

1100

B18 (8) (5) (1)

Crash Data Analysis & Prediction in Zimbabwe: An Application of ARIMA Models

Traffic Safety Council of Zimbabwe Operations Research and Marketing

Abstract

Based on annual and quarterly data sets, this study employed Autoregressive Integrated Moving Average (ARIMA) models in order to analyze and predict road traffic crashes in Zimbabwe. Our annual data covers the period 1997 to 2020 while the quarterly data spans over the period March 2016 to September 2021. Optimal models were carefully selected using the Akaike Information Criterion (AIC). The most striking feature of our results is that road traffic fatalities have been forecasted to generally increase over the out-of-sample period. This is paradoxical in the sense that we would normally expect the opposite to happen during a pandemic such as the COVID-19 scourge where restrictions and lockdowns limit the volume of traffic on the road. The overall implication of the study is the need for new and or revised road safety polices that take into account COVID-19 dynamics. A number of policy suggestions have been put forward in order to ensure maximum road safety in Zimbabwe and ultimately save precious lives. The study is quite important as it will foster evidence-based decision making with respect to crash data analysis and prediction in Zimbabwe, in order to strategically reposition the country in terms of road safety policy formulation and implementation.

Keywords: ARIMA Models, road traffic crashes, prediction, Zimbabwe.

1. Introduction

Road traffic today is inherently dangerous (Wegman & Aarts, 2006). It is well-known and appreciated that the road, together with the road user and the vehicle, plays a key role in the cause of crashes on the roads and highways of the world (Asian Development Bank, 2018). In fact, unlike other modes of transport such as railways and air traffic, the road traffic system was not designed with safety as a jumping-off point (Wegman & Aarts, 2006). The geometric design of new roads and the safe management of traffic on existing roads have been critical safety considerations in global efforts to reduce crashes on roads of the world (Asian Development Bank, 2018). Therefore, in road traffic, human beings make the difference between hazard and safety, with little keeping them from harm should we make a mistake (Wegman & Hoekstra, 2011). Each year nearly 1.3 million people die as a result of a road traffic crash—more than 3000 deaths each day—and more than a half of these people are vulnerable road users: pedestrians, cyclists, and motorcyclists. 20 to 50 million more people sustain non-fatal injuries from a crash. These injuries are an important cause of disabilities worldwide. 90% of road traffic deaths occur in low and middle-income countries, which owns about 54% of the world's vehicles. Unless immediate and effective action is taken, road traffic crashes are likely to become the 7th leading cause of death in the world by 2030, resulting in an estimated 1.9 million deaths each year (UN, 2018). Unfortunately, road traffic crashes are often covered in the media simply as events-not as a leading killer of people and an enormous drain on a country's human, health and financial resources (WHO, 2017; Hassouna & Pringle, 2019). The occurrence of crashes can be attributed to driver, vehicle, environment and roadway characteristics (Abdulhafedh, 2016). In Zimbabwe, most crashes are as a result of over-speeding. Furthermore, the state of roads is appalling, characterised by a myriad of potholes in all major roads causing crashes to motorists and pedestrians alike (Mutangi, 2015).

Road traffic crashes continue to increase in developing countries (Bunn *et al.*, 2003; Garcia-Altes & Perez, 2007; Agyemang *et al.*, 2013; Agbeboh & Osarumwense, 2013; Avuglah *et al.*, 2014) such as Zimbabwe (Mutangi, 2015). This problem requires special attention, particularly in low and middle income countries (Bishai & Hyder, 2006). Therefore, crash data analysis and prediction is an essential step towards helping government and transportation agencies in allocating scarce resources more efficiently and also finding preventive measures. Indeed, having predictive information about road traffic crashes (case volumes, deaths and injuries) can provide important information on upcoming changes in crash trends (Zolala *et al.*, 2016).

Research Objectives

The general objective of this study is to analyze and predict crash data in Zimbabwe. Therefore, the following three-fold specific objectives will be pursued:

i. To analyze crash trends in Zimbabwe over the study period.

ii. To develop and estimate a reliable road traffic crash prediction model/(s) for Zimbabwe based on the Box-Jenkins ARIMA technique.

iii. To project road traffic crashes (case volumes, deaths and injuries) in Zimbabwe over the next 5 to 10 years.

Relevance & Timeliness of the Research

Road traffic crash is the 8th cause of mortality around the world (Haghighi *et al.*, 2020). Indeed, serious injuries and mortality in road crashes remain a major health problem with consequences similar to those of major diseases such as cancer and cardiovascular diseases (World Medical Association, 2006). The burden of mortalities, disabilities and injuries due to road traffic crashes has a large effect on health and on the social and economic development of many nations, especially low and middle income countries (Peden, 2005). In fact, Africa grabs the lion's share of the road crash burden relative to its low level of motorization and road network density and experiences the highest per capita rate of road fatalities. The demographic characteristics of road crash victims in the region signifies that over 75% of the casualties are of productive age between 16-65 years; and the vulnerable road users constitute over 65% of the deaths. Road crash costs African countries 1-5% of their Gross Domestic Product (GDP) every year. These figures explicitly indicate the direct linkage and the impact of road crash in worsening poverty in Africa. The regional features such as road network expansion and improvement, rapid motorization, population growth, urbanization, unsafe vehicle fleet and mixed traffic inevitably will worsen road crash deaths and injuries unless African countries significantly invest on road safety (African Development Bank Group, 2014). Road traffic crashes cause not only grief and suffering but also economic losses to victims, their



families, communities and nations, costing countries on average 3% of their GDP. Indirect costs, such as loss of productivity, damage to vehicles and property, reduced quality of life and other factors, must also be included in calculating the true cost to society (WHO, 2014). Despite the growing burden of road traffic crashes, crash data analysis and prediction has received insufficient attention at both the international and national levels. The main aim of the study is to analyze and forecast crash data in Zimbabwe. This is envisioned to go a long way in helping the Government in planning ahead and mobilizing enough resources to fight road carnage in the country.

Contribution of the Paper

There is a dearth of knowledge in Zimbabwe, particularly, on crash data analysis and prediction despite its overwhelming importance in road safety promotion (Zolala *et al.*, 2016). Availability of research-driven evidence in this domain is likely to go a long way in helping policy makers to formulate policies that would have significant impact on the transport sector and thus contribute to sustainable economic growth and development. This paper is apparently in line with the Traffic Act of Zimbabwe, the law that regulates the conduct of motorists on the roads of Zimbabwe. The research also comes at a time when the Government recently launched the National Development Strategy 1 (NDS1) (2021 – 2025) which, amongst other priorities; aims to reduce road crashes and fatalities by 25% margin per annum. Therefore, this study is essential because it is a direct response to national initiatives such as the Traffic Act and NDS 1 and is envisioned to enhance the success of these initiatives. Indeed, the paper will foster evidence-based decision making with respect to crash data analysis and prediction in Zimbabwe, in order to strategically reposition the country in terms of road safety policy formulation and implementation.

2. Methodology

The Autoregressive Integrated Moving Average (ARIMA) Model

The Autoregressive (AR), Moving Average (MA) and the Autoregressive Moving Average (ARMA) processes are normally not applied empirically due to the fact that in most cases many time series data are not stationary; hence the need for differencing until stationarity is achieved.

When the actual data series is differenced "d" times before fitting an ARMA (p, q) process, then the model for the actual undifferenced series is called an ARIMA (p, d, q) model as follows:

ARIMA models have been proposed in literature for both modeling and prediction purposes (Zhou *et al.*, 2014). The Box – Jenkins technique was proposed by Box & Jenkins (1970) and is widely used in many forecasting contexts, including Transport Economics. In this paper, hinged on this technique; the researchers will use the ARIMA approach for analyzing crash data in Zimbabwe. ARIMA models do not assume knowledge of any underlying model or relationship. It is assumed that past values of the series plus previous error terms contain information for the purposes of prediction. The main advantage of ARIMA forecasting is that it requires data on the time series in question only. This avoids the problems such as data unavailability that sometimes occur with multivariate models (Meyler, 1998). In this study, ARIMA models are also considered due to their high level of performance in analyzing linear data sets.

Data Issues

Crash data analysis and prediction has extensively been considered within the framework of time series analysis (Avuglah *et al.*, 2014), a common technique used by numerous research studies to analyze trends of certain phenomena and to predict future conditions (Hassouna & Al-Sahili, 2020). The word "accident" may be taken as meaning that a car crash happened through the fault of nobody in particular; hence the commonly used phrase, "it was just an accident". On the other side of the same coin, the word "crash" implies that someone caused the car wreck to happen, or that someone is actually at fault. According to Nyanzira (2021), approximately 90% of road traffic accidents in Zimbabwe are due to human error. This indicates that most accidents in Zimbabwe occur due to someone's fault. For this reason, and for purposes of this paper, the researchers will proxy crash data using available road traffic accident data. Hence, in this paper, the researcher used secondary data on road traffic accident (crash) case volumes, deaths and injuries. All the data was gathered from the Zimbabwe National Statistics Agency (ZimStats) head office in Harare and covers the period 1997 to 2021. The E-Views version 12 software (automatic ARIMA forecasting option) was used for data analysis.



Evaluation of ARIMA Models

The table below shows the characteristics of the variables, X (annual road traffic crash case volumes), Y (annual road traffic crash deaths), W (annual road traffic crash injuries), M (quarterly road traffic crash case volumes), K (quarterly road traffic crash deaths) and J (quarterly road traffic crash injuries) as well as their corresponding Akaike Information Criteria (AIC) for the selected optimal model. The study used the AIC for optimal model selection instead of using other criteria such as the Schwarz Information Criteria (SIC) because the criteria is more tolerant with higher numbers, and is also good for making asymptotically equivalent cross-validation. When comparing SIC and AIC, penalty for additional parameters is more in SIC than AIC. Furthermore, AIC approach to model selection is best at finding unknown models that have high dimensional reality.

Dependent Variable	Level of Stationarity	Selected ARIMA Model	AIC Statistic		
Annual Road Traffic	I(0)	$(2,0,2)^1$	21.691095		
Crash Case Volumes					
Annual Road Traffic	I(0)	$(2,0,1)^2$	15.414182		
Crash Deaths					
Annual Road Traffic	I(1)	(0,1,1)	20.122647		
Crash Injuries					
Quarterly Road Traffic	I(0)	$(1,0,0)^3$	18.256877		
Crash Case Volumes					
Quarterly Road Traffic	I(0)	$(0,0,0)^4$	11.865250		
Crash Deaths					
Quarterly Road Traffic	I(0)	(1,0,2) ⁵	15.239705		
Crash Injuries					

 Table 1 Evaluation of ARIMA models

As show in table 1 above, most of the variables under consideration are stationary in levels. Only annual road traffic crash injuries data was stationary after taking first differences. The study made use of both the Augmented Dickey Fuller (ADF) and the Dickey Fuller – Generalized Least Squares (DF-GLS) tests for stationarity. The two tests, while different, complement each other in the analysis of unit roots. The ARIMA models, automatically chosen by the E-Views software based on the AIC are as shown above.

3. Analysis and Results

Model Forecasts of the ARIMA (2, 0, 2) Model

¹ This model is the same as the ARMA (2,2) model;

² This model is the same as the ARMA (2,1) model;

³ This model is the same as the ARMA (1,0) model;

⁴ This model is the same as the ARMA (0,0) model;

⁵ This model is the same as the ARMA (1,2) model.





Figure 1 Graphical Presentation of the Out of Sample Forecasts of the ARIMA (2, 0, 2) Model Figure 1 above indicates that road traffic crash case volumes may follow downwards trajectory from 2021 until 2026 after which the case volumes may begin to rise until 2030.



Figure 2 Graphical Presentation of the Out of Sample Forecasts of the of the ARIMA (2, 0, 1) Model

Figure 2 above shows that the road traffic crash deaths are likely to continue on a sharp upwards trajectory until 2024. What is apparent is that throughout the out-of-sample period, crash fatalities will be on the rise. This is not a good sign in as far as road safety is concerned. It is apparently worrisome to note that even during the COVID-19 era, where there are relatively less vehicles on the road; road traffic crash fatalities are increasing.





Figure 3 Graphical Presentation of the Out of Sample Forecasts of the of the ARIMA (0, 1, 1) Model **Table 2** Presentation of the Out of Sample Forecasts of the of the quarterly crash data sets under

ation of the Out of Sample Porecasts of the of the quarterry cras					
	consideration				
	Predicted Quarterly Road	Predicted			

Period	Predicted Quarterly Road	Predicted	Predicted
	Traffic Crash Case Volumes	Quarterly Road	Quarterly Road
		Traffic Crash Deaths	Traffic Crash Injuries
December 2021	11 539.22	417.6957	2 239.312
March 2022	11 529.27	417.6957	2 390.766
June 2022	11 524.71	417.6957	2 423.203
September 2022	11 522.63	417.6957	2 448.052
December 2022	11 521.68	417.6957	2 467.088
March 2023	11 521.25	417.6957	2 481.670
June 2023	11 521.05	417.6957	2 492.841
September 2023	11 520.96	417.6957	2 501.398
December 2023	11 520.91	417.6957	2 507.954
March 2024	11 520.90	417.6957	2 512.976
June 2024	11 520.89	417.6957	2 516.823
September 2024	11 520.88	417.6957	2 519.770
December 2024	11 520.88	417.6957	2 522.027
March 2025	11 520.88	417.6957	2 523.757
June 2025	11 520.88	417.6957	2 525.082
September 2025	11 520.88	417.6957	2 526.097

Table 2 above indicates that quarterly road traffic crash case volumes, deaths and injuries will most likely hover around 11 500, 418, and 2500; respectively, over the out-of-sample period. For a small country like Zimbabwe, these are very high numbers, that apparently signify the need for urgent action if road safety is anything to go by in the country.



4. Conclusions

There is no doubt, transport has become an essential part of everyone's life. With this ever-increasing importance of road usage, the number of road traffic crashes have certainly increased, not only in Zimbabwe but also across the globe. Indeed, car crashes are incredibly common. To minimize these unfortunate eventualities, road safety has been one of the greatest concerns for everyone, particularly, governments. In this study, we examined both annual and quarterly crash data for Zimbabwe. Six optimal univariate Box-Jenkins models were finally presented. Indeed, road traffic crashes have remained a major cause of death globally, even though every one of those deaths and injuries is preventable (WHO, 2021). Below are practical recommendations emanating from the study; The Government of Zimbabwe should:

- Make sure that road safety is one of the country's major political priorities;
- With the help from the Zimbabwe Republic Police (ZRP), enforce safe driving practices;
- Install and maintain appropriate road signs on new and existing roads;
- Ensure the installation of visible road markings and signs that are less prone to vandalism;
- Establish and implement a national road safety plan to achieve the target of reducing road traffic crashes by at least 25% by the year 2025;
- Support the creation of road safety advocacy groups. A decision can be made to have at least 2 road safety advocacy groups per province, countrywide;
- Through the Road Traffic Act, strictly enforce legislation requiring the use of seat-belts and child restraints, adherence to appropriate speed limits, preventing alcohol-impaired driving and the wearing of motorcycle helmets and bicycle helmets;
- Set and enforce strict and uniform vehicle safety standards;
- Establish data collection systems designed to collect and analyse data and use the data to improve safety. In this regard, the Government, through its lead agency, the Traffic Safety Council of Zimbabwe (TSCZ), should have dedicated data collection officers stationed, preferably at the Police General Head Quarters (PGHQ) and be solely responsible for gathering data using a TSCZ-tailored questionnaire. This will go a long way in enabling TSCZ in effectively collecting, collating and analysing road traffic crash data;
- Through the Ministry of Transport and Infrastructural Development, set appropriate design standards for roads that promote safety for all;
- Through the Ministry of Health and Child Care, include road safety in health promotion and disease prevention activities;
- Through the Ministry of Health and Child Care, develop trauma care skills of medical personnel at the primary, secondary, tertiary and quaternary health care levels throughout the country. This will go a long way in improving the management of survivors of road traffic crashes;

Donors, civil society groups and private sector players have a role to play in reducing road traffic crashes in Zimbabwe by:

- Including road safety components in grants for health, transport, environmental and educational programmes;
- Support research, programmes and policies on road safety in the country;
- Set up mechanisms to fund the sharing of knowledge and the promotion of road safety in the country;
- Facilitate road traffic safety management capacity building at community, district, provincial and national levels.

Communities and individuals should:

- Encourage governments to make the roads safe;
- Identify road traffic safety problems;
- Help plan safe and efficient transport systems that accommodate drivers as well as vulnerable road users, such as bicyclists and pedestrians;
- Demand the provision of safety features, such as seat-belts, in cars.
- Encourage enforcement of traffic safety laws and regulations, and campaign for firm and swift punishment for traffic offenders;
- Behave responsibly by abiding by the speed limit on roads; never driving when over the legal alcohol limit; always wearing a seat-belt and properly restraining children, even on short trips; as well as wearing a crash helmet when riding a two-wheeler.

Given these policy directions, the ball is now in the hands of those responsible for policy implementation. However, the study is not without its own limitation, firstly, in terms of its methodological approach: we do not compare our optimal ARIMA models with other applicable predictive models such as Artificial Neural Network (ANN) models. Secondly, the study only includes aggregated data, that is, our data sets are not split according to any demographic characteristics of road traffic crash victims, that is, gender (males and females), and or age groups (adults, children and or young persons) involved in road traffic crashes. When data is split according to gender,

for example; the researcher can be able to analyze and forecast road traffic crashes based on gender. In such a scenario, researchers would be able to determine which gender is at more risk of being involved in RTAs in both the in-and-out- of sample data sets. Similarly, age groups help the researcher to analyze and forecast crash data for young persons or youths and compare results using data sets for adults. Indeed, more insights could be drawn. Thirdly, the study exclusively resorted on the AIC for model selection. However, further studies may need to explore more robust model selection criteria such as the Minimum Description Length (MDL), which is capable of overcoming the limitations of AIC and other criteria such as the SIC. Last but not least, it is important for everyone to remember that road safety is a shared responsibility. Everyone has a role to play despite the fact that the Government will definitely take a leading role in the fight against road crashes.

Acknowledgment

The authors wish to express their gratitude to the management and staff of ZimStats, Harare, for availing data on road traffic crashes.

References

- 1. Abdulhafedh, A. Crash Frequency Analysis. Journal of Transportation Technologies, 2016. 6, p. 169 180.
- 2. African Development Bank Group. New Roads and Schemes: Road Safety Audit, Transport and ICT Department, 2014. African Development Bank Group.
- 3. Agbeboh, G. U., and Osarumwense, O. I. Empirical Analysis of Road Traffic Accidents: A Case Study of Kogi State, North Central Nigeria. International of Physical Sciences, 2013. 8, p. 1923 1933.
- 4. Agyemang, B., Abledu, G. K., and Semevoh, R. Regression Analysis of Road Traffic Accidents and Population Growth in Ghana. International Journal of Business and Social Research, 2013. 3, p. 41 50.
- 5. Asian Development Bank 2018. Roadside Hazard Management. Asian Development Bank, Metro Manilla.
- 6. Avuglah, R. K., Poku, K. A., and Harris, E. Application of ARIMA Models to Road Traffic Accident Cases in Ghana, International Journal of Statistics and Applications, 2014. 4, p. 233 239.
- Bishai, D. M., and Hyder, A. A. Modeling the Cost Effectiveness of Injury Interventions in Lower- and Middle-Income Countries: Opportunities and Challenges, Cost Effective Resource Allocation, 2016. 4, p. 1–9.
- 8. Box, D. E., & Jenkins, G. M. C. Time Series Analysis, Forecasting and Control, Holden Day, London, 1970.
- 9. Bunn, F., Collier, T., Frost, C., Ker, K., Roberts, I., and Wentz, R. Traffic Calming for the Prevention of Road Traffic Injuries: Systematic Review and Meta Analysis. Injury Prevention, 2003. 9, p. 200 204.
- 10. Garcia-Altes, A., & Perez, K. The Economic Cost of Road Traffic Crashes in an Urban Setting. International Journal of Injury Prevention, 2007. 13, p. 65 – 68.
- 11. Getahun, K. A. Time Series Modeling of Road Traffic Accidents in Amhara Region. Journal of Big Data, 2021. 8, p. 1 15.
- 12. Haghighi, M. R., Sayari, M., Ghahramani, S., and Lankarani, K. B. Social, Economic and Legislative Factors and Global Road Traffic Fatalities. BMC Public Health, 2020. 20, p. 1 12.
- 13. Halim, H., Bustam, B., and Saing, Z. The Forecasting of a Traffic Accident in the Pandemic of COVID-19. Turkish Journal of Computer and Mathematics Education, 2021. 12, p. 3664 – 3669.
- 14. Hassouna, F. M. A., and Al-Sahili, K. Practical Minimum Sample Size for Road Crash Time Series Prediction Models. Advances in Civil Engineering, 2020. P. 1 12.
- 15. Hassouna, F. M. A., and Pringle, I. Analysis and Prediction of Crash Fatalities in Australia. The Open Transportation Journal, 2019. 13, p. 134 140.
- Husin, W. Z. W., Afdzal, A. S., Azmi, N. L. H., and Hamadi, S. A. T. S. Box-Jenkins and State Space Models in Forecasting Malaysia Road Accident Cases. Journal of Physics: Conference Series, 2021. 2084, p. 1 – 15.
- 17. Jomnonkwao, S., Uttra, S., and Ratanavaraha, V. Forecasting Road Traffic Deaths in Thailand: Application of Time Series, Curve Estimation, Multiple Linear Regression and Path Analysis Models. Sustainability, 2020. 12, p. 1 − 17.
- 18. Meibner, K., and Rieck, J. Multivariate Forecasting of Road Accidents Based on Geographically Separated Data. Vietnam Journal of Computer Science, 2021. 8, p. 1 22.
- 19. Meyler, A., Geof, K., and Quin, T. Forecasting Irish Inflation Using ARIMA Models, University Library of Munich, MPRA Paper No. 11359.



- 20. Mutangi, K. Time Series Analysis of Road Traffic Accidents in Zimbabwe. International Journal of Statistics and Applications, 2015. 5, p: 141 149.
- 21. Nyanzira, M. Traffic Safety Council Sets Accident Reduction Target, Chronicle Newspapers, Harare, 2021.
- 22. Nyoni, T. Box-Jenkins ARIMA Approach to Predicting Net FDI Inflows in Zimbabwe, 2018. University Library of Munich, MPRA Paper No. 87737.
- Okoli, C. N. Modeling and Forecasting of Road Traffic Crash and Death Occurrences in Anambra State, Nigeria. COOU Journal of Physical Sciences, 2020. 3, p. 493 – 506.
- 24. Peden, M. Global Collaboration on Road Traffic Injury Prevention. International Journal of Injury Control and Safety Promotion, 2005. 12, p. 85 91.
- 25. Sarani, R., Rahim, S. A. S., Marjan, J. M., and Voon, W. S. Predicitng Malaysian Road Fatalities for the Year 2020, 2012. Research Report No. 6, Malaysian Institute of Road Safety Research (MIROS), Negara.
- Sekadakis, M., Katrakazas, C., Michelaraki, E., Kehagia, F., and Yannis, G. Analysis of the Impact of COVID-19 on Collisions, Fatalities and Injuries Using Time Series Forecasting: The Case of Greece. Accident Analysis and Prevention, 2021. 162, p. 1 – 12.
- 27. Twenefour, F. B. K., Ayitey, E., Kangah, J., and Brew, L. Time Series Analysis of Road Traffic Accidents in Ghana. Asian Journal of Probability and Statistics, 2021. 11, p. 12 20.
- 28. UN. Road Safety Strategy, 2018. UN, New York.
- Wegman, F., and Aarts, L. T. Advancing Sustainable Safety: National Road Safety Outlook for 2005 2020, SWOV Institute for Road Safety Research, 2006. Leidschendam.
- 30. Wegman, F., and Hoekstra, T. Improving the Effectiveness of Road Safety Campaigns: Current and New Practices, IATSS Research, 2011. 34, p. 80 86.
- 31. WHO. Road Safety: Basic Facts, WHO, Geneva, 2017.
- 32. WHO. The Decade of Action for Road Safety: 2021 2030, WHO, Geneva, 2021.
- 33. WHO. Road Safety: Basic Facts, WHO, Geneva, 2014.
- World Medical Association. Statement on Traffic Injury. World Medical Association General Assembly, 2006. Pilanesburg.
- 35. Zhou, B., He, D., and Sun, Z. Traffic Modeling and Prediction Using ARIMA/GARCH Model, Centre for Communication System Research, 2014. University of Surrey, Guildford.
- 36. Zolala, F., Haghdoost, A. A., Touraj, A., Salari, A., Abbas, B., Baneshi, M. R., and Razzaghi, A. Forecasting the Trend of Traffic Accident Mortality in West Iran, 2016. Health Scope, 5, p. 1 5.