

Prediction of Road Traffic Accidents Outcomes on Selected Routes in Ibadan Environs Using Time-Series Models

*¹Akintayo, F. O., ²Ajaelu, C. S., ¹Okunlola, O. and ¹Komolafe, O.

¹Department of Civil Engineering, University of Ibadan, Ibadan, Nigeria

²Department of Biomedical Engineering, University of Ibadan, Ibadan, Nigeria

*Corresponding Author: fo.akintayo@ui.edu.ng

ABSTRACT

Road traffic accidents (RTAs) has emerged as a major public concern due to the associated burdens of injuries, loss of lives and properties. The objective of this paper is to use time-series technique applying autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA) models to predict accidents outcomes within and around Ibadan metropolis for appropriate mitigation measures. Data on daily recorded cases of road traffic accidents on seven specified routes between 2019 and 2021 were collected from the Federal Road Safety Corps, Ibadan Zonal Office. The RTAs outcomes were grouped into three: injuries, fatalities and vehicle occupants. The training set consisted of data from December 2019 to June 2021, whereas the testing set was composed of data from July 2021 to December 2021. During the study, 427 fatalities and 2,245 injuries were caused by traffic accidents involving a total of 5,577 vehicle occupants. By the end of 2025, the study projected a 172% rise in the death rate and a 38% increase in the number of vehicle occupants involved in RTAs in the study area, and approximately 32% of those occupants suffering injuries. There was no discernible seasonal pattern in the trends of injury and fatality rates. There is an immediate need for a road traffic accident prevention and evaluation programme in the study region given the projected number of injuries and fatalities.

Keywords: Prediction, Road safety, Road traffic accidents, Time-series models.

1. INTRODUCTION

Annually, about 1.35 million death and 20 to 50 million non-fatal injuries are attributed to road traffic accidents (RTAs) globally (Adejugbagbe *et al.*, 2015; Akinleye and Tijani, 2017; He *et al.*, 2021). The RTAs have become a major public concern of this century and no nation seems to be exempted from its burden. Studies have estimated RTA as the fifth leading cause of death globally by 2030. Road accidents account for about 162 deaths per 100,000 people in Nigeria (Ogbodo and Nduoma, 2011). According to the Road Transport Data Q4 2023 report from the Federal Road Safety Corps (FRSC) to the National Bureau of Statistics (NBS, 2024), there were 4,195 vehicles involved in traffic crashes in Q4 2023, which was higher than the 3,371 vehicles recorded in the Q3 quarter, indicating a 24.44% increase. It is critical to note that the estimated rates of RTA in Nigeria is notably higher than the global average of 22 fatalities per 100,000 people (Shaik *et al.*, 2021). United States loses about \$871 billion annually to traffic accidents and it also costs most other countries about 3% of their GDP (He *et al.*, 2021). Given the burden of the RTAs on individuals, nations and the economy, there is need for adoption of traffic management system which uses predictive information to forecast the possible RTAs and to gain an understanding of the severity and speed-up decision making for mitigation measures (Çodur and Tortum, 2015; Chavhan and Venkataram 2020). Prediction models can help in traffic planning and management and hence reducing RTAs (Santos *et al.*, 2021).

Machine learning techniques such as Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving Average (SARIMA) and Artificial Neural Networks (ANNs) are becoming very useful tools in different sectors such as transportation and health to analyze, visualize and

make predictions from big data which are generated from various organizations. Machine learning algorithms can analyze a large volume of data beyond the capacity of traditional statistical methods (Dipto *et al.*, 2020). These techniques have been applied in predictions of road accidents from weather conditions and other variables influencing road accidents (Moleshi *et al.*, 2021). Gatarić *et al.* (2023) constructed two ANN models to forecast traffic accidents on common roads in the Republic of Serbia and the Republic of Srpska (Bosnia and Herzegovina) utilising objective criteria and terrain type, road width, average daily traffic volume, and speed limit. Both the quantity and severity of traffic collision outcomes, such as injuries, property damage, and fatalities, in the study were predicted by the models. The SARIMA and ARIMA algorithms have been used in the R software package to forecast injuries from traffic accidents involving various modes of transport (Parvareh *et al.*, 2018). However, this study focused on Ibadan with the aim of providing modelled predictive traffic management intervention. This study used time series approaches to estimate RTAs fatalities and injuries, as well as to model the patterns of RTAs on selected roads in Ibadan.

2. METHODOLOGY

2.1 Study Design and Population

The study area in Ibadan (7.3775° N, 3.9470° E) covers the following routes: Ibadan-Ife, Oyo-Ibadan, Eleyele-Iwo Road, Eleyele-Ologuneru, Lagos-Ibadan, Ibadan-Abeokuta, Old-Oyo and Lagun-Iyana Offa (see Figure 1). A cross-sectional study was carried out to analyze the time-series observations of road traffic accidents outcomes (injuries, fatalities and vehicle occupants) from December 2019 to June 2021 from RTAs data obtained from the FRSC, Ibadan Zonal Office. The FRSC is responsible for keeping daily record of RTAs data in Nigeria and the database usually contain variables such as date, vehicle category, number of injuries, number of fatalities, number of people involved and routes. Data cleaning was done to remove duplicates and redundant data. The RTAs outcomes were divided into three categories including number of fatalities, number of injuries and vehicle occupants involved and each category was analyzed separately.

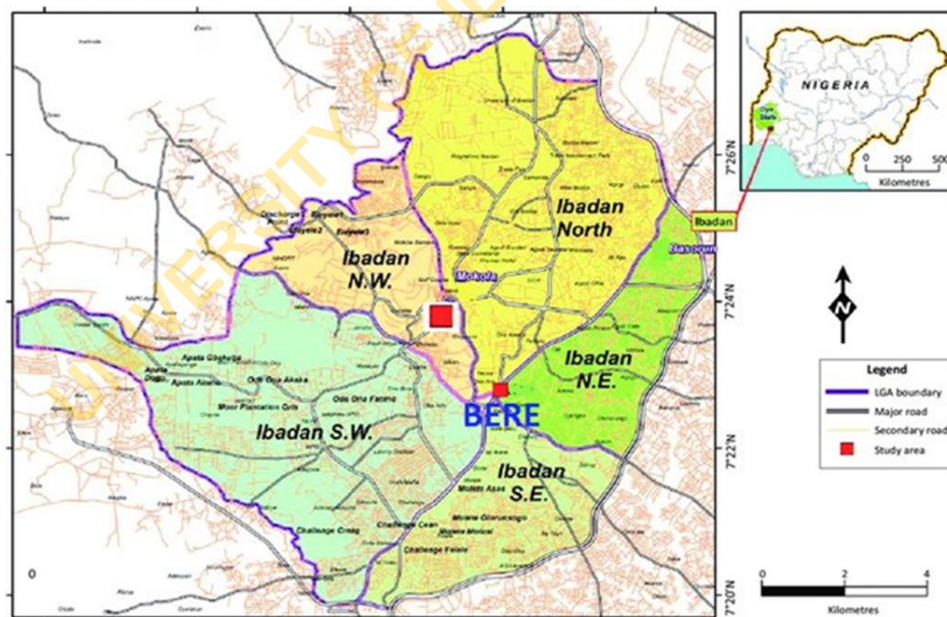


Figure 1: Map of Ibadan showing the Road Networks in major Local Government Areas
(Source: Salami *et al.*, 2017)

2.2 Data Analysis

Time series analysis was carried out to identify patterns of RTAs outcomes (injuries, fatalities and vehicle occupancy) and also to predict future occurrence till 2025 as demonstrated by Twenefour *et al.* (2021). The ARIMA and SARIMA models, which are mathematical models that use parametric method for time-series analysis were used to model the number of injuries, fatalities and vehicle occupants involved in RTAs in

the study area. The mixture of autoregressive (AR), moving average (MA), auto regressive moving average (ARMA), autoregressive integrated moving average (ARIMA) and seasonal autoregressive integrated moving average (SARIMA) were used as described by Parvareh *et al.* (2018). The non-stationarity in time-series is achieved by integrating ARMA (p, q) to d^{th} order to create ARIMA (p, d, q) model, where d is the number of non-seasonal differences needed for stationarity (Bozkurt *et al.*, 2017). The ARIMA model is mathematically expressed as in equation 1. The ARIMA model incorporates the AR model, integration (I), and MA model (Getahun, 2021). The SARIMA model (equation 2) is an extended version of ARIMA with additional seasonal terms.

$$R_t = C + \sum_{i=1}^p \phi_i Y_{t-i} + \epsilon_t - \sum_{j=1}^q \theta_j \epsilon_{t-j} \quad (1)$$

Where:

- ϕ_i ($i = 1, 2 \dots p$) is autoregressive model parameters,
- θ_j ($j = 1, 2 \dots q$) is moving average model parameters,
- R_t is the current time-series value,
- $R_{t-1}, R_{t-2} \dots, R_{t-p}$ is past values,
- ϵ_t is a random error,
- C is a constant.

$$\text{SARIMA}, Y_t = \text{ARIMA}(p, d, q) \times (P, D, Q)_s \quad (2)$$

Where:

P is the degree of seasonal AR model, Q is the degree of seasonal MA model, D is the degree of seasonal integration, s is the span of repeating seasonal pattern, Y_t represents outcomes of interest in each model (number of injuries, number fatalities and vehicle occupants).

The daily observation time-series was used in this study. The following steps were employed;

- i. Model identification using plots of the data, the autocorrelation function (ACF) and the partial autocorrelation function (PACF), which helped in the selection of appropriate values for p, d and q.
- ii. Model estimation: The model parameters were estimated by least-squared or maximum likelihood.
- iii. Performance evaluation: The models were trained with observed frequency of the variables of interest between December 2019 and December 2020 (592 daily observations) and tested with January 2021 to June 2021 dataset (328 daily observations). The models were compared with naïve model using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics, then the forecasting model was accepted if it has better performance than the naïve model. Otherwise, new parameters for model were estimated with a repeated procedure until a set of model parameters that satisfy our predefined criteria was found (Ndume *et al.*, 2022).
- iv. Between 2022 and 2025, future trends in RTA injuries, fatalities, and vehicle occupants were predicted using the fitted models.
- v. The fitted accident injuries model was then used to forecast number of injuries from 2022 to 2025. All analyses and forecast were computed using Python 3.0 version.

3. RESULTS AND DISCUSSION

3.1 Accident Outcomes, Injuries, Fatality and Vehicle Occupants

Table 1 shows the accident outcomes between December 2019 and June 2021 in the study area. A total of 5,577 individuals were reported to have been involved in traffic accidents between December 2019 and June 2021 (Table 1); 2,290 people were hurt (2,192 adults and 98 children), while there were 427 fatalities (396 adults and 31 children). This high percentage of RTAs outcomes in the study area is a major threat to the public health and is consistent with a WHO (2023) estimate that 92% of child fatalities resulting from

road traffic accidents occur in low-income countries and that younger road users will participate in RTAs in these countries as reported and other similar studies also emphasis the increase in RTAs in Nigeria (Onyemaechi and Ofoma, 2016; NBS, 2023).

Table 1: Accident outcomes between December 2019 and June 2021 in the study area

Accident Outcomes	Adult	Children	Total
Vehicle Occupants	4,354	1,223	5,577
Injuries	2,192	98	2,290
Fatalities	396	31	427

Figure 2 presents the trend of accident injuries including (a) actual injury trend and predictions using ARIMA and SARIMA (b) ACF and PACF plots with (c and d) showing a non- seasonal pattern in data. Figure 3 shows the trend of accident fatality with (a) displaying actual fatality trend and predictions using ARIMA and SARIMA (b) ACF and PACF plots and (c and d) showing a non-seasonal pattern in data. Consequently, Figure 4 displays the trend of accident vehicle occupants including (a) actual vehicle occupant trend and predictions using ARIMA and SARIMA (b) ACF and PACF plots with (a and d) presenting a seasonal pattern in data. The accident injuries and fatalities data showed non-seasonality with lower frequency between March and July 2021 which peaked in August (Figure 2a and 3a respectively). The ACF and PACF plots shown in Figures 2c, 2d, 3c and 3d reflected stationarity in the mean of the accident injuries and fatality data. The final models which showed the best goodness of fit (Table 2) for accident injuries and fatalities and vehicle occupants are ARIMA (1, 0, 1), ARIMA (2, 1, 3) and SARIMA (4, 0, 3) (1, 0, 1)₁₂ respectively. The models were used to forecast the accident outcomes between January and June 2021 and compared with the test data (Figures 2b, 3b and 4b).

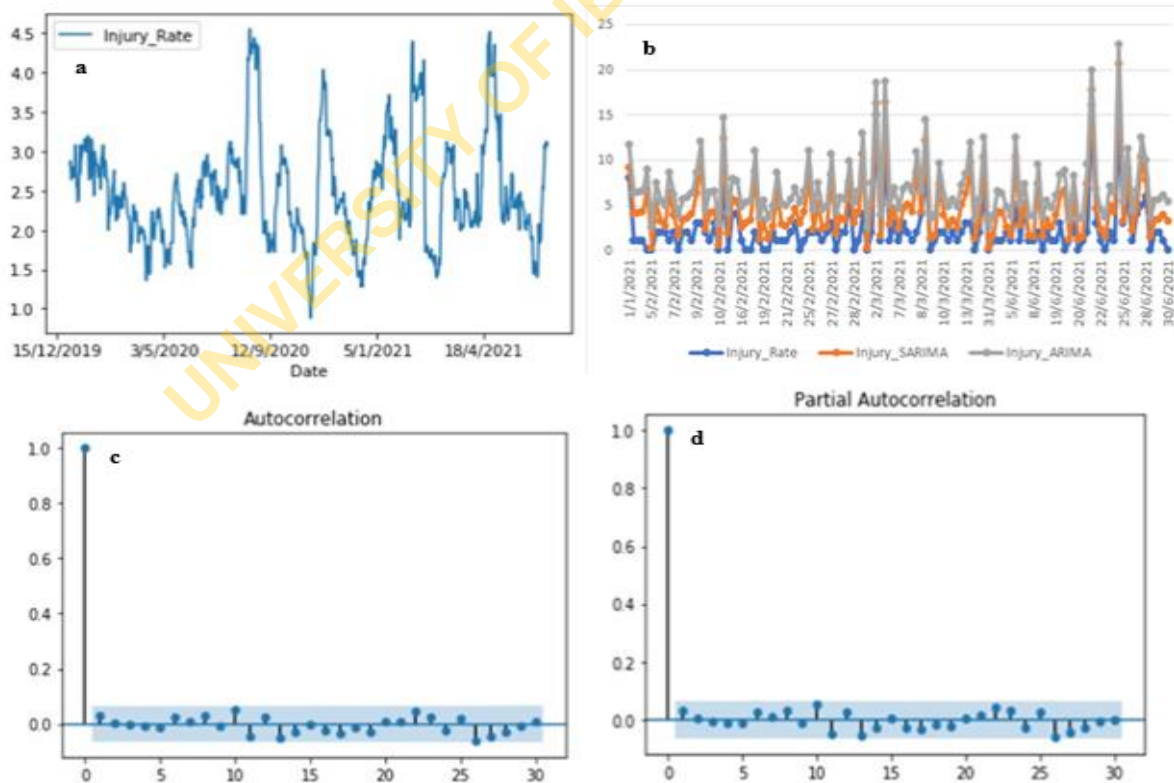


Figure 2: The trend of Accident injuries (a), Actual Injury trend and predictions using ARIMA and SARIMA (b), autocorrelation function (ACF) and the partial autocorrelation function (PACF) plots (c and d) show a non- seasonal pattern in data

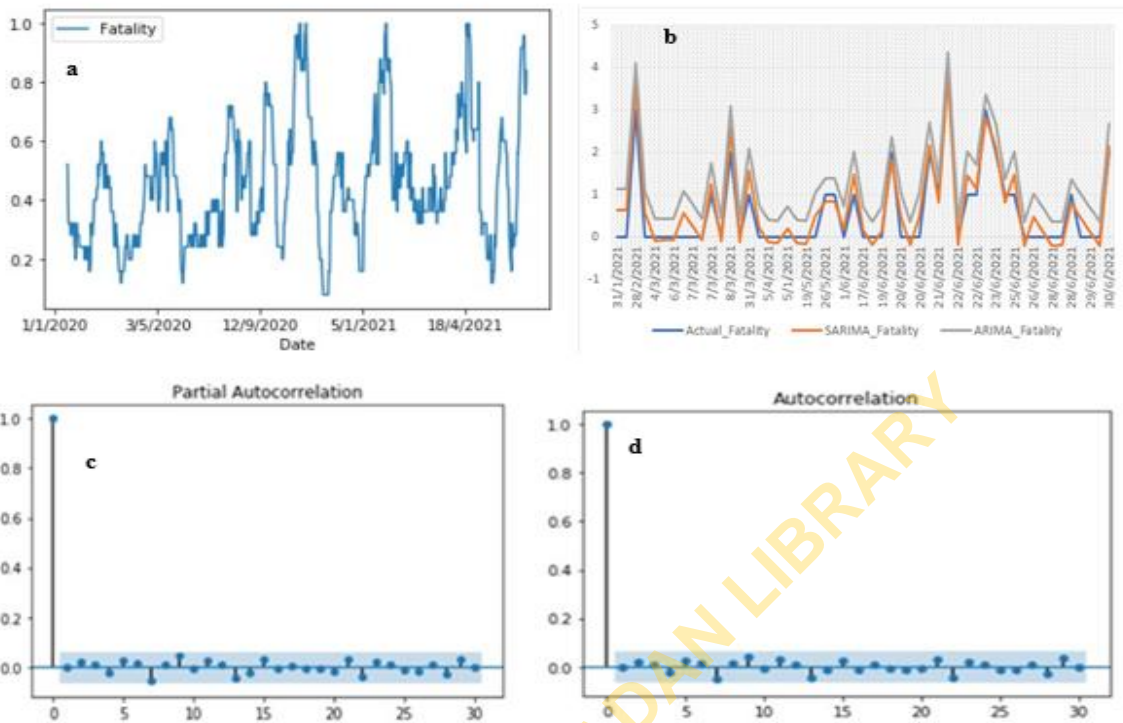


Figure 3: The trend of Accident fatality (a), Actual fatality trend and predictions using ARIMA and SARIMA (b), autocorrelation function (ACF) and the partial autocorrelation function (PACF) plots (c and d) show a non-seasonal pattern in data

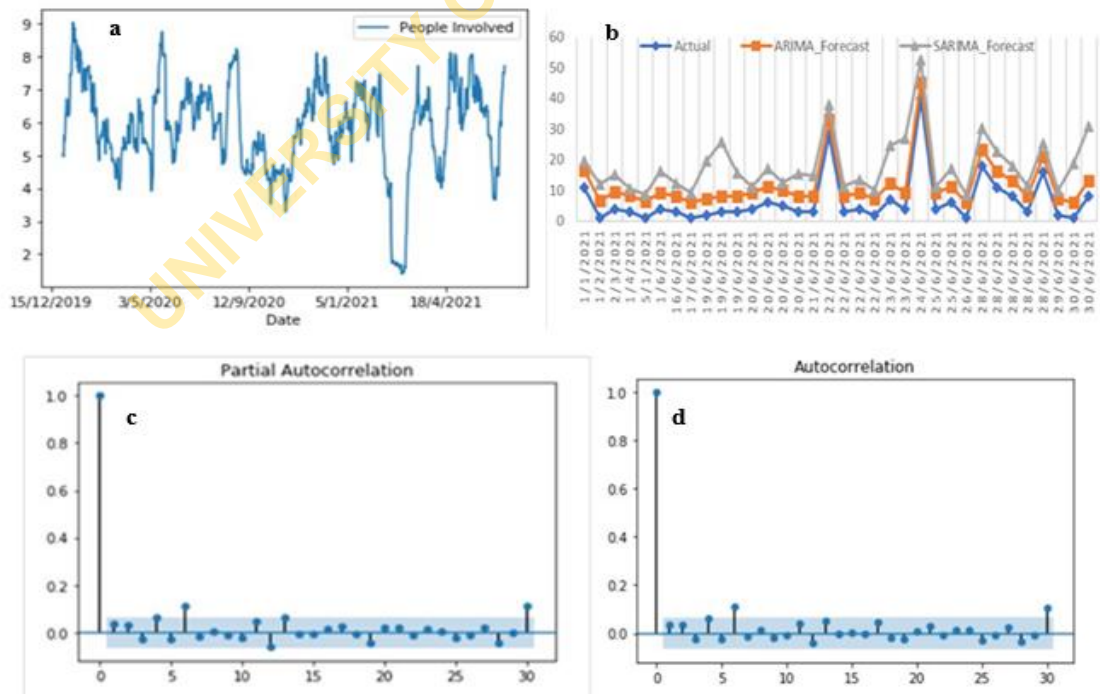


Figure 4: The trend of Accident vehicle occupants (a), Actual vehicle occupant trend and predictions using ARIMA and SARIMA (b), autocorrelation function (ACF) and the partial autocorrelation function (PACF) plots (a and d) show a seasonal pattern in data

Table 2: Parameter Estimation and Goodness of Fit Measures for the Prediction Models

Models	Parameters	Constant	Coefficient	SE	P-value
Accident-Injury	AR ₁	2.4326	-0.7298	0.119	0.000
	MA ₁		0.7650		
Accident-Fatality	AR ₁	0.0002	0.2012	0.032	0.000
	MA ₁		-1.1944		
Vehicle Occupants	AR ₂	5.9479	0.7603	0.337	0.024
	MA ₂		-0.7486		
	SAR ₁₂		0.1888		
	SMA ₁₂		-0.2446		

Note: AR_i and MA_i are the *i*th order of autocorrelation and partial autocorrelation, respectively. Also, SAR_i and SMA_i are the *i*th order of seasonal autocorrelation and seasonal partial autocorrelation, respectively.

3.2 Future Prediction

Using the built models, the predicted RTAs outcomes for four years (2022 to 2025) are shown in Table 3. The models forecast a 172% increase in fatality, a 38% increase in vehicle occupants and 32% increase injury attributed to road traffic accidents by the end of 2025 in the study area (Table 4). The predictions are in agreement with the WHO Global Burden of Disease project and the World Bank's Traffic Fatalities and Economic Growth (TFEC) project, using transport, population and economic data. The models from the two projects forecast that there will be increase in road traffic fatalities in low-income countries by 2024 (WHO, 2023; Twenefour et al, 2021). However, a study conducted in Tanzania showed that in 2030, the nation will be able to reduce RTA fatalities by 50% with a decreasing rate of 0.360 per year (Ndume *et al.*, 2022). Other studies on road traffic accidents and injuries in other nations also predicted an increase in road traffic accidents injuries and fatalities (Al-Zyood, 2017; Mohammadian *et al.*, 2014). These differences in outcomes may be attributed to geographic factors, available transportation systems in different nations, road safety policies and implementations.

Table 3: Future Prediction of Road Traffic Accident Outcomes (2022 – 2025)

Year	Accident Injuries	Accident Vehicle Occupants	Accident Fatality
2022	804	2020	234
2023	768	1963	272
2024	733	1891	310
2025	706	1828	347

Table 4: Road Traffic Accident Outcomes between 2021 and 2025

	End of 2021	End of 2025	Percentage Difference (%)
Accident Vehicle Occupants	5,577	7702	38
Accident Injury	2,290	3030	32
Accident Fatality	427	1162	172

3.3 Performance Evaluation

The performance of ARIMA and SARIMA models were compared for all the RTAs outcomes considered in this study. The SARIMA and ARIMA models had similar performance for the forecast of the RTAs outcomes, but ARIMA models had the least values for the root mean square errors (RMSE) and mean absolute errors (MAE) for all the variables considered except in vehicle occupant where SARIMA had better performance (Table 4). The SARIMA model performed better on vehicle occupants' data which support the fact that SARIMA models are more effective for time-series data with seasonal trends.

SARIMA and ARIMA models had close accuracy values in this study but ARIMA models performed better on injuries and fatalities data than SARIMA models.

This study supported the fact that machine learning techniques such as ARIMA and SARIMA models can be effective tools for forecasting future occurrences and patterns of road traffic accident injuries, fatalities and the number of vehicle occupants in the study area (Mora *et al.*, 2020). This technology can be developed for policy formulation on transportation and traffic management in the study area, to reduce the incidences of road traffic accidents to a very minimal level. However, this study showed that other than the increased need to travel during festive seasons, there is a need to investigate the effect of other factors such as the road design, poor state of the road, uncompleted road construction projects, non-use of protective devices in vehicles, presence of alcohol, excessive and inappropriate speed, mechanical failures on increasing rates of road traffic accidents injuries and fatality in the study area (Parvareh *et al.*, 2018; WHO, 2023).

Table 5: Evaluation Metrics for the Prediction Models

Model	RMSE	MAE
SARIMA_Vehicle Occupants	6.37	4.33
ARIMA_Vehicle Occupants	6.39	4.45
SARIMA_Injury	4.07	2.15
ARIMA_Injury	4.03	2.20
SARIMA_Fatality	1.25	0.67
ARIMA_Fatality	1.16	0.73

4. CONCLUSION

This paper investigated the use of time-series technique applying ARIMA and SARIMA models to predict accidents outcomes within and around Ibadan metropolis using data of daily recorded cases of road traffic accidents on seven specified routes between 2019 and 2021 collected from the FRSC. The study area has a high rate of RTAs between December 2019 and June 2021. SARIMA and ARIMA models had close accuracy values in this study but ARIMA models performed better on injuries and fatalities data than SARIMA models. This study supported the fact that machine learning techniques such as ARIMA and SARIMA models can be effective tools for forecasting future occurrences and patterns of RTAs. The findings showed that ARIMA and SARIMA models can be effectively developed to make acceptable decision on road safety issues in the study area. Given the predicted RTA outcomes in the study area, in particular the number of injuries and fatalities, there is a need for a road traffic accident prevention and evaluation programme by the FRSC and all agencies on promoting safety on the roads.

REFERENCES

- Adejugbagbe, A. M., Fatiregun, A. A., Rukewe, A. and Alonge T. (2015). Epidemiology of Road Traffic Crashes among Long Distance Drivers in Ibadan, Nigeria. *African Health Sciences*, 15(2): 481-488.
- Akinleye M. T and Tijani M. A. (2017). Assessment of Quality of Asphalt Concrete used in Road Construction in South West Nigeria. *Nigerian Journal of Technological Development*, 14(2): 51-54.
- Al-Zyood, M. (2017). Forecast Car Accident in Saudi Arabia with ARIMA Models. *International Journal of Soft Computing and Engineering*, 7(3): 30-33.
- Bozkurt, Ö. Ö., Biricik, G. and Tayşi, Z. C. (2017). Artificial Neural Network and SARIMA Based Models for Power Load Forecasting in the Turkish Electricity Market. *Plos One* 12(4). <https://doi.org/10.1371/journal.pone.0175915>.
- Çodur, M. Y. and Tortum, A. (2015). An Artificial Neural Network Model for Highway Accident Prediction: A Case Study of Erzurum, Turkey. *Promet-Traffic and Transportation*, 27(3): 15-25. <https://doi.org/10.7307/ptt.v27i3.1551>.

- Dipto, I. C., Rahman, M. A., Islam, T. and Rahman H. M. M., (2020). Prediction of Accident Severity Using Artificial Neural Network: A Comparison of Analytical Capabilities between Python and R. *Journal of Data Analysis and Information Processing*, 8(1): 134-157.
- Gatarić, D., Ruškić, N., Aleksić, B., Đurić, T., Pezo, L., Lončar, B. and Pezo, M. (2023). Predicting Road Traffic Accidents - Artificial Neural Network Approach. *Algorithms*, 16(5): 257. <https://doi.org/10.3390/a16050257>.
- Getahun K. A. (2021). Time Series Modeling of Road Traffic Accidents in Amhara Region. *Journal of Big Data* 8:102. <https://doi.org/10.1186/s40537-021-00493-z>.
- He, S. Sadeghi, M. A., Chawla, S., Alizadeh, M., Balakrishnan, H. and Madden, S. (2021). Inferring High-Resolution Traffic Accident Risk Maps Based on Satellite Imagery and GPS Trajectories. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 11977-1198. <https://doi.org/10.4236/jdaip.2020.83008>.
- Mohammadian, M., Hajare, A. and Hafshejani, A. M. (2014). Incidence Trends of Injury and Mortality from Traffic Accidents in Urban and Suburban Areas of Isfahan Province during 2002 - 2010. *Journal of Police Medicine*, 3(1): 1-10.
- Mora, N., Sood H. and Goyal, T. (2020). Application of Machine Learning Technique for Prediction of Road Accidents in Haryana - A Novel Approach. *Journal of Intelligent and Fuzzy Systems*. 38(1): 6627 – 6636. doi:10.3233/JIFS-179742
- Moslehi, S., Gholami, A., Haghdoost, Z., Abed, H., Mohammadpour, S. and Moslehi, M. A. (2021). Prediction of Traffic Accidents Based on Weather Conditions in Gilan Province Using Artificial Neural Network. *Journal of Health Administration*, 24 (3): 67-78.
- National Bureau of statistics (NBS) (2023). Road Transport Data Q2 2023 (Accessed: 23 September 2023). <https://nigerianstat.gov.ng/elibrary/read/1241395>
- National Bureau of Statistics (NBS) (2024). Road Transport Data: Q4 2023. Retrieved February 17, 2024, from <https://nigerianstat.gov.ng/elibrary/read/1241462>.
- Ndume, V. A., Rutalebwa, E. C. and Runyoro, A. K. (2022). Prediction of Road Accidents Trend in Tanzania Using ARIMA Model: The Road Safety Implication by 2021-2030. *International Journal of Traffic and Transportation Engineering*, 11(1): 1-7. doi: 10.5923/j.ijtte.20221101.01.
- Ogbodo, D. and Nduoma, E. (2011). FRSC. Nigerian Roads Are Second-Worst in the World. *Journal of Sustainable Development*, 7(4): 163-164.
- Onyemaechi, N., and Ofoma, U. R. (2016). The Public Health Threat of Road Traffic Accidents in Nigeria: A Call to Action. *Ann Med Health Sci Res*, 6:199-204.
- Parvareh, M., Karimi, A., Rezaei, S., Woldemichael, A., Nili, S., Nouri, B. and Nasab, N. E. (2018). Assessment and Prediction of Road Accident Injuries Trend Using Time-Series Models in Kurdistan, *Burns and Trauma*, 6(1): 1-5. <https://doi.org/10.1186/s41038-018-0111-6>.
- Peden M. and Sminkey L. (2004). World Health Organization dedicates World Health Day to Road Safety. *Injury Prevention*, 10: 67. doi: 10.1136/ip.2004.005405.
- Salami, R. O., Von Meding, J. K. and Giggins, H. (2016). Assessing Habitats of Vulnerability in African Cities: A Case of Poverty Housing in Ibadan Metropolis, Nigeria. doi:[10.13140/RG.2.1.3185.3683](https://doi.org/10.13140/RG.2.1.3185.3683).
- Santos, D., Saias, J., Quaresma, P. and Nogueira, V. B. (2021). Machine Learning Approaches to Traffic Accident Analysis and Prediction. *Computers*, 10(1): 157-165. <https://doi.org/10.3390/computers10120157>
- Shaik, M. E., Islam, M. M. and Hossain, Q. S. (2021). A Review on Neural Network Techniques for the Prediction of Road Traffic Accident Severity. *Asian Transport Studies*. 7: 100040. <https://doi.org/10.1016/j.eastsj.2021.100040>.
- Twenefour, F. B. K., Ayitey, E., Kangah, J. and Brew, L. (2021). Time Series Analysis of Road Traffic Accidents in Ghana. *Asian Journal of Probability and Statistics*, 11(2): 12-20.
- World Health Organization (WHO) (2023). Global Status Report on Road Safety 2023. Geneva. ISBN 978-92-4-008651 (electronic version)
- World Health Organization (WHO) (2023). Road Traffic Injuries. <https://www.who.int/news-room/factsheets/detail/road-traffic-injuries>.