months	Minimum	Maximum	Mean	Std. Deviation
Response Variable					
Number of Crashes	108	1	41	14.63	9.815
Explanatory Variables					
Over speeding	108	0	17	4.56	3.880
Tyre burst/ failure	108	0	5	0.80	1.092
Loss of control	108	0	17	3.00	3.496
Wrongful overtaking	108	0	6	0.57	1.061
Brake failure	108	0	8	1.69	1.921
Dangerous overtaking	108	0	5	0.67	0.917
Weather Condition	108	0	5	0.20	0.623
Route violation	108	0	8	0.86	1.300
Obstruction	108	0	3	0.30	0.600
Dangerous driving	108	0	11	2.85	2.335
Sign light violation	108	0	4	0.34	0.763