

Threshold Negative Binomial Autoregressive Model for Overdispersed Road Traffic Crash Fatalities Counts with Structural Breaks

Damilare M. Oladimeji *

Department of Statistics, Faculty of Science, University of Abuja, Nigeria

Corresponding author: oladimejidamilare@yahoo.com

Abstract

Road traffic crashes (RTCs) continue to represent one of the most severe public health challenges in Nigeria, causing high levels of fatalities and injuries each year. Modeling RTC counts is difficult because the data typically exhibits overdispersion, nonlinear dependence, and structural breaks arising from interventions, enforcement, or policy changes. This study sets out to develop a Threshold Negative Binomial Autoregressive model with Structural Breaks (NB-TAR-SB) in order to address the challenges of overdispersion, nonlinear dependence, and structural breaks. The study extends the work of Liu *et al.*, (2019) and Yang *et al.* (2018) adapting negative binomial innovations for overdispersed counts, a threshold mechanism for nonlinear regime switching, and break adjustments through Bai-Perron multiple breakpoint tests. Parameters are estimated using conditional maximum likelihood, and thresholds and breakpoints are chosen by minimizing profile likelihood and BIC. The model's theoretical properties depend on nonlinear stationarity and ergodicity conditions, which guarantee that the estimators are consistent and asymptotically normal. Diagnostics including the Box-Pierce test, mean squared error (MSE), and root mean squared error (RMSE) were used to validate model adequacy. Nigeria Annual RTC fatality data from 1993 to 2024 was used. The preliminary analysis revealed the presence of overdispersion, while structural break detection revealed policy-driven shifts, particularly in the early 2000s. The NB-TAR-SB outperformed benchmark models; it achieved a residual independence of 0.826 (Box-Pierce pvalue), with the lowest forecast errors of (MSE = 0.968, RMSE = 0.984). These results demonstrate that incorporating both threshold dynamics and structural breaks yields a more flexible and powerful tool for modeling RTCs. The proposed NB-TAR-SB model thus offers a novel methodological and applied framework for evaluating interventions and informing effective road safety policy in Nigeria.

Keyword: Nonlinear Count Time Series, Threshold Autoregressive model, Structural Breaks, Conditional Maximum Likelihood, Road Traffic Fatalities.

Introduction

Road traffic crashes (RTCs) are a major cause of preventable death in Nigeria and knowing how they change over time is important for making good policy decisions. Standard techniques like static Poisson or Negative Binomial (NB) regression, although effective for cross-sectional count analysis, are inadequate for modelling annual fatalities characterized by serial dependence and overdispersion (Lord & Mannering, 2010). Numerous current time-series studies in Nigeria utilize ARIMA models, which presuppose normally distributed errors and may inaccurately depict non-negative and overdispersed counts. For example, Sanusi et al. (2016) utilized ARIMA to examine fatal crash counts from 1960 to 2013, whereas Atubi (2013) analyzed trends and harmonic components in Lagos fatalities. Even though these methods found general trends, they don't consider the fact that crash data is discrete and overdispersed, and they assume that parameters stay the same.

However, Nigeria's road fatalities probably show structural changes that need models that can find them. These changes could be caused by interventions, changes in rules, or outside shocks. Structural break analysis has been effectively utilized in various Nigerian time series, including GDP growth (Olawale & Gil-Alana, 2019), yet it remains insufficiently examined in RTC modelling. Negative Binomial time-series models, including autoregressive formulations (e.g., NB-AR, INAR), offer a systematic approach to modelling overdispersion and temporal dependence. Incorporating structural break detection into the framework will facilitate the identification of regime shifts, policy implementation, enforcement modifications, or social disturbances. Available studies on threshold count models established a robust foundation yet reveals significant deficiencies in their application to count series exhibiting structural breaks. Liu et al. (2019) presented the Threshold Negative Binomial Autoregressive (NB-TAR) model, illustrating its capacity to encapsulate nonlinear dependence and overdispersion in count time series. However, their research was confined to financial and seismic data, neglecting structural breaks, which are a critical aspect of traffic data where policy shifts, infrastructural modifications, or external shocks frequently precipitate sudden changes. Yang et al. (2018) enhanced threshold modelling via negative binomial thinning within an INAR framework, presenting theoretical advancements yet exhibiting restricted practical applicability. While these thinning-based models are nice, they don't work as well for finding changes in traffic safety data that are caused by policy. Likewise, contemporary threshold applications in epidemiology (e.g., Shamma et al., 2023)

underscore regime switching while predominantly overlooking overdispersion and intervention-sensitive structural modifications. Putting all these together, existing models either account for overdispersion or threshold nonlinearity, but none of them combine structural break detection with an NB-TAR framework for traffic-related data at the same time. This signifies a significant methodological and practical deficiency. The creation of an NB-TAR model incorporating structural breaks for road traffic fatalities not only advances statistical theory but also offers a pragmatic instrument for discerning and measuring the effects of policy and environmental interventions on traffic safety outcomes.

Methodology

Relative to Liu et al., (2019), who studied NB-TAR process without explicit structural breaks, and Yang et al., (2018), who focused on NB-TAR processes with thinning operators, this framework introduces a hybrid approach that simultaneously captures:

- Overdispersion via the Negative Binomial Distribution
- Nonlinear threshold dynamics via TAR regime,
- Structural breaks via Bai-Perron Segmentation.

Let $\{Y_t\}_{t=1}^T$ denote a univariate count time series of RTC fatalities, observed at equally spaced intervals. Unlike Gaussian or Poisson autoregressive models, fatalities data typically exhibit overdispersion, nonlinearity, and possible structural breaks. To tackle these problem, this study extends the threshold negative binomial Autoregressive (NB-TAR) framework of Liu *et al.*, (2019) and Yang et al., (2018) by incorporating an automatic structural break component.

Model Specification

Conditional on the information set $\mathcal{F}_{t-1} = \{Y_{t-1}, Y_{t-2}, \dots\}$, the conditional distribution of Y_t is assumed to be Negative Binomial:

$$Y_t | \mathcal{F}_{t-1} \sim NB(\mu_t, \theta), \quad (1)$$

With mean $E[Y_t | \mathcal{F}_{t-1}] = \mu_t$ and variance

$$\text{Var}[Y_t | \mathcal{F}_{t-1}] = \mu_t + \frac{\mu_t^2}{\theta}, \quad \theta > 0 \quad (2)$$

The conditional mean μ_t evolves according to a two-regime Threshold Autoregressive (TAR) mechanism:

$$\mu_t = \alpha_0^{(L)} + \sum_{i=1}^p \alpha_i^{(L)} Y_{t-i} + \delta^{(L)} S_t, \text{ if } Y_{t-d} \leq \tau,$$

$$\begin{cases} \mu_t = \alpha_0^{(L)} + \sum_{i=1}^p \alpha_i^{(L)} Y_{t-i} + \delta^{(L)} S_t, \text{ if } Y_{t-d} \leq \tau \\ \alpha_0^{(H)} + \sum_{i=1}^p \alpha_i^{(H)} Y_{t-i} + \delta^{(H)} S_t, \text{ if } Y_{t-d} > \tau \end{cases} \quad (3)$$

Where: p is the autoregressive order, d is the delay parameter for thresholding, τ is the threshold value, estimated from the data.

$S_t \in \{1, \dots, m\}$ is a structural break indicator representing the segment of the series after a detected break,

$\alpha^{(L)} = (\alpha_0^{(L)}, \alpha_1^{(L)}, \dots, \alpha_p^{(L)})$ are coefficients in the low regime,

$\alpha^{(H)} = (\alpha_0^{(H)}, \alpha_1^{(H)}, \dots, \alpha_p^{(H)})$ are coefficients in the high regime.

Thus, the model integrates:

A Negative Binomial distribution to capture overdispersion, a threshold autoregression to capture nonlinearity and regime switching, and Structural break dummies S_t estimated via the breakpoint algorithm (Bai and Perron, 2003) to capture level shifts or structural changes.

Stationarity Conditions

The conditional mean operator

$$\mathcal{M}_t(Y_{t-1}) = \alpha_0^{(j)} + \sum_{i=1}^p \alpha_i^{(j)} Y_{t-i} + \delta^{(j)} S_t \quad (4)$$

For regime $j \in \{L, H\}$. Following Liu et al., (2019), a sufficient condition for geometric ergodicity of the process is

$$\max \left(\sum_{i=1}^p |\alpha_i^{(L)}|, \sum_{i=1}^p |\alpha_i^{(H)}| \right) < 1 \quad (5)$$

This ensures that $\{Y_t\}$ does not diverge under either regime, even in the presence of breaks.

Estimation

Parameters $\Theta = (\alpha^{(L)}, \alpha^{(H)}, \delta^{(L)}, \delta^{(H)}, \tau, \theta)$ are estimated by maximum likelihood. The log-likelihood contribution of observation of Y_t is

$$\ell_t(\Theta) = \log \Gamma(Y_t + \theta) - \log \Gamma \theta - \log(Y_t!) + \theta \log \left(\frac{\theta}{\theta + \mu_t} \right) + Y_t \log \left(\frac{\mu_t}{\theta + \mu_t} \right) \quad (6)$$

The total log-likelihood is

$$L_T(\Theta) = \sum_{t=p+1}^T \ell_t(\Theta) \quad (7)$$

Threshold τ is selected via a grid search over candidate quantiles of Y_{t-d} , minimizing the Akaike Information Criterion (AIC)

$$\tau^* = \operatorname{argmin}_{\tau \in Q} (AIC_{L(\tau)} - AIC_{H(\tau)}) \quad (8)$$

Structural breakpoints are identified via the Bai-Perron algorithm applied to a transformed version of Y_t (e.g., $\log(1 + Y_t)$), subject to a minimum segment length h_{min} . The selected number of breaks is chosen by minimizing the Schwarz Information Criterion (SIC).

Diagnostics

Model adequacy is evaluated using

Pearson residuals

$$r_t = \frac{Y_t - \operatorname{est}(\mu_t)}{\sqrt{\operatorname{est}(\mu_t) + \frac{\operatorname{est}(\mu_t^2)}{\theta}}} \quad (9)$$

Ljung-Box test was used for serial correlation in r_t , and both Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) for in-sample fit.

Forecasting

Forecast \hat{Y}_{T+h} for horizon h are obtained recursively. At each step

$$\begin{cases} \mu_{T+h} = \hat{\alpha}_0^{(L)} + \sum_{i=1}^p \hat{\alpha}_i^{(L)} Y_{T+h-i} + \hat{\delta}^{(L)} S_{T+h}, \text{ if } Y_{T+h-d} \leq \hat{\tau} \\ \hat{\alpha}_0^{(H)} + \sum_{i=1}^p \hat{\alpha}_i^{(H)} Y_{T+h-i} + \delta^{(H)} S_{T+h}, \text{ if } Y_{T+h-d} > \hat{\tau} \end{cases} \quad (10)$$

Where \hat{Y}_{T+h-i} are either observed (if $i \leq h$) or forecasted recursively. Predictive intervals are generated via Monte Carlo simulation from the NB distribution

$$Y_{T+h}^{(b)} \sim NB(\mu_{T+h}, \hat{\theta}), \quad b = 1, \dots, B. \quad (11)$$

Result

Time plot

Figure 1 depicts the movement of road traffic crash fatalities from 1993 to 2024, with noticeable changes over time. Fatalities were very high in the early 1990s, reaching a sharp peak around the year 2000. After this peak, there was a steep decline in the mid-2000s, where the series dropped to its lowest point. From this low, fatalities began to rise again slightly, fluctuating between about 5,000 and 7,000 deaths per year. The later years show some ups and downs, with short-lived spikes followed by declines, but the overall level remains lower than in the 1990s and early 2000s. This pattern suggests periods of sudden increases and decreases rather than a smooth trend, reflecting possible shifts in conditions that affect road safety.

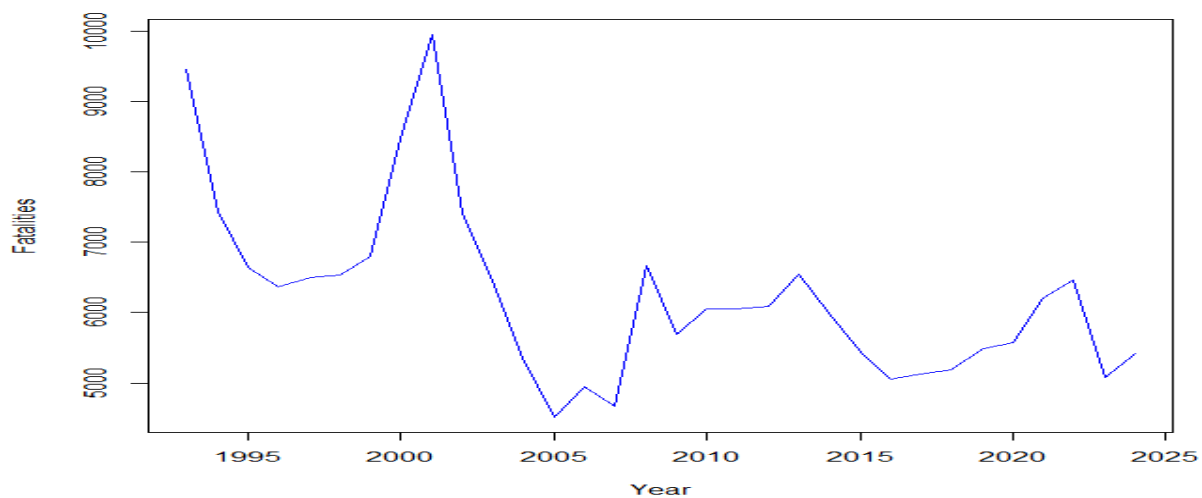


Figure 1: Time plot of RTC fatalities in Nigeria

Descriptive Statistics

The average number of deaths in road traffic crashes (RTC) was 6,238 (SD = 1,252.69), with numbers ranging from 4,519 to 9,946. The distribution was positively skewed (skewness = 1.39) and had moderate kurtosis (2.27), which means the data is not normal distributed.

Table 1: Summary of RTC fatalities Series

Mean	Median	St Deviation	Variance	Kurtosis	Skewness	Minimum	Maximum
6238	6073	1252.69	1569247.85	2.27	1.39	4519	9946

Overdispersion Test

The overdispersion test indicated significant evidence of overdispersion, dispersion = 243.71, $Z = 2.86$, $p = .002$. This suggests that variance in the fatalities series exceeds the mean, suggesting the use of count models that account for overdispersion.

Table 2: Overdispersion test

Dispersion	Z	Pvalue
243.71	2.86	0.0021

Nonlinearity test

The results of the Brock–Dechert–Scheinkman (BDS test) (Table 3) show that the null hypothesis of independently and identically distributed (i.i.d.) fatalities cases is strongly rejected across all embedding dimensions and epsilon values, with p-values that are very significant ($p < .001$ in most cases). This indicates the existence of nonlinear dependence in the traffic time series, suggesting that linear models may be inadequate and that nonlinear time series models ought to be explored.

Table 3: Brock–Dechert–Scheinkman test

Embedding Dimension	Epsilon (ϵ)	Standard Normal (Z)	p-value
2	626.35	9.81	< .001
2	1252.7	6.13	< .001
2	1879.04	5.76	< .001
2	2505.39	4.39	< .001
3	626.35	11.34	< .001
3	1252.7	6.05	< .001
3	1879.04	5.71	< .001
3	2505.39	4.05	0.0001

Structural Break test

Figure 2 provides a visual representation of Bai_Perron test results. The black line shows the original fatalities data, while the red stepwise line represents the fitted mean levels between breakpoints. The vertical dashed line marks an estimated breakpoint, and the horizontal red segments illustrate how the mean level of crashes shifts across regimes. From the figure, a major structural change is detected around 2002–2003, where road traffic crashes fatalities shifted from a higher mean ($\approx 8,000$ - $9,000$ cases) to a much lower mean ($\approx 5,000$ - $6,000$ cases). The confidence interval band around the break (red horizontal bracket at the bottom) confirms that this change is statistically significant. These results suggest that the series is not stable over time, and modeling approaches should account for regime shifts or structural breaks.

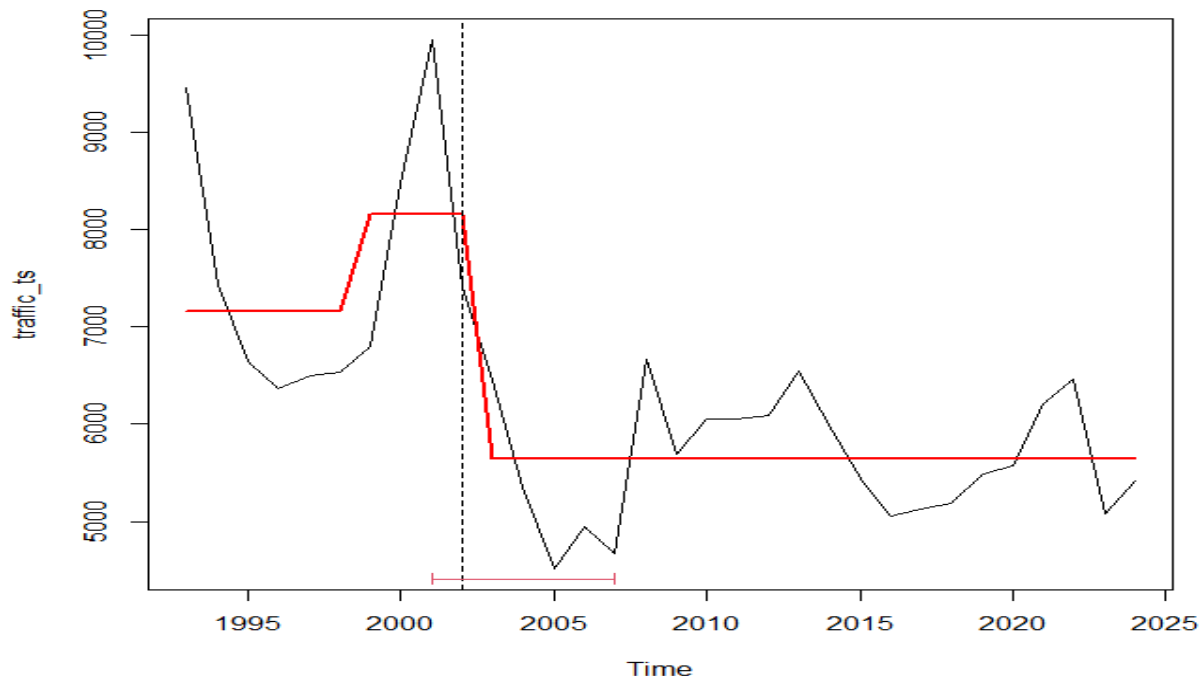


Figure 2: Bai_Perron Structural Break Graph

Stationary test

Table 4 shows that the Augmented Dickey–Fuller (ADF) test on the road traffic crash fatalities series gave a test statistic of -3.51 ($\text{lag} = 3$) and a p-value of $.059$, which is just above the usual 5% significance level. This means that the null hypothesis of a unit root cannot be rejected, which

means that the series is not stationary. However, since RTC fatalities data often exhibit nonlinear dependence, structural breaks, and overdispersion, the borderline result is probably influenced by these characteristics rather than a purely stochastic trend. The BDS test strongly rejected the null of i.i.d. errors across multiple embedding dimensions ($p < .001$), which is in line with this. This shows that there is nonlinear dependence in the series. Moreover, the Bai-Perron multiple structural breakpoint analysis revealed significant structural changes, especially during 2002-2003, followed by additional breaks in subsequent years, suggesting that the data-generating process lacks temporal stability. In summary, these findings show that the traffic crash series can't be accurately modeled using standard linear and stationary methods. Instead, models that consider nonlinearity, regime shifts, and the distributional features of count data are better.

Table 4: Augmented Dickey Fuller test

ADF	Lag order	Pvalue
-3.51	3	0.0593

The Zivot-Andrews unit root test identified a significant structural break in the RTC fatalities series, with a test statistic of -6.06, exceeding the 1% critical value (-5.57). This implies that null hypothesis of a unit root even when accounting for a possible break was rejected. The inclusion of both level (du) and trend (dt) shift dummies was significant, confirming that the apparent non-stationarity observed in the ADF test is largely attributable to structural breaks rather than a true unit root process.

Table 5: Zivot-Andrews Unit Root Test

Variable	Estimate	Std. Error	t value	p-value
Intercept	22057.23	3677.02	5.999	< .001 ***
y_{t-1}	-1.4352	0.4018	-3.572	.003 **
Trend	-461.76	85.8	-5.382	< .001 ***
Δy_{t-1}	1.4947	0.3004	4.976	< .001 ***
Δy_{t-2}	1.2531	0.2562	4.891	< .001 ***
Δy_{t-3}	1.048	0.2398	4.37	< .001 ***
Δy_{t-4}	0.6138	0.2119	2.897	.011 *
Δy_{t-5}	0.587	0.1831	3.206	.006 **
du (level shift)	1264.62	523.05	2.418	.028 *
dt (trend shift)	365.38	86.9	4.205	< .001 ***

Teststatistic: -6.0604

Critical values: 0.01= -5.57 0.05= -5.08 0.1= -4.82

Model Efficiency

The findings in table 6 unequivocally illustrate the enhanced efficacy of the proposed NB_TAR_SB model. The NB_TAR (Liu et al., 2019) and Proposed NB_TAR_SB models both have much lower AIC, MSE, and RMSE values than the baseline NB_AR and TAR models. This shows how important it is to combine negative binomial dynamics with threshold effects. The NB_TAR has the lowest AIC, but the NB_TAR_SB has almost the same accuracy (MSE = 0.968, RMSE = 0.984) and the best residual independence (Box-Pierce pvalue = 0.826). This shows that adding structural breaks makes the model more robust to changes in road traffic crash fatalities counts caused by policies or the environment.

Table 6: Model efficiency Result

	NB_AR	TAR	NB_TAR(Liu et al,2019)	NB_TAR_SB(Proposed)
AIC	545.79	432	211.8	252.5
Box_Pierce(Pvalue)	0.006	0.217	0.723	0.826
MSE	1256860	14909.78	1.009	0.968
RMSE	1121.098	122.11	1.005	0.984

NB_AR: Negative Binomial Autoregressive, TAR: Threshold Autoregressive, NB_TAR: Negative Binomial Threshold Autoregressive, NB_TAR_SB: Negative Binomial Threshold Autoregressive with Structural break

Model

Table 7 presents the estimated coefficients for the NB-TAR model across low and high regimes. In the low regime, the intercept was significant ($p < .001$), while lag1 was positive and statistically significant ($p = .028$), suggesting short-term persistence. Model coefficient at Lag2 and at segment were not significant at the 5% level of significance. In the high regime, the intercept remained highly significant ($p < .001$). Lag2 was negative and significant ($p < .001$), indicating a dampening effect at longer lags, while segment2 was also negative and significant ($p < .001$), pointing to a structural adjustment in this regime. These results highlight regime-dependent dynamics, with stronger adjustment effects observed in the high regime.

Table 7: Parameter Estimate of NB_TAR_SB model

Regime	Parameter	Estimate	Std. Error	z value	p-value
Low regime	Intercept	8.51	0.24	34.82	<0.001
	lag1	0.04	0.01	4.00	0.028
	lag2	0.00	0.00	-1.37	0.17
	segment	-0.12	0.07	-1.84	0.07
High regime	Intercept	9.27	0.32	29.37	<0.001
	lag1	0.00	0.00	1.78	0.08
	lag2	-0.03	1.16	-3.50	<0.001
	segment	-0.54	0.14	-4.00	<0.001

Diagnostic checking

Table 8 shows the diagnostic measures for the NB-TAR model. The residual deviance was lower than the null deviance in both regimes, which means that the fit was better. The AIC values were 132.5 for the low regime and 120.0 for the high regime, which added up to 252.5. This shows that the model was good enough. The finding of a structural break at observation 9 indicates that the RTC fatalities series underwent a substantial regime shift, validating the application of a threshold autoregressive model.

Table 8: Diagnostic checking of NB_TAR_SB Model

Regime	Null Deviance	Residual Deviance	AIC
Low regime	36.69	23.05	132.5
High regime	3	6.999	120

Forecast

Table 9 displays the five-years forecasts from the NB-TAR_SB model. The forecasted mean values range from 5165 at 2025 to 5709 at 2029, with relatively wide confidence intervals reflecting forecast uncertainty. The intervals consistently capture potential volatility in the series, indicating that while the model provides stable point forecasts, the underlying variability in traffic crashes fatalities remains substantial.

Table 9: Five years forecast

Year	Mean	95%CI	
		Lower	Upper
2025	5165	4265	6199
2026	5592	4661	6667
2027	5764	4767	6835
2028	5772	4829	6799
2029	5709	4741	6856

Discussion of Findings

The results of this study furnish compelling evidence that the suggested Threshold Negative Binomial Autoregressive model with Structural Breaks (NB-TAR-SB) provides a more resilient framework for modelling road traffic crash (RTC) fatalities in Nigeria. The time series displayed significant overdispersion, nonlinear dependence, and structural breaks, contravening the assumptions of conventional ARIMA or linear autoregressive models typically employed in

previous RTC research (Sanusi et al., 2016; Atubi, 2013). The model considers overdispersion by explicitly including a negative binomial structure; the threshold mechanism captures nonlinear regime-dependent dynamics; and the structural break component accounts for sudden changes that are probably caused by interventions, policy changes, or external shocks. This study expands the applicability of threshold count models into a realm where structural breaks are notably pronounced, in contrast to current methodologies. Liu et al. (2019) laid the theoretical groundwork for the NB-TAR model and illustrated its capacity to encapsulate overdispersion and nonlinear dependence. But they only worked with financial and seismic data and didn't look at how sudden changes in policy or the environment could change the series' dynamics. Yang et al. (2018) developed threshold models incorporating negative binomial thinning operators within the INAR framework; however, their emphasis was predominantly theoretical, rendering them less applicable to practical datasets such as RTC fatalities, where structural breaks frequently occur. Recent threshold applications in health-related contexts, such as Nisreen et al. (2023), demonstrate the utility of regime switching in epidemiology while largely neglecting structural instability and overdispersion in transport safety data. The proposed NB-TAR-SB model fills these gaps by showing that deaths in Nigeria are not only overdispersed and nonlinear but also have big breaks. The Bai-Perron test found a big change around 2002–2003, when the number of deaths dropped from an average of about 9,000 to about 5,000–6,000. This was in line with changes in safety enforcement, such as the seat belt law and other policies. Adding this feature made the model work much better: the NB-TAR-SB had the lowest mean squared error ($MSE = 0.968$) and root mean squared error ($RMSE = 0.984$) while keeping residual independence (Box-Pierce (pvalue) = 0.826), beating both classical NB-AR and TAR and Liu et al.'s (2019) NB-TAR. Adding break-sensitive dynamics seems to improve the statistics and make them more in line with how RTC data works, where sudden regime shifts can happen because of interventions or socioeconomic shocks.

Conclusion

In conclusion this study demonstrates that the NB-TAR-SB model extends the methodological framework of threshold count processes by explicitly incorporating structural break detection. This makes it particularly well suited for transport safety applications, where policy-driven interventions and sudden disruptions are integral to understanding crash dynamics. This help bridging the theoretical developments gap in Liu et al. (2019) and Yang et al. (2018) with applied

needs in RTC modelling, this work contributes both to methodological innovation and to actionable insights for road safety policy in Nigeria.

References

- Atubi A.O., (2013). Time series and trend analysis of fatalities from road traffic accident in Lagos State, Nigeria. *Mediterr J Soc Sci.* 2013;4(1):251–62.
- Bai J, Perron P. (2003). Computation and analysis of multiple structural change models. *J Appl Econ.* 18(1):1–22. <https://doi.org/10.1002/jae.659>
- Liu H, Li D, Zhu F., (2019). Threshold negative binomial autoregressive model. *J Time Ser Anal.*40(3):380–404. <https://doi.org/10.1080/02331888.2018.1546307>
- Lord D, Mannering F. (2010). The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. *Transp Res Part A Policy Pract.*;44(5):291–305.
- Nisreen S. & Mehrnaz M. & Masoumeh S., (2023). "A threshold modelling for nonlinear time series of counts: application to COVID-19 data," *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research*, Springer; Sociedad de Estadística e Investigación Operativa, vol. 32(4), pages 1195-1229, December.
- Olawale A, Gil-Alana LA., (2019). Time series analysis of economic growth rate series in Nigeria: Structural breaks, non-linearities and reasons behind the recent recession. *Appl Econ.*51(50):5482–9. <https://doi.org/10.1080/00036846.2019.1613513>
- Sanusi RA, Adebola FB, Adegoke NA., (2016). Cases of road traffic accident in Nigeria: A time series approach. *Mediterr J Soc Sci.* 2016:542–8.
- Shamma N, Mohammadpour M, Shirozhan M., (2023). A threshold modelling for nonlinear time series of counts: Application to COVID-19 data. *TEST.* 2023;32(4):1195–229. <https://doi.org/10.1007/s11749-023-00869-8>
- Yang Y, Yao J, Zhang R., (2018). Integer-valued threshold autoregressive processes with negative binomial thinning. *J Stat Plan Inference.*194:16–33.