FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 1, January, 2026, pp 231 – 242
DOI: https://doi.org/10.33003/fjs-2026-1001-4325

THE PERFORMANCE OF INTEGER-VALUED AUTO-REGRESSIVE (INAR) MODEL IN ZERO INFLATED DATA

***[1]Ibrahim Modu, [1]Hamidu Aliyu Chamalwa, [1]Harun Rann Bakari and [2]Alhaji Umar Abubakar**

[1]Department of Statistics, Faculty of Physical Sciences, University of Maiduguri, Borno, Nigeria.
[2]Department of Mathematics and Computer Sciences, Faculty of Sciences, Kashim Ibrahim University, Maiduguri, Nigeria.

*Corresponding authors' email: alajistat@gmail.com

ABSTRACT

Time series count data frequently exhibits zero inflation and even heavy-tailedness in practical applications. Many models have been proposed for modelling count data, but heavy-tailedness is less considered. The effect of excess zeros on time series count data cannot be disregarded. Thus, there is a need for a model that would cater for excess zeros in the time series data. The proposed model, a new integer-valued autoregressive process, is expected to capable of capturing these features. This study therefore investigates the effectiveness of Integer-Valued Autoregressive (INAR) models in handling time series count data at different proportions of excess zeros, determine the predictive ability of INAR models at different steps ahead and compare its performance with orders of model {INAR (1), INAR (2), INAR (3) and INAR (4)} being used for the data. The effects of sample sizes n $= 20, 40, \dots, 200,$ on the performance of the models were also studied through simulation. At every sample size, the best status of the orders p, where p = 1, 2, 3, 4 are respectively determined for 20%, 30% and 40% proportions of the excess zeros using information criteria AIC, BIC and HQIC. Forecast accuracy was assessed using the Thiel's U statistic, where lower values indicate better performance. INAR (3) achieved the lowest AIC, BIC and HQIC values across most scenarios indicating a strong model fit and is suggested for use in fitting any time series of count to the underlying features given in this dissertation. Similarly, INAR (3) has the best predicting capacity because of its lower value at some point in the steps ahead. However, the Theil values of INAR (1), INAR (2), and INAR (4) are improving better at larger steps ahead as the percentage of excess zero increases.

Keywords: Data, Fitting, Order, Valued, Zero

INTRODUCTION

Count time series data—non-negative integer observations such as daily disease cases or financial transactions—present unique statistical challenges that render traditional continuous-data models (e.g., ARIMA) unsuitable. The application of such model often leads to inefficient or biased parameter estimates, as count data frequently exhibit overdispersion and, critically, excess zeros (zero inflation). While specialized models like the Zero-Inflated Poisson (ZIP) have been developed to handle zero inflation, they are inadequate for overdispersed data (Lambert, 1992; Ndwiga et al., 2019). This leaves a significant methodological gap in accurately modeling real-world count series, such as those seen in epidemiology or finance, where both zero inflation and overdispersion coexist (Saleh et al., 2021; Tawiah, 2021). A promising yet underexplored avenue for such data is the class of Integer-Valued Autoregressive (INAR) models, explicitly designed for discrete counts. However, the existing literature lacks a clear, comparative evaluation of these models' performance under varying and realistic conditions of zero inflation. Consequently, policymakers and analysts relying on forecasts (e.g., for disease outbreaks or transaction volumes) may be using suboptimal models.

Therefore, this study is motivated by the need for a systematic, simulation-based comparative evaluation of INAR models. The aim is to identify the most effective INAR specification for forecasting count time series data across different empirically observed levels of zero inflation, thereby providing a robust statistical tool for applications where accurate discrete-count forecasting is essential.

MATERIALS AND METHODS

A Simulation-Based Evaluation of INAR Models for Zero-Inflated Count Data

This study employs a rigorous Monte Carlo simulation framework in R to systematically evaluate the fitting and forecasting performance of Integer-Valued Autoregressive (INAR) models under controlled conditions of zero inflation. The methodology is designed for full transparency and reproducibility, with each component detailed below.

Data Generation Process

The core data-generating process (DGP) simulates count time series that combine genuine INAR autocorrelation structure with artificially induced zero inflation.

Base INAR Process

We simulate the fundamental count series $\{X_t\}$ using a Poisson INAR (2) model as the baseline DGP. The model is defined as:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \epsilon_t \qquad (1)$$

where:

i. ∘∘ denotes the binomial thinning operator ($\alpha \circ Y = \sum_{i=1}^{x} B_i$, with $B_i \sim$Bernoulli(α)).
ii. $\epsilon_t \sim$i.i.d. Poisson (λ) represents the innovation term.
iii. The autoregressive parameters are fixed at $\alpha_1 = 0.6$ and $\alpha_1 = -0.3$ to ensure a stationary process with moderate persistence.
iv. The Poisson innovation rate is fixed at λ=3.0, determining the marginal mean of the non-inflated process.

Induction of Zero Inflation:
To create the final observed series $\{X_t\}$with excess zeros, we apply a deterministic random replacement algorithm to the base series $\{X_t\}$. For a target zero-inflation proportion π_{ZI} (e.g., $20\%, 30\%, 40\%$), a corresponding percentage of randomly selected observations in $\{X_t\}$ are replaced with zero. This procedure directly manipulates the empirical probability mass at zero, mimicking real-world scenarios where a latent process (e.g., non-occurrence of an event) generates structural zeros beyond those expected from the Poisson-INAR model.

Simulation Design and Model Specification
Experimental Factors:
The simulation varies two key factors in a full-factorial design:

i. Sample Size (n): The sample size in simulation studies is generally determined using the formula

$$n = \frac{Z^2\sigma^2}{E^2} \quad (2)$$

Where (Z) is the standard normal deviate at the desired confidence level, σ^2 is the variance of the process, and (E) is the allowable margin of error. In this study, rather than fixing a single value of (n), we systematically varied the sample size across (n = 20, 40, . . ., 200) to evaluate the robustness of INAR models under both small and large sample conditions.

ii. Zero-Inflation Proportion (πZIπZI): $\pi_{ZI}\{0\%, 20\%, 30\%, 40\%\}$ For each unique (n, π_{ZI}) combination, N=1000 independent time series are generated.

Candidate Models:
The study fit INAR models of orders p=1,2,3,4 to each simulated series to assess order selection performance. The general INAR(p) model is:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \epsilon_t, \quad \epsilon_t \sim poisson(\lambda) \quad (3)$$

$$Y_t = \alpha_1 \circ Y_{t-1} + \alpha_2 \circ Y_{t-2} + \cdots + \alpha_p \circ Y_{t-p} + \epsilon_t, \ \epsilon_t \sim \text{Poisson}(\lambda)$$

The upper limit of p≤4 is justified by parsimony and common practice in applied count time series analysis, where higher-order dependencies are often captured by lower-order models or are not empirically prevalent in datasets of the sizes considered.

Model Estimation and Selection
Estimation Procedure
All INAR models are estimated via Conditional Maximum Likelihood (CML), as implemented in the tscount package (version x.y.z) in R. The CML method conditions on the first pp observations and maximizes the likelihood of the remaining observations, providing consistent and efficient estimators for INAR processes.

Model Selection Criteria
For each fitted model, we compute three standard information criteria to evaluate in-sample fit and penalize overfitting:

i. Akaike Information Criterion (AIC): AIC= $-2log(L) + 2k$
ii. Bayesian Information Criterion (BIC): BIC= $-2log(L) + klog(n)$
iii. Hannan–Quinn Criterion (HQIC): HQIC= $-2log(L) + 2klog(log(n))$ where L is the maximized likelihood value and k is the number of estimated parameters (i.e., $k = p + 1$ for pp thinning parameters and one innovation rate λ). The model with the smallest criterion value is considered optimal for a given series.

Performance Evaluation
Performance is assessed along two dimensions:
Order Selection Accuracy: The percentage of simulations where the true data-generating order ($p = 2$) is correctly identified by each information criterion across different n and πZI levels.

Forecasting Ability**:** For each selected model, we generate one-step-ahead forecasts. Forecasting accuracy is measured using the Mean Absolute Error (MAE) and the Mean Squared Error (MSE) on a hold-out sample, evaluated against the known DG. Their respective formulars are stated below:

Mean Absolute Error (MAE)
The Mean Absolute Error (MAE) measures the average magnitude of the forecast errors, without considering their direction. It is calculated as:

$$MAE = \frac{1}{n}\sum |X_t - \hat{X}_t| \quad (4)$$

Where: n = number of forecasted points, y_t = actual observed value at time t and $\hat{y}_t$ = forecasted value at time t.

Mean Squared Error (MSE)
The Mean Squared Error (MSE) measures the average of the squares of the forecast errors. It penalizes larger errors more than MAE. It is calculated as:

$$MSE = \frac{1}{n}\sum (X_t - \hat{X}_t)^2 \quad (5)$$

Where: n = number of forecasted points, y_t = actual observed value at time t and $\hat{y}_t$ = forecasted value at time t

RESULTS AND DISCUSSION
The performance of INAR models were determined through simulations on the count data with excess zeros. The effect of sample sizes $n = 20, 40, \ldots, 200$, on the performance of the models were studied. At every sample size, the best status of the p, where p = 1, 2, 3, 4 are respectively determined for the levels of the excess zero in the data generated using criteria like AIC, BIC and HQIC as presented in table 1 and plotted on graphs 1,2 and 3. 20% of excess zero were injected in the data so as to determine the best INAR model for each category. The simulation study was carried out with 1000 iteration on each case in R statistical software. For each iteration, the values of the criteria for the assessment (AIC, BIC and HQIC) were computed and their average values were recorded according to sample sizes as shown in tables below. The values from the tables were plot with their figures respectively. The model with lowest criteria is considered as the best.
Samples of data generated across the sample sizes of 20 and 200 with different proportions/ levels of excess zeros (20%, 30% and 40%) are presented in figures 1,2 and 3 respectively as follows

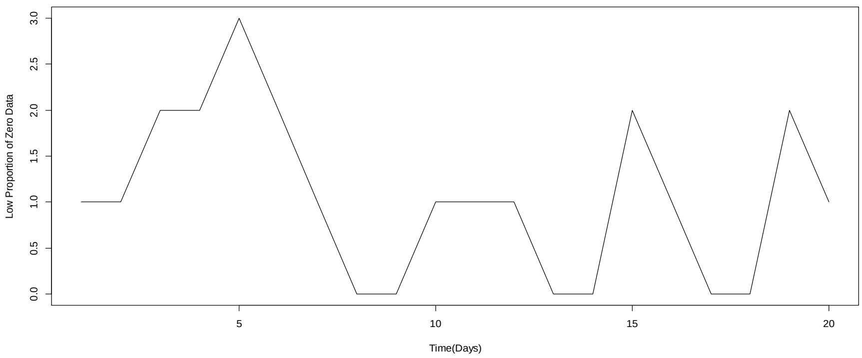

Figure 1: Plots of Sample Generated Time Series Count Data

The figure above shows the data simulated with Poisson distribution, 20% of excess zeros were injected in the algorithms and the sample of the simulated data were displayed as shown in the graph above. The zero values were clearly seen randomly on the graph

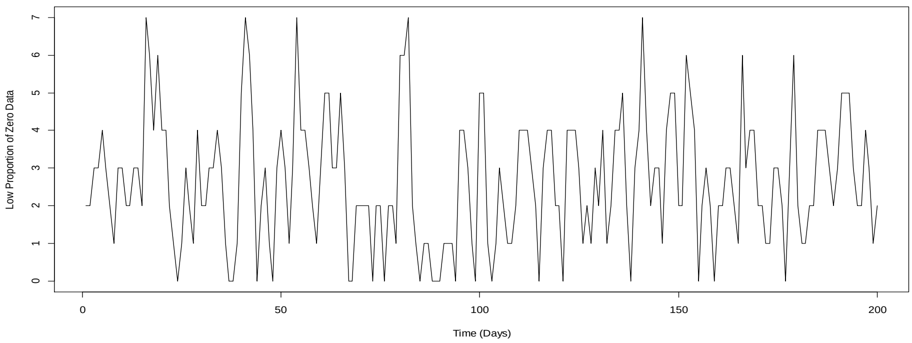

Figure 2: Simulated Count Data with 20% Excess Zero and Large Sample Size

Figure 2 shows the data simulated with Poisson distribution, 20% of excess zeros and sample size of 200 were injected in the algorithms and the sample of the simulated data were displayed as shown in the graph above. The zero values were clearly seen randomly on the graph

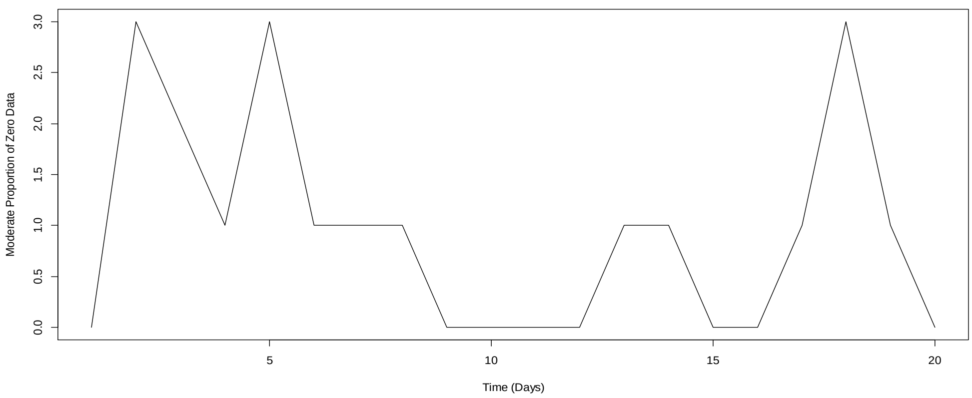

Figure 3: Simulated Count Data with 30% Excess Zeros and Small Sample Size

The figure above shows the data simulated with Poisson distribution, 30% of excess zeros and sample size of 20 were injected in the algorithms and the sample of the simulated data were displayed as shown the graph above. The zero values were clearly seen randomly on the graph

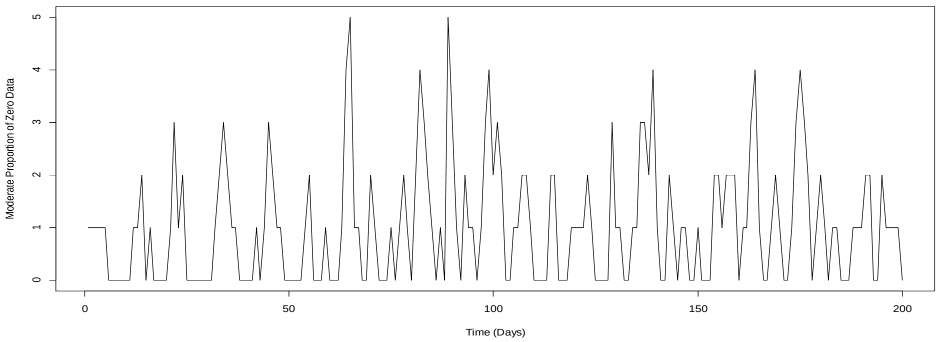

Figure 4: Simulated Count Data with 30% Excess Zeros and Large Sample Size

Figure 4 shows the data simulated with Poisson distribution, 30% of excess zeros and sample size of 200 were injected in the algorithms and the sample of the simulated data were displayed as shown the graph above. The zero values were clearly seen randomly on the graph

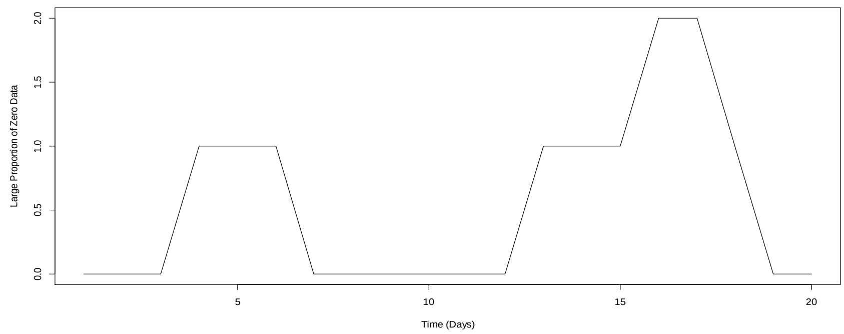

Figure 5: Simulated Count Data with 40% Excess Zeros and Small Sample Size

The figure 5 above shows the data simulated with Poisson distribution, 40% of excess zeros and sample size of 20 were injected in the algorithms and the sample of the simulated data were displayed as shown the graph above. The zero values were clearly seen randomly on the graph

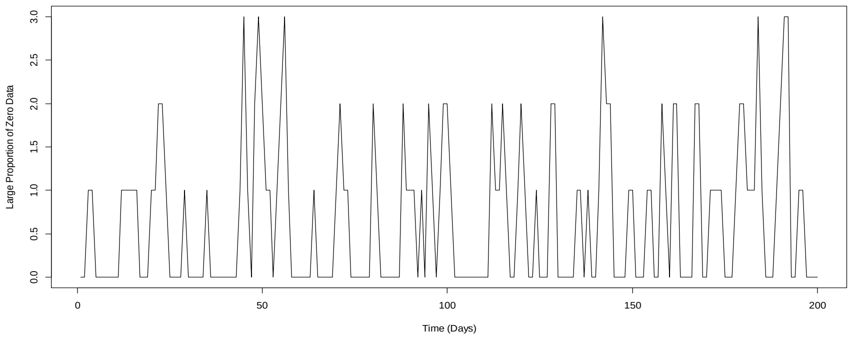

Figure 6: Simulated Count Data with 40% Excess Zeros and Large Sample Size

The figure 6 above show the data simulated with Poisson distribution, 40% of excess zeros and sample size of 200 were injected in the algorithms and the sample of the simulated data were displayed as shown the graph above. The zero values were clearly seen randomly on the graph

Fitting INAR model to Count Data with No Zero

The performance of INAR models were determined through simulations on the count data with excess zeros. The effect of sample sizes $n = 20, 40, \ldots, 200$, on the performance of the models were studied. At every sample size, the best status of the p, where p = 1, 2, 3, 4 are respectively determined for the levels of the excess zero in the data generated using criteria like AIC, BIC and HQIC as presented in table 1 and plotted on graphs in figures 7, 8 and 9 The simulation study was carried out with 1000 iteration on each case in R statistical software. For each iteration, the values of the criteria for the assessment (AIC, BIC and HQIC) were computed and their average values were recorded according to sample sizes as shown in table 1. The values from the tables were plot in figures 7, 8 and 9. The model with lowest criteria is considered as the best.

Table 1: Comparative Analysis of INAR (p) Model on Data without Zeros

N	AIC				BIC				HQIC			
	INAR (1)	INAR (2)	INAR (3)	INAR (4)	INAR (1)	INAR (2)	INAR (3)	INAR (4)	INAR (1)	INAR (2)	INAR (3)	INAR (4)
20	1.009	1.070	1.034	1.042	1.014	1.218	1.075	1.045	1.106	1.116	1.125	1.135
40	1.000	1.091	1.044	1.054	1.015	1.247	1.091	1.076	1.136	1.149	1.163	1.176
60	0.990	1.111	1.053	1.066	1.016	1.276	1.108	1.106	1.166	1.183	1.200	1.217
80	0.980	1.132	1.062	1.078	1.017	1.305	1.124	1.136	1.196	1.217	1.237	1.258
100	0.970	1.152	1.071	1.090	1.018	1.333	1.141	1.166	1.226	1.250	1.274	1.299
120	0.960	1.173	1.081	1.102	1.019	1.363	1.157	1.196	1.256	1.284	1.312	1.340
140	0.950	1.193	1.090	1.113	1.019	1.391	1.174	1.226	1.286	1.318	1.349	1.380
160	0.940	1.214	1.099	1.125	1.020	1.420	1.190	1.256	1.316	1.351	1.386	1.421
180	0.931	1.234	1.108	1.137	1.021	1.449	1.206	1.286	1.346	1.385	1.424	1.462
200	0.921	1.255	1.117	1.149	1.022	1.477	1.223	1.316	1.376	1.444	1.461	1.503

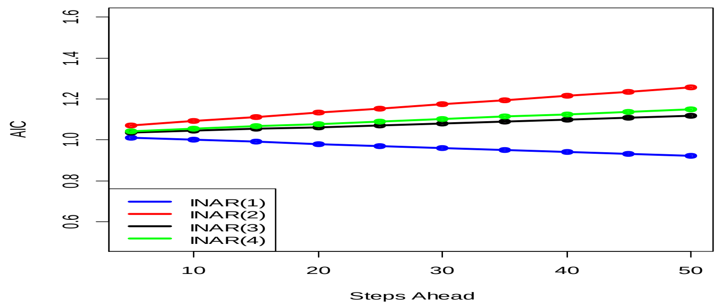

Figure 7 Plot of AIC of INAR (p) on Data without Zeros

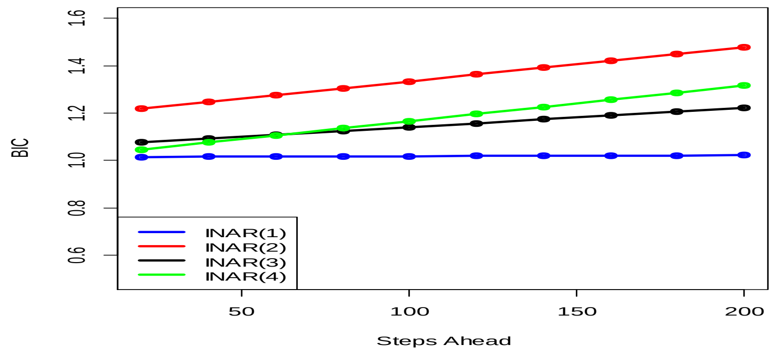

Figure 8: Plot of BIC of INAR (p) on Data without Zeros

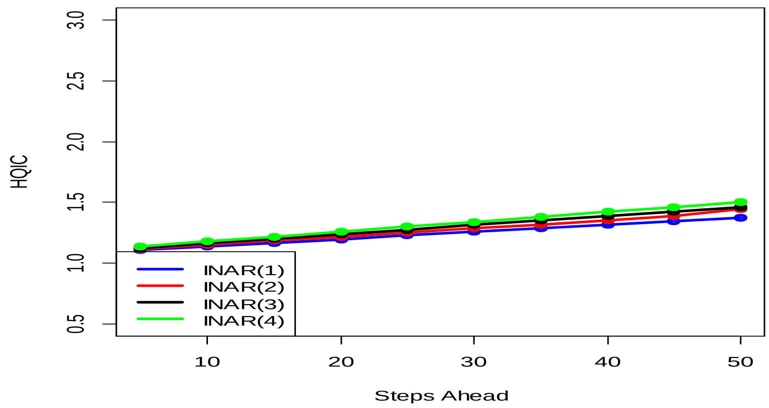

Figure 9: Plot of HQIC of INAR (p) on Data without Zeros

The Table above shows the relative performance of the INAR models under different criteria of selection at different sample size when the data is generated from a Poisson distribution without a zero response. From the table, it is observed that INAR (1) is the best at different sample sizes especially on the basis of AIC, BIC and HQIC criteria. This is followed by INAR (3) as the second best in terms of all criteria for the data without zeros. This best model selected is in line with theory of parsimony.

Fitting INAR model to Count Data with 20% of Excess Zero

The performance of INAR models were determined through simulations on the count data with 20% of excess zero. The effect of sample sizes $n = 20, 40, \ldots, 200,$ on the performance of the models were studied. At every sample size, the best status of the p, where p = 1, 2, 3, 4 are respectively determined for the levels of excess zero in the data generated using criteria like AIC, BIC and HQIC as presented in table 2 and plotted on graphs 10, 11 and 12. 20% of excess zero were injected in the data so as to determine the best INAR model for each category. The simulation study was carried out with 1000 iteration on each case in R statistical software. For each iteration, the values of the criteria for the assessment (AIC, BIC and HQIC) were computed and their average values were recorded according to sample sizes as shown in table 2. The values from the tables were plot in figures 10, 11 and 12. The model with lowest criteria is considered as the best.

Table 2: Comparative Analysis of INAR (P) Model On Data with 20% Of Excess Zeros

N	AIC				BIC				HQIC			
	INAR (1)	**INAR (2)**	**INAR (3)**	**INAR (4)**	**INAR (1)**	**INAR (2)**	**INAR (3)**	**INAR (4)**	**INAR (1)**	**INAR (2)**	**INAR (3)**	**INAR (4)**
20	1.484	0.990	0.680	1.160	1.434	1.550	1.044	1.420	1.430	1.400	1.368	1.464
40	1.440	0.987	0.705	1.153	1.392	1.517	1.038	1.388	1.401	1.371	1.340	1.432
60	1.400	0.979	0.722	1.136	1.351	1.478	1.032	1.356	1.371	1.342	1.313	1.400
80	1.359	0.972	0.739	1.119	1.309	1.438	1.026	1.324	1.340	1.313	1.285	1.368
100	1.318	0.964	0.756	1.103	1.268	1.398	1.020	1.293	1.310	1.284	1.258	1.336
120	1.277	0.957	0.774	1.086	1.226	1.359	1.014	1.261	1.279	1.255	1.230	1.304
140	1.236	0.949	0.789	1.070	1.185	1.319	1.008	1.229	1.249	1.226	1.203	1.271
160	1.195	0.942	0.806	1.053	1.143	1.279	1.002	1.197	1.219	1.1973	1.175	1.240
180	1.154	0.934	0.823	1.037	1.102	1.240	0.996	1.165	1.188	1.168	1.148	1.208
200	1.481	0.994	0.688	1.169	1.434	1.557	1.044	1.420	1.432	1.400	1.368	1.464

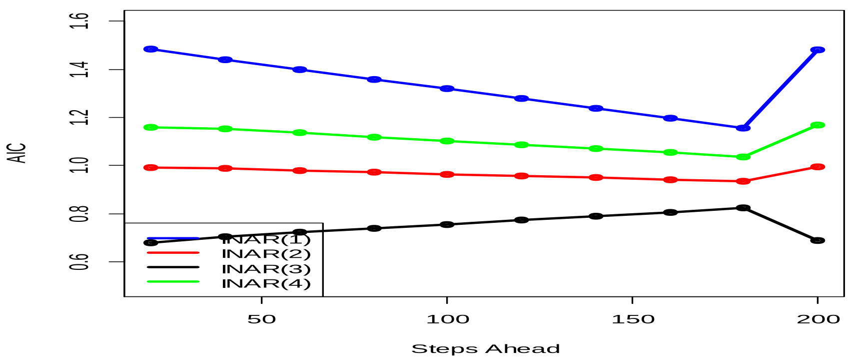

Figure 10: Plot of AIC of INAR (p) on Data with 20% of Zeros

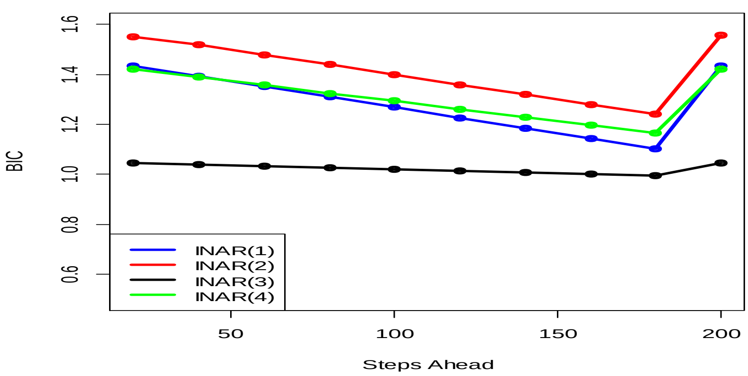

Figure 11: Plot of BIC of INAR (p) on Data with 20% of Zeros

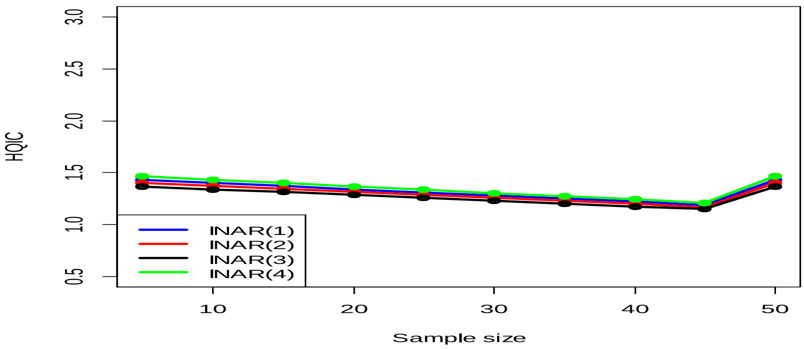

Figure 12: Plot of HQIC of INAR (p) on Data with 20% of Excess Zeros

Table 2 displays the average values of AIC, BIC, and HQIC of the fitted models calculated from 1000 iterations of data simulated from a Poisson distribution with 20% excess zero. The criteria values from Table 2 were plotted on Figures 13,14 and 15 respectively. It is noted that the model follows a similar pattern of fit across sample sizes. However, INAR (3) outperforms the other models at moderate sample sizes especially on the basis of two criteria, AIC and BIC across sample sizes suggesting its flexibility in moderate zero-inflation contexts. This indicate that the various sample sizes considered INAR (3) model shows resilience.

Fitting INAR Model to Count Data with 30% of Excess Zero

The performance of INAR models were determined through simulations on the count data with 40% of excess zeros. The effects of sample sizes $n = 20, 40, \ldots, 200,$ on the performance of the models were studied. At every sample size, the best status of the p, where p = 1, 2, 3, 4 are respectively determined for the levels of excess zero in the data generated using criteria like AIC, BIC and HQIC as presented in table 3 and plotted on graphs 16, 17 and 18. 30% of excess zero was injected respectively in the data so as to determine the best INAR model for each category. The simulation study was carried out with 1000 iteration on each case in R statistical software. The values of the criteria for the assessment (AIC, BIC and HQIC) were computed for iteration performed and their average values were recorded according to sample sizes as shown in table 3. The values from the tables were plot in figures 16, 17 and 18. The model with lowest criteria is considered as the best.

Table 3: Comparative Analysis of INAR (p) Model on Data with 30% of Excess Zeros

N	AIC				BIC				HQIC			
	INAR (1)	**INAR (2)**	**INAR (3)**	**INAR (4)**	**INAR (1)**	**INAR (2)**	**INAR (3)**	**INAR (4)**	**INAR (1)**	**INAR (2)**	**INAR (3)**	**INAR (4)**
20	9.287	8.046	6.165	8.975	7.372	2.940	1.525	6.693	12.740	8.943	8.138	6.332
40	15.150	28.420	11.700	21.840	18.790	2.787	2.461	25.130	9.890	8.583	7.273	5.963
60	14.460	25.930	10.770	20.250	17.300	2.834	2.367	22.930	9.021	7.779	6.537	5.295
80	14.170	23.940	9.840	18.670	15.800	2.881	2.274	20.730	8.148	6.975	5.802	4.628
100	13.280	20.940	8.920	17.080	14.310	2.974	2.180	18.520	7.276	6.171	5.066	3.961
120	12.390	18.450	7.991	15.500	12.820	2.964	2.086	16.320	6.404	5.367	4.330	3.293
140	12.100	15.960	7.063	13.920	11.330	3.021	1.993	14.120	5.532	4.563	3.594	2.626
160	11.110	13.470	6.135	12.330	9.840	3.068	1.895	11.910	4.659	3.759	2.859	1.997
180	10.510	10.980	5.207	10.750	8.3470	3.115	1.805	9.620	3.787	2.955	2.123	1.291
200	11.280	13.570	5.716	12.380	9.720	3.118	1.940	12.100	4.696	3.805	2.914	2.023

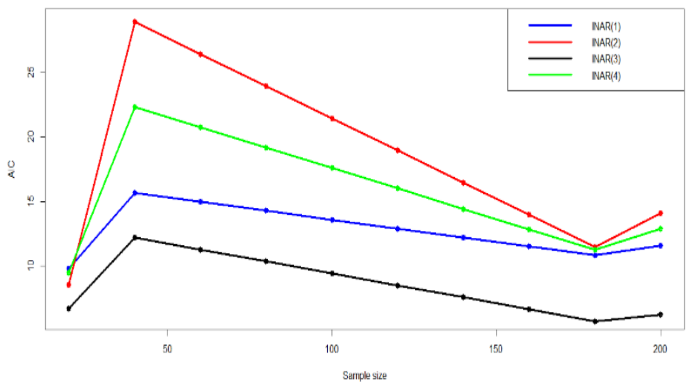

Figure 16: Plot of AIC of INAR (p) on Data with 30% of Zeros

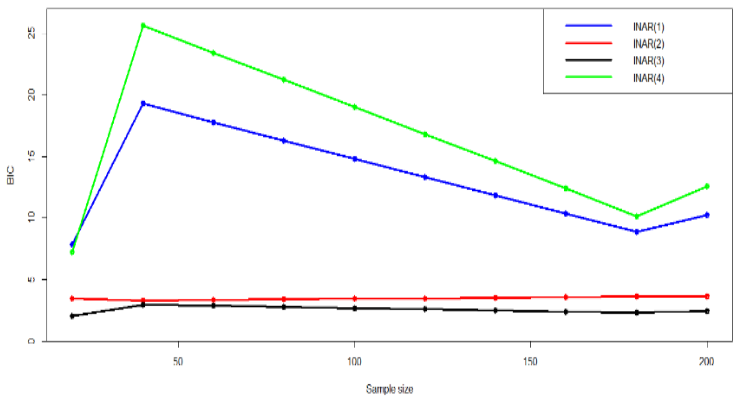

Figure 17: Plot of BIC of INAR (p) on Data with 30% of Excess Zeros

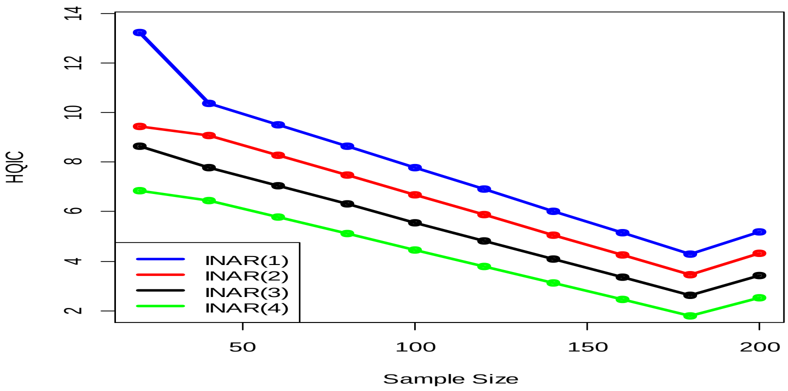

Figure 18: Plot of HQIC of INAR (p) on Data with 30% of Zeros

Figures 16, 17 and 18 shows the plots of the criteria values from Table 3. INAR (3) model is more robust to larger levels of excess zero than other competing models, particularly due to the lowest AIC, BIC and HQIC values and therefore appear to show the best fit for count time series data with 30% excess zero for small and large sample sizes, respectively.

Fitting INAR model to Count Data with 40% of excess zero

The performance of INAR models were determined through simulations on the count data with 40% of excess zeros. The effect of sample sizes $n = 20, 40, \ldots, 200,$ on the performance of the models were studied. At every sample size, the best status of the p, where p = 1, 2, 3, 4 are respectively determined for the levels of excess zero in the data generated using criteria like AIC, BIC and HQIC as presented in table 4 and plotted on graphs 19, 20 and 21. 40% of excess zero were injected respectively in the data so as to determine the best INAR model for each category. The simulation study was carried out with 1000 iteration on each case in R statistical software. For each iteration, the values of the criteria for the assessment (AIC, BIC and HQIC) were computed and their average values were recorded according to sample sizes as shown in table 3. The values from the tables were plot in figures 19, 20 and 21. The model with lowest criteria is considered as the best.

Table 4: Comparative Analysis of INAR (p) Model on Data with 40% of Excess Zeros

n	AIC				BIC				HQIC			
	INAR (1)	**INAR (2)**	**INAR (3)**	**INAR (4)**	**INAR (1)**	**INAR (2)**	**INAR (3)**	**INAR (4)**	**INAR (1)**	**INAR (2)**	**INAR (3)**	**INAR (4)**
20	10.820	12.600	5.219	11.750	9.090	13.060	1.914	11.220	12.360	12.090	2.575	10.230
40	10.570	11.630	4.723	11.120	8.459	12.640	1.888	10.400	11.470	11.400	2.001	9.570
60	10.310	10.660	4.226	10.510	7.827	12.680	1.863	9.570	10.590	10.710	1.573	9.910
80	10.060	9.690	3.730	9.890	7.196	12.610	1.838	8.747	9.710	10.020	2.148	9.740
100	9.810	8.731	3.233	9.274	6.565	12.520	1.813	7.917	10.170	9.660	1.723	9.401
120	9.550	7.764	2.737	8.654	5.933	10.310	1.788	7.086	10.050	9.350	2.298	9.061
140	9.303	6.797	2.240	8.033	5.302	10.020	1.763	6.256	8.935	9.047	1.872	8.720
160	9.049	5.830	1.744	7.413	4.671	9.550	1.737	5.425	8.818	8.729	1.547	8.389
180	8.796	4.862	1.247	6.792	4.039	9.610	1.712	4.595	8.701	8.419	1.522	8.039
200	8.542	3.895	0.751	6.172	3.408	9.520	1.687	3.764	8.584	8.108	1.507	7.699

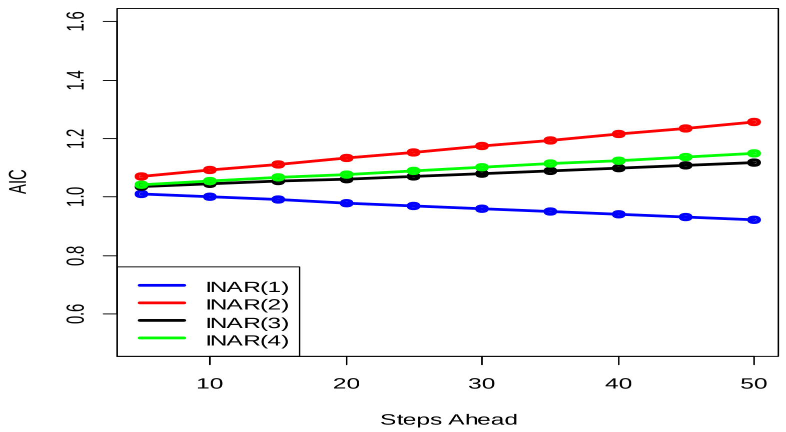

Figure 19: Plot of AIC of INAR (p) on Data without Zeros

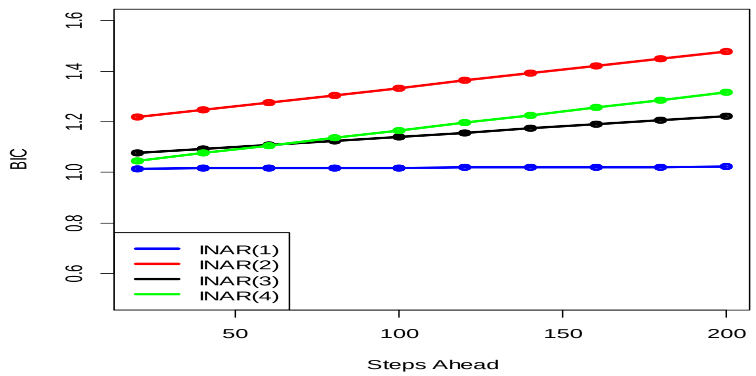

Figure 20: Plot of BIC of INAR (p) on Data without Zeros

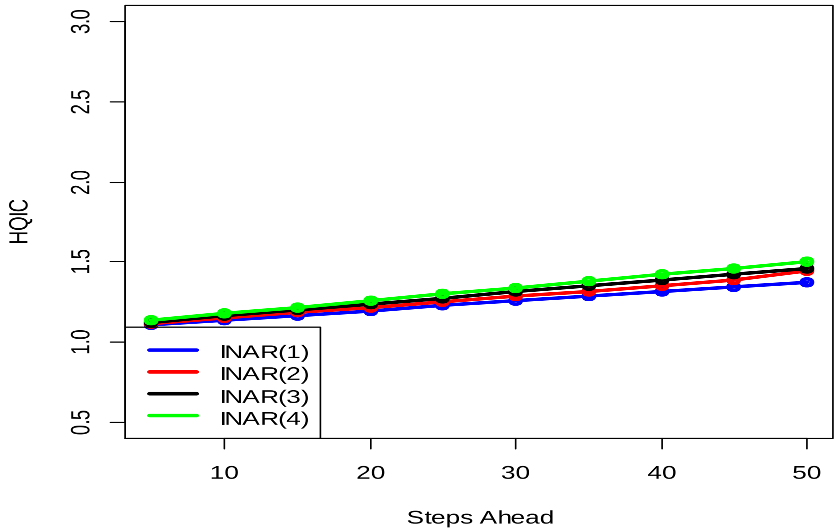

Figure 21: Plot of HQIC of INAR (p) on Data without Zeros

The average values of AIC, BIC and HQIC recorded in table 4 revealed that INAR (3) is the model best fit, because it has the minimum values of the three criteria used for the assessment and therefore chosen as the most robust model to data with high proportions of excess zeros.

Forecast Ability of the Models
The predictive ability of the models selected from INAR (1), INAR (2), INAR (3) and INAR (4) were examined using Theil U statistics. Theil U statistics is the relative accuracy measure that compares forecasted results with the results of forecasting with minimal historical data. It also requires the deviations to give more weight to large errors and to exaggerate errors, which can help eliminate methods with large errors. Theil's U smaller values indicates a better forecasting technique whereas higher values of Theil U indicate that the forecasting technique is worse than naive model.

Table 5: Forecast Performance of the Models

	20% of excess zero				30% of excess zero				40% of excess zero			
Steps Ahead	INAR (1)	INAR (2)	INAR (3)	INAR (4)	INAR (1)	INAR (2)	INAR (3)	INAR (4)	INAR (1)	INAR (2)	INAR (3)	INAR (4)
5	2.024	2.498	1.943	2.724	2.044	2.264	1.330	2.997	1.904	2.365	0.432	2.999
10	2.004	2.399	1.923	2.706	2.024	2.243	1.303	2.988	1.925	2.347	0.404	2.993
15	1.984	2.311	1.903	2.684	2.004	2.224	0.273	2.879	1.804	2.324	0.376	2.909
20	1.963	2.252	1.883	2.665	1.603	2.204	0.204	2.838	1.705	2.300	0.307	2.878
25	1.949	2.194	1.108	2.649	1.583	2.184	0.226	2.779	1.684	2.294	0.326	2.799
30	1.868	2.135	0.838	2.624	1.563	2.164	0.156	2.721	1.664	2.275	0.257	2.742
35	1.843	2.077	0.766	2.604	1.543	2.142	0.138	2.662	1.644	2.253	0.238	2.685
40	1.827	2.018	0.612	2.584	1.523	2.124	0.122	2.604	1.625	2.234	0.225	2.634
45	1.803	1.960	0.590	2.564	1.504	2.104	0.115	2.545	1.605	2.215	0.215	2.576
50	1.863	1.901	0.573	2.544	1.483	2.084	0.088	2.487	1.588	2.195	0.189	2.497

Based on the Theil's Analysis above, the INAR (3) has the highest forecasting power around 30 steps ahead under 20% of excess zero, 15 step ahead under 30% of excess zero and all step ahead under 40% of excess zero in the model. However, the Theil values of INAR (1), INAR (2) and INAR (4) is not good in forecasting, but their ability increase as steps ahead increases.

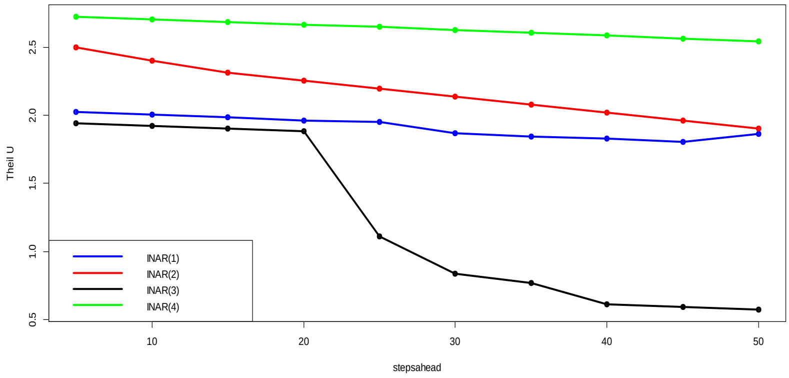

Figure 22: Forecast performance of the INAR (p) model with 20% of excess zero

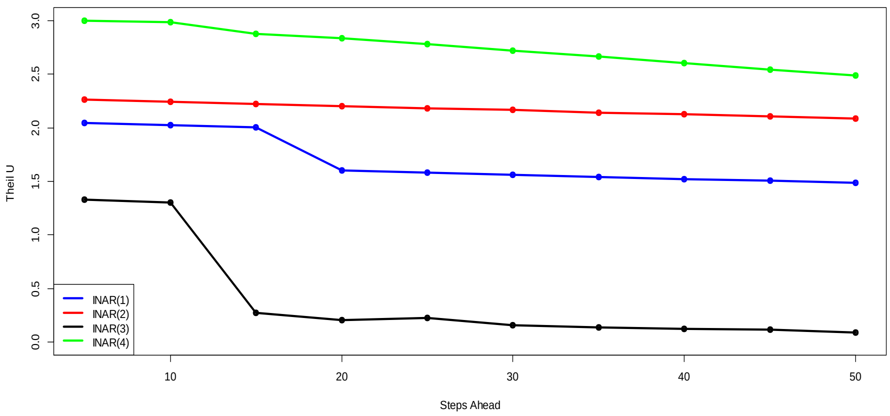

Figure 23: Forecast performance of the INAR (p) model with 30% of excess zero

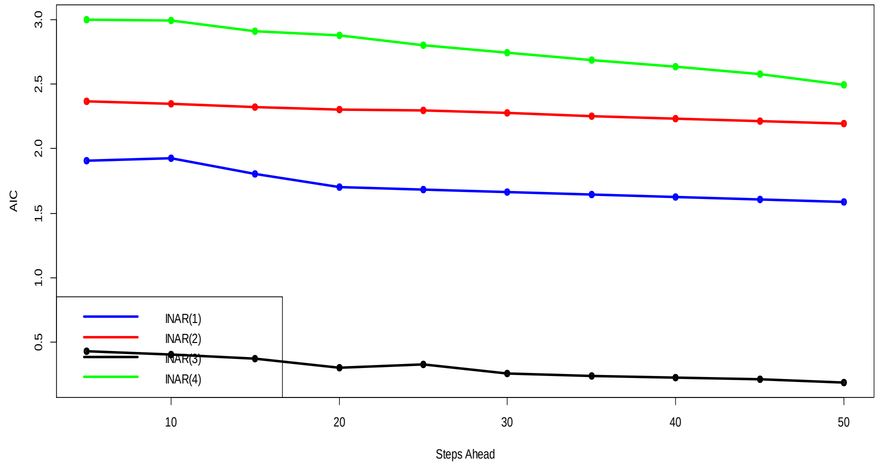

Figure 24: Forecast Performance of The INAR (P) Model with 40% Of Excess Zero

From figure 22, 23 and 24, based on the Theil's U Analysis above, the INAR (3) has the highest forecasting power due to the value is less than 1 at some point in the step ahead for all percentages of excess zeros.

CONCLUSION

This study addresses the critical challenge of modeling and forecasting discrete count time series with excess zeros—a common feature in data such as daily disease cases, financial transactions, or rare event counts. Through a rigorous simulation-based evaluation of Integer-Valued Autoregressive (INAR) models under varying levels of zero inflation, the research provides clear, evidence-based guidance for applied researchers and analysts. The study established that while simpler models like INAR (1) perform adequately for standard count data, they deteriorate as zero inflation severity increases. Notably, the INAR (3) model emerges as the most robust choice for data with moderate to high zero inflation (20–40%). It consistently demonstrates superiority in-sample fit and provides the most accurate short- to medium-term forecasts, owing to its flexibility in capturing autocorrelation patterns distorted by structural zeros. The forecasting advantage of INAR (3) is most pronounced in the near term, with competing models gaining relative performance only over longer horizons.

These findings actively engage with and extend contemporary methodological discourse. They affirm recent calls for specialized discrete models over traditional approaches, resonating with reviews by Weiß (2023). The empirical superiority of INAR (3) under contamination provides concrete support for theoretical analyses, such as those by Karlis & Tsiamyrtzis (2024), on the behavior of thinning operators in zero-inflated contexts. Furthermore, the study's scenario-specific model ranking refines the general selection process discussed by Aleksandrov & Weiss (2024). Finally, the demonstrated need for more sophisticated frameworks aligns with advancements like the GLAR models of Liboschik et al. (2023) and points toward the next logical innovation: developing mechanistic hybrid models, such as Zero-Inflated Thinning INAR (ZIT-INAR), as previewed in work like Chen et al. (2022).

REFERENCES

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving-average models. Biometrika, 60, 255–266.

Al-Osh, M. A., & Alzaid, A. A. (1987). First-order integer-valued autoregressive INAR (1) process. *Journal of Time Series Analysis*, 8, 261–275

Cameron, A. C., & Trivedi, P. K. (2013). *Regression analysis of count data* (2nd ed.). Cambridge University Press.

Harvey, A., & Fernandes, C. (1989). Time series models for count or qualitative observations. *Journal of Business & Economic Statistics*, 7(4), 407–417.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. Technometrics, 34(1), 1–14.

Silva, M. E. (2015). Modelling time series of counts: An INAR approach. Textos De Matemática, DMUC, 47, 107–122.

Ridout, M., Demétrio, C. G. B., & Hinde, J. (1998). Models for count data with many zeros. International Biometric Conference, Cape Town, December 1998.

McKenzie, E. (1985). Bayesian predictions of low count time series. *International Journal of Forecasting*, 21, 315–330.

Ndwiga, A. M., Ngesa, O., Wanjoya, A., & Mulwa, D. F. (2019). Comparison of statistical models in modelling over dispersed count data with excess zeros. International Journal of Research and Innovation in Applied Science, 4(5), 80–90.

Phang, Y. N., & Loh, E. F. (2023). Zero-inflated models for overdispersed count data. *International Journal of Mathematical, Computational, Physical and Quantum Engineering*, 7(8)

Saleh, I. M., & Nweze, N. O. (2021). Model selection for time series count data with over-dispersion. *Asian Journal of Probability and Statistics*, 14(2), 60–73.

Saleh, I. M., Akeyede, I., Abubakar, A. M., & Sulaiman, A. I. (2024). Exploring the INAR model on heavy-tailed time series data with outliers. FULafia *Journal of Science & Technology*, 8(1).

Schwarz, G. (1978). Estimating the dimension of a model. Annals of Statistics, 6(2), 461–464.

Hannan, E. J., & Quinn, B. G. (1979). The determination of the order of an autoregression. Journal of the Royal Statistical Society: Series B (Methodological), 41(2), 190–195.

Silva, I., Silva, M. E., Pereira, I., & Silva, N. (2005). Replicated INAR (1) process. Methodology and Computing in Applied Probability, 7, 517–542.

Steutel, F. W., & Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. The Annals of Probability, 5, 893–899.

Tawiah, R. (2021). Zero-inflated time series modelling of COVID-19 deaths in Ghana. Journal of Environmental and Public Health, 2021, Article ID 5543977. https://doi.org/10.1155/2021/5543977

Theil, H. (1966). Applied economic forecasting. Rand McNally.

Xekalaki, E. (2014). On the distribution theory of over-dispersion. *Journal of Statistical Distributions and Applications*, 1(2), 2–22. https://doi.org/10.1186/s40488-014-0019-z

Aleksandrov, B., & Weiss, C. H. (2024). Model selection for count time series: A focus on information criteria and residual analysis. Statistical Papers.

Chen, Y., Kang, J., & Wang, X. (2022). Bayesian inference for zero-inflated negative binomial INAR (1) models with application to infectious disease data. Journal of Applied Statistics.

Karlis, D., & Tsiamyrtzis, P. (2024). On the use of thinning-based models for integer-valued time series with excess zeros. Statistics and Computing.

Liboschik, T., Kerschke, P., & Fried, R. (2023). Modeling zero-inflated and overdispersed time series of counts: A new R package for generalized linear autoregressive models. Journal of Statistical Software.

Weiß, C. H. (2023). Flexible count time series models: A review and benchmark. Wiley Interdisciplinary Reviews: Computational Statistics.