Asian Journal of Probability and Statistics

14(2): 60-73, 2021; Article no.AJPAS.72702
ISSN: 2582-0230

Model Selection for Time Series Count Data with Over-Dispersion

Saleh Ibrahim Musa[1*] and N. O. Nweze[2]

[1]Department of Statistics, Federal University of Lafia, PMB 146 Lafia, Nigeria.
[2]Department of Statistics, Nasarawa State University Keffi, PMB 122 Keffi, Nigeria.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2021/v14i230326
Editor(s):
(1) Dr. Dariusz Jacek Jakóbczak, Koszalin University of Technology, Poland.
Reviewers:
(1) Jyoti U Devkota, Kathmandu University, Nepal.
(2) Puneet Kumar Gupta, The ICFAI University, India.
(3) Cliff Richard Kikawa, Kabale University, Uganda.
Complete Peer review History: https://www.sdiarticle4.com/review-history/72702

Original Research Article

Received 14 June 2021
Accepted 19 August 2021
Published 24 August 2021

Abstract

Time series of count with over-dispersion is the reality often encountered in many biomedical and public health applications. Statistical modelling of this type of series has been a great challenge. Rottenly, the Poisson and negative binomial distributions have been widely used in practice for discrete count time series data, their forms are too simplistic to accommodate features such as over-dispersion. Unable to account for these associated features while analyzing such data may result in incorrect and sometimes misleading inferences as well as detection of spurious associations. Therefore, the need for further investigation of count time series models suitable to fit count time series with over-dispersion of different level. The study therefore proposed a best model that can fit and forecast time series count data with different levels of over-dispersion and sample sizes Simulation studies were conducted using R statistical package, to investigate the performances of Autoregressiove Conditional Poisson (ACP) and Poisson Autoregressive (PAR) models. The predictive ability of the models were observed at different steps ahead. The relative performance of the models were examined using Akaike Information criteria (AIC) and Hannan-Quinn Information Criteria (HQIC). Conclusively, the best model to fit was ACP at different sample sizes. The predictive abilities of the four fitted models increased as sample size and number of steps ahead were increased

Keywords: Time series; count data; over-dispersion; forecasting.

**Corresponding author: Email: salehibro22@gmail.com;*

1 Introduction

Over-dispersion is a phenomenon that occurs in count data from binomial, poison or negative binomial distributions. Data is overdispersed, if the variance of the data distribution is greater than the mean. In other words, if the estimated dispersion after fitting is not near the assumed values, then the data may be overdispersed, the value is greater than the expected value. It is underdispersed, if the value is less than expected. It is generally caused by positive correlation between responses or by excess variation between response probabilities or counts. It also arises when there are violations in the distributional assumptions of the data [1]. Violations of Poisson assumptions usually result in over-dispersion, where the variance of the model exceeds the value of the mean. Excess or (deficiency) of zero counts result in over-dispersion. Violations of equidispersion indicate correlation in the data, which affect standard errors of the parameter estimates. Model fit is also affected [2].

In medical and health related research, data are often collected in the form of counts which are related to the number of times that an event of interest occurs. Because of their simplicity, one-parameter distributions for which the variance is directly determined by the mean are often used at least in the first method to model this data. However, the equal mean-variance relationship rarely happens with real-life data [3,4,5]. In most cases, the observed variance is larger than the assumed variance, which is known as over-dispersion. If the over-dispersion is ignored, statistical inference results in an inaccurate conclusion by underestimating the variability of the data [3]. If this dispersion is not taken into account, then using these models may lead to biased estimates of the parameters and consequently incorrect inferences about the parameters. Several statistical methods have been proposed for analysis of count data with over-dispersion. Many of them used negative binomial distribution to model the count data [6,7,8]. In their studies, they demonstrated the use of various models for overdispersed count data. These are Poisson, negative binomial, Quasi-Poisson, and Zero-inflated models. The models underestimated the standard errors and overstated the significance of some covariates.

Ndwiga et al. [9] confirmed the in appropriate use of negative binomial distributions and poison distributions in modelling count time series especially with over dispersion. The researcher further proposes the use of hurdle poison model for analysing data with ove-dispersion. Qian et al. [10] considered modelling of heavy tailed count time series data on number of traded stock in 5 min for interval Empire District Electric Company using heavy tailed probabilities, he further recommend the use of INAR of order p to analyse heavy tailed count time series data. The commonly used INAR and ACP in aforementioned literature is of order one [INAR(1) and ACP(1,1)]. This study therefore, aimed at extending the order of the models in order to determining the best model to fit and forecast count data at different levels of over-dispersion, sample sizes and steps aheads.

2 Methodology

Data set were simulated in R statistical software with sample sizes of 30, 60, 90, … and 300, from poison and negative binomial distributions to produce count data with equidispersion and over-dispersion respectively. The two models under study, namely: PAR and ACP were fitted to the simulated data so as to examine the effect of the proportion of over-dispersion on their performances. Levels of over-dispersion were imposed with difference between means and variances to be 5, 10 and 20 from the simulated data on observation of yi in the different data sets generated, which were randomized and replicated 1000 times each for the respective selected sample sizes.

In simulation, we set our parameters to be $\emptyset_1$ =1 $\emptyset_2$ = 1 to ensure discrete nature of count data generated. The response Y_{ti} in 3.1 were generated from poison and negative binomial distributions. The two models under study were considered to analyze how well each of the model fits the selected data sets having some degree of over dispersion and excess zeros.

Data were generated from linear second order of autoregressive function given as follows:

Model 1. AR (2): $Y_{ti} = 0.2Y_{ti-1} + 0.4Y_{ti-2} + e_t$ 1

$t = 30, 60, 90, 120, 150, 180, 210, 240, 270, 300. \; i = 1,2, \dots, 1000$

Where Y_{ti} will be simulated from poison and negative binomial families for equidispersion/excess zeros and over-dispersion respectively as follows:

The basic count model is the Poisson regression model which is based on the Poisson distribution with probability density function

$$\frac{\lambda^{y_i} e^{-\lambda_i}}{y_i!}, for\ y_i = 0,1,2,\dots \tag{2}$$

Thus, for the Poisson models $E(y_i) = V(y_i) = \mu_i$. The restrictive condition that the mean must equal the variance is often violated by overdispersed data (where variance exceeds the mean). As a result of that Poisson model is generally considered inappropriate for count data, which are usually highly skewed and overdispersed [11].

The over-dispersion is achieved from the Negative binomial distribution function given as follows;

$$p(y_i; \lambda_i, \alpha_i) = \frac{\Gamma\left(y_i + \frac{1}{\alpha_i}\right)}{\Gamma(y_i + 1)\Gamma\left(\frac{1}{\alpha_i}\right)} \left(\frac{1}{1 + \alpha_i \lambda_i}\right)^{\frac{1}{\alpha_i}} \left(\frac{\alpha_i \lambda_i}{1 + \alpha_i \lambda_i}\right)^{y_i}, i = 1, \dots, 1000 \tag{3}$$

Here, the dispersion parameter$\alpha_i > 0,\ \lambda_i = E(Y_i)$; and $V(Y_i) = \lambda_i + \alpha_i \lambda_i^2$

The Negative Binomial model can be used to impose the over-dispersion problem on y_i by creating larger values of variances than means. Lawal [12] argued that the Negative Binomial (NB) model might be a suitable alternative to the Poisson model especially for overdispersed count data. This is because the NB model in this case would account for the heterogeneity in the data by introducing the dispersion parameter α. In order to compare the modeling and forecasting accuracy of the models, AIC and HQIC criteria for performance evaluation procedure were used in this study. The model with the minimum criteria values were considered as the best for the fitting and forecasting. Note that a number of steps ahead were forecasted from each model.

2.1 Autoregressive Conditional Poison (ACP) model

The ACP Model proposed in this study has counts follow a Poisson distribution with an autoregressive mean. Let F_t denote the information available on the series up to and including time t. In the simplest model, the counts are generated by a Poisson distribution

$$Y_t / F_{t-j} \sim P(y_t, \mu_t) = \frac{\mu^y e^{-\mu}}{y!} \tag{4}$$

with an autoregressive conditional intensity as in the ACD model of Engle and Russell [13] or the conditional variance in the GARCH (Generalised Autoregressive Conditional Heteroskedasticity) model of Bollerslev [14]:

$$E(Y_t / F_{t-j}) = \mu_t = \sum_{j=1}^{p} \alpha_j Y_{t-j} + \sum_{j=1}^{p} \beta_j \mu_{t-j} + \omega \tag{5}$$

for positive αj's, βj's and ω.

We call this model the Autoregressive Conditional Poisson (ACP (p,q)). The following properties of the unconditional moments of the ACP can be established.

Unconditional mean of the ACP (p,q)). Provided that

$$\sum_{j=0}^{\max(p,q)} \left(\alpha_j + \beta_j\right) < 1 \qquad 6$$

the ACP(p,q) is stationary and its unconditional mean is

$$E(Y_t) = \mu = \frac{\omega}{1 - \sum_{j=0}^{\max(p,q)} \left(\alpha_j + \beta_j\right)} \qquad 7$$

This proposition shows that, as long as the sum of the autoregressive coe□cients is less than 1, the model is stationary and the expression for its mean is identical to the mean of an ARMA process. For instance, the mean equation of ACP (1,1) is then given as:

$$E\left(Y_t / F_{t-j}\right) = \mu_t = \omega + \alpha_1 Y_{t-1} + \beta_1 \mu_{t-1} \qquad 8$$

2.2 Poisson Autoregressive (PAR) model

The poison autoregressive or PAR (p) model can be define as

$$P\left({}^{Q_t}/_{s_t}\right) = \frac{{s_t}^{q_t} e^{-s_t}}{q_t!} \qquad 9$$

Where s_t is the conditional mean of the linear autoregressive AR process with $E(q_t/Q_{t-1})$ in (16)

This represent the measurement equation for the observed data. The one step ahead for the conditional PAR (p) model forecast is given by

$$E(q_{t+1}/Q_t) = s_{t+1/t} = \sum_{i=1}^{k} \rho_i s_{\frac{t}{t}-1} + (1 - \sum_{i=1}^{k} \rho_i)\mu \qquad 10$$

$$Var(q_{t+1}/Q_t) = \frac{1 + \sigma_{t+1/t}}{\sigma_{t+1/t}} s_{t+1/t} \qquad 11$$

Where $\rho, \delta, s_t and\ \sigma_t$ are the optimized values of a PAR series, the induced covariance X_t has the $\mu = e^{(X_t \delta)}$. See [15].

2. 2.1 PAR (p) Forecast density for the one step ahead distribution

The PAR (p) forecast density is given by

$$P(q_t/Q_{t-1}) = \int_{\emptyset} \Pr\left(q_t/\emptyset_t\right) \Pr\left(\emptyset_t/Q_{t-1}\right) d\emptyset \qquad 12$$

$$= \int_{\emptyset} \frac{{\emptyset_t}^{q_t} e^{-\emptyset_t}}{q_t} \cdot \frac{e^{-\sigma_{t/t-1}\emptyset} {\emptyset_t}^{\sigma_{t/t-1}s_{t/t-1}} {\sigma_t}^{\sigma_{t/t-1}s_{t/t-1}}}{\Gamma\left(\sigma_{t/t-1}s_{t/t-1}\right)}$$

$$= \frac{\Gamma\left(\sigma_{t/t-1}s_{t/t-1} + q_t\right)}{\Gamma(q_t+1)\Gamma\left(\sigma_{t/t-1}s_{t/t-1}\right)} (\sigma_{t/t-1})^{\sigma_{t/t-1}s_{t/t-1}} \times (1 + \sigma_{t/t-1})^{\sigma_{t/t-1}s_{t/t-1} + q_t} \qquad 13$$

This is a negative binomial distribution function with a gamma function $\Gamma(\cdot)$.

The forecast function for the conditional mean and variance of a PAR (p) series realizations are based on the optimized values of $\rho, \delta, s_t and\ \sigma_t$. The log-likelihood function for the PAR (p) model is given as

$$L(s_{t-1,}\sigma_{t-1}/q_t, \dots, q_T;\ Q_{t-1}\ ) = ln \prod_{t=1}^{T} P(q_t/Q_{t-1}) \qquad 14$$

$$= \sum_{t=1}^{T} ln\Gamma(\sigma_{t/t-1}s_{t/t-1} + q_t) - ln\Gamma(q_t + 1) - ln\Gamma(\sigma_{t-1}s_{t-1})$$
$$+\sigma_{t-1}s_{t-1}\ln(\sigma_{t-1}) - (\sigma_{t-1}s_{t-1} + q_t) + \ln(1 + \sigma_{t-1}) \qquad 15$$

Using the linear autoregressive equation

$$E(q_t/Q_{t-1}) = \sum_{i=1}^{p} \rho_i Q_{t-i} + \lambda \qquad 16$$

Where ρ_i and λ are real number values.

We can obtain AR (1) for q_t which yield PAR (1) model with a negative binomial predictive distribution, for order p can also be generated as well. There is no restriction for the linear AR process with respect to the density$P(q_t/Q_{t-1})$. The q_t density choice resulted constraints to ρ_i and λ to require admissible values.

3 Analyses, Results and Discussion

The results of simulation and analysis for relative performances of different orders of the ACP and PAR models based on the criteria of the assessment at different sample sizes and levels of over-dispersion are presented in Tables 1-10. The corresponding values of the tables are plotted in graphs for more clarity.

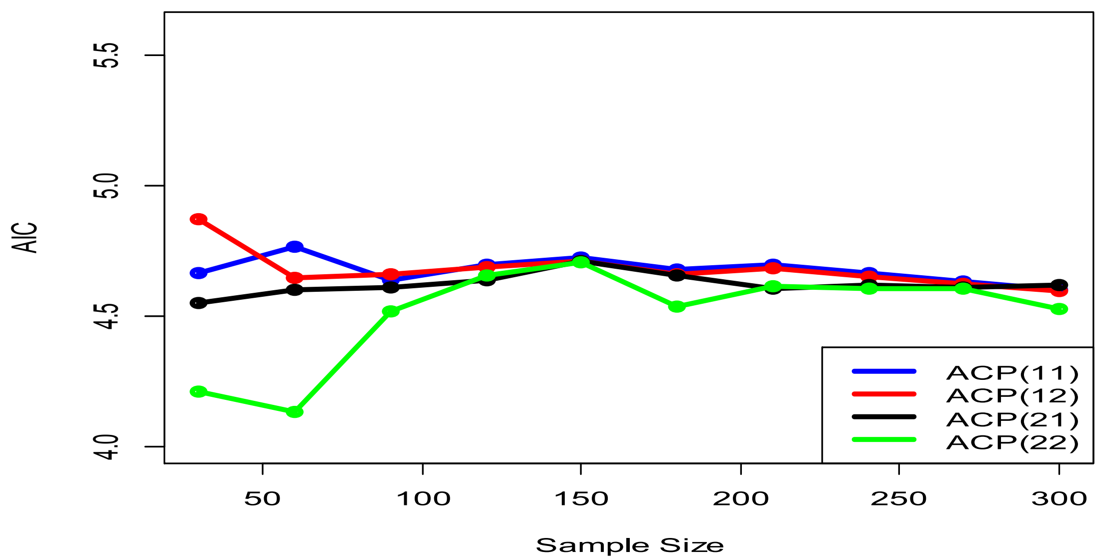

Fig. 1a. AIC of the fitted ACP (p, q) models when there is no over-dispersion

The AIC values tabulated in Table 1 and the values were shown in Fig. 1a and 1b for the ACP and PAR respectively, the ACP(2, 2) model performed well compared to ACP (2, 1) at all scenarios of sample sizes. The ACP (1, 1) and ACP (1, 2) gave similar results in most scenarios. As per the results of the AIC, ACP (2, 2) seems the preferred choice for equidispersed count time series data. Relatively, PAR (4) performed greatly in comparison to their PAR models followed closely by PAR (2) while PAR (1) performed poorly. Generally, ACP (2, 2) is the best among all at different sample sizes.

Table 1. AIC values of the fitted models in the absence of over-dispersion

Sample Sizes Model	**30**	**60**	**90**	**120**	**150**	**180**	**210**	**240**	**270**	**300**
ACP(1,1)	4.665	4.764	4.639	4.694	4.723	4.678	4.697	4.6647	4.633	4.6018
ACP(1,2)	4.869	4.645	4.661	4.6877	4.709	4.661	4.68	4.6491	4.6225	4.594
ACP(2,1)	4.549	4.601	4.607	4.6379	4.711	4.656	4.603	4.6172	4.61	4.6191
ACP(2,2)	4.208	4.134	4.519	4.6539	4.705	4.536	4.612	4.6035	4.6057	4.5243
PAR(1)	59.488	107.060	150.541	226.18	289.07	350.652	408.34	460.980	510.301	555.605
PAR(2)	57.829	104.560	147.503	205.91	261.39	312.037	359.51	407.400	456.553	500.005
PAR(3)	58.890	104.920	145.985	208.334	263.35	314.528	362.89	409.740	456.815	505.011
PAR(4)	55.230	101.370	142.812	203.193	259.34	306.787	354.85	400.030	436.834	493.270

Table 2. HQIC of models performance when there is no over-dispersion

Sample Sizes Models	**30**	**60**	**90**	**120**	**150**	**180**	**210**	**240**	**270**	**300**
ACP(1,1)	152.923	291.968	423.519	569.806	715.264	849.148	993.675	1127.13	1258.695	1388.477
ACP(1,2)	152.823	291.892	423.148	567.047	711.275	844.178	988.272	1117.579	1253.854	1380.763
ACP(2,1)	154.505	294.462	423.978	568.601	709.826	841.957	985.563	1118.652	1252.291	1381.776
ACP(2,2)	151.497	294.454	423.954	538.327	701.812	831.718	980.457	1108.31	1242.015	1371.521
PAR(1)	61.281	110.335	154.573	230.708	293.962	355.831	413.750	466.589	516.081	561.534
PAR(2)	59.622	107.833	151.535	210.438	266.284	317.216	364.921	413.007	462.333	505.934
PAR(3)	60.683	108.193	150.017	212.862	268.243	319.706	368.299	415.351	462.595	510.94
PAR(4)	57.023	104.642	146.844	207.721	264.228	311.966	360.262	405.635	442.614	499.199

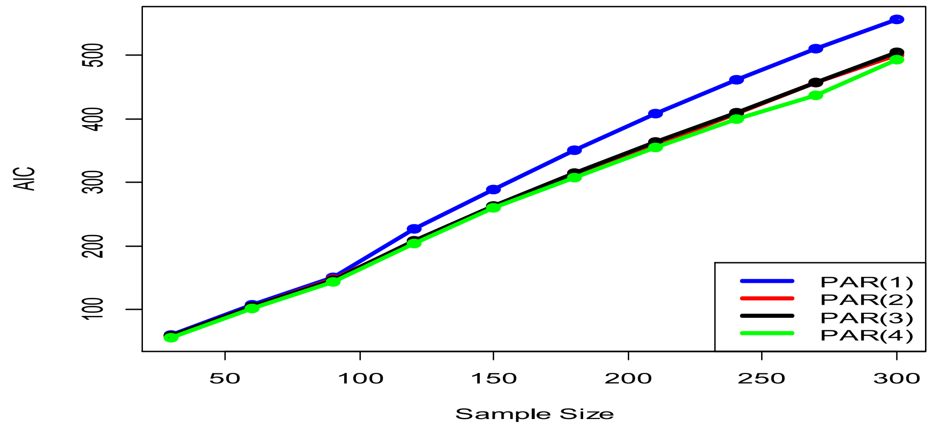

Fig. 1b. AIC of PAR Model across Sample Sizes when there is No Over-Dispersion

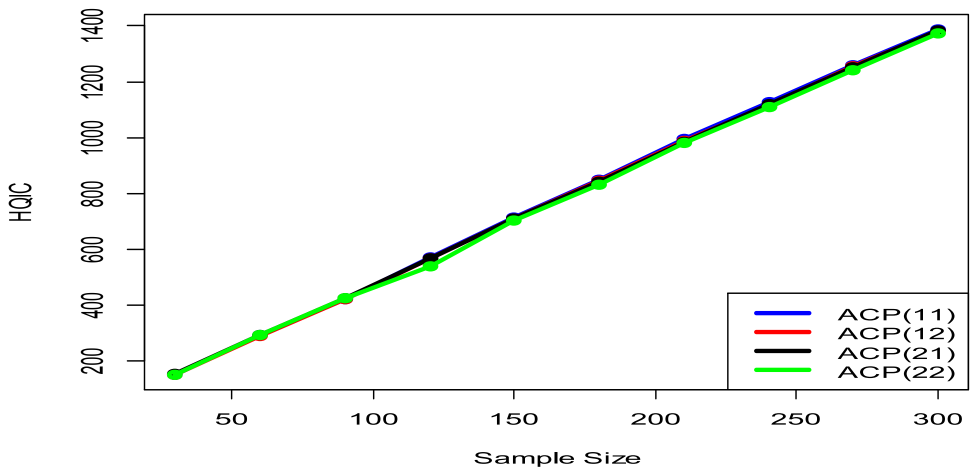

Fig. 2a. HQIC of the fitted ACP (p, q)) models when there is no over-dispersion

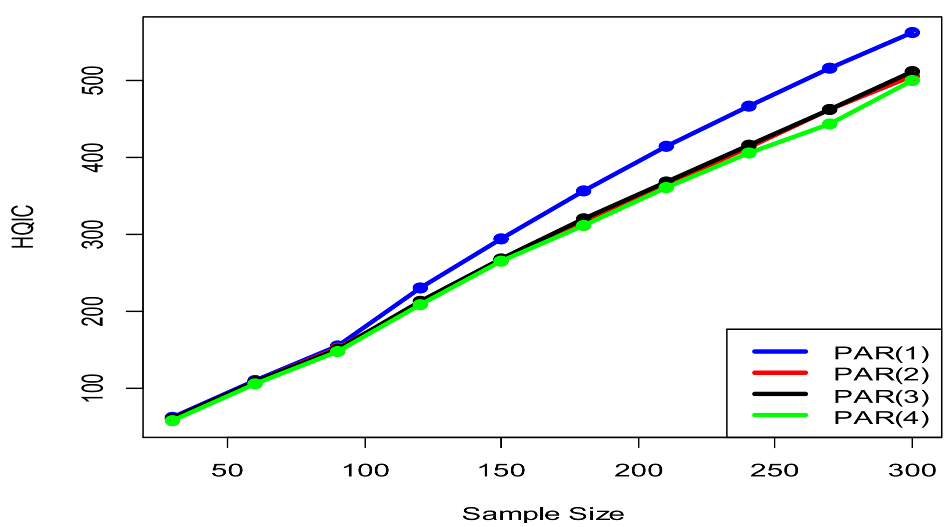

Fig. 2b. HQIC of the Fitted PAR (p) Models When There is No Over-dispersion

The plot of HQIC values from Table 2 of the fitted APC (p, q) and PAR (p) models are displayed in Fig. 2a and 2b. ACP (2, 2) has the best fits across samples with no over-dispersion, followed by ACP (2, 1) which exhibits good performance closely to rest of the ACP models across sample sizes. The PAR (p) model HQIC values follows similar pattern with the earlier reported criterion with PAR (4) as the better fitted, having the minimum HQIC values across sample sizes, followed by improved PAR (2) especially at sample sizes above 240.

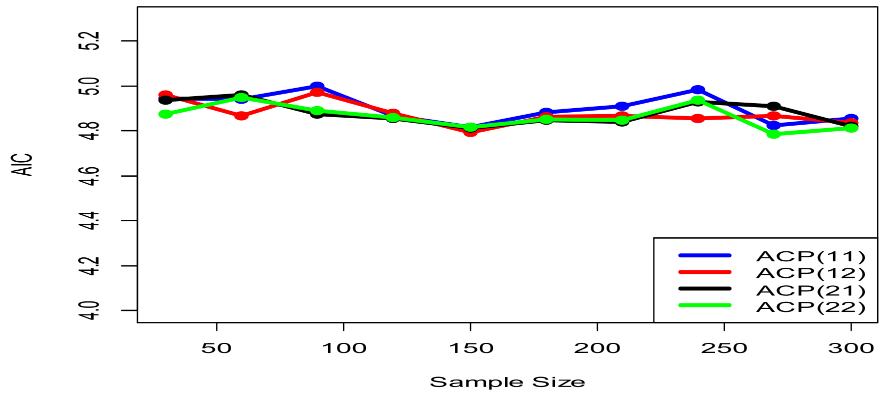

Fig. 3a. AIC of the fitted ACP (p, q) models when there is low over-dispersion

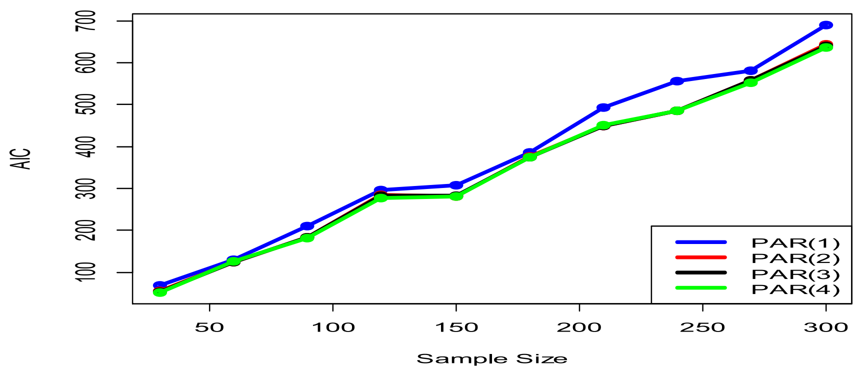

Fig. 3b. AIC of the fitted PAR (p) models when there is low over-dispersion

The average values of AIC of each model at various sample sizes when there is low over-dispersion are presented in Table 3. ACP (2,1) model exhibits great performance especially below the sample size of 250 followed by ACP (2, 2). When the sample size increases the performances of ACP (2, 2) and other models increase. In, PAR models, the trend shows PAR (4) as the most well performed followed by PAR (3) as reported by the AIC values. Generally ACP (2, 1) is the best among all the models.

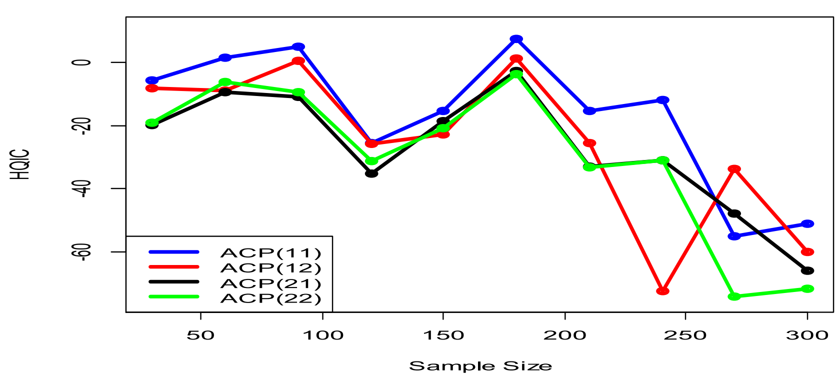

Fig. 4a. HQIC of the fitted ACP (p, q) models when there is low over-dispersion

Table 3. AIC of performance of models when there is low over-dispersion

Sample Sizes Models	30	60	90	120	150	180	210	240	270	300
ACP(1,1)	4.946	4.942	5.001	4.867	4.818	4.882	4.908	4.983	4.823	4.854
ACP(1,2)	4.961	4.869	4.972	4.880	4.795	4.862	4.869	4.857	4.868	4.835
ACP(2,1)	4.939	4.960	4.875	4.857	4.813	4.849	4.839	4.928	4.908	4.819
ACP(2,2)	4.875	4.948	4.889	4.861	4.817	4.853	4.847	4.936	4.787	4.812
PAR(1)	69.384	130.109	209.532	296.340	308.388	385.975	493.652	556.198	581.594	689.140
PAR(2)	56.257	124.194	183.449	283.977	280.763	377.155	448.157	484.546	558.775	642.966
PAR(3)	53.218	125.165	183.628	282.210	282.121	375.353	449.984	484.928	558.506	640.447
PAR(4)	52.407	125.471	182.338	277.485	281.155	375.229	451.131	485.677	552.496	636.167

Table 4. HQIC of performance of models when there is low over-dispersion

Sample Sizes Models	30	60	90	120	150	180	210	240	270	300
ACP(1,1)	-5.748	1.6404	5.0758	-25.616	-15.41	7.3478	-15.277	-11.945	-55.056	-51.105
ACP(1,2)	-8.118	-8.7969	0.511	-25.728	-22.84	1.2982	-25.519	-72.293	-33.616	-60.073
ACP(2,1)	-19.69	-9.322	-10.875	-35.165	-18.61	-2.804	-33.012	-31.068	-47.874	-65.905
ACP(2,2)	-19.17	-6.1306	-9.3957	-31.294	-20.86	-3.688	-33.271	-31.044	-74.047	-71.685
PAR(1)	71.177	133.386	213.564	300.868	313.28	391.154	499.065	561.808	587.374	695.07
PAR(2)	58.05	127.471	187.481	288.505	285.655	380.334	453.569	490.156	564.555	648.895
PAR(3)	55.011	128.442	187.661	286.738	287.014	380.531	455.397	490.538	564.286	646.376
PAR(4)	54.200	128.748	186.370	282.013	286.047	380.408	456.544	491.287	558.276	642.096

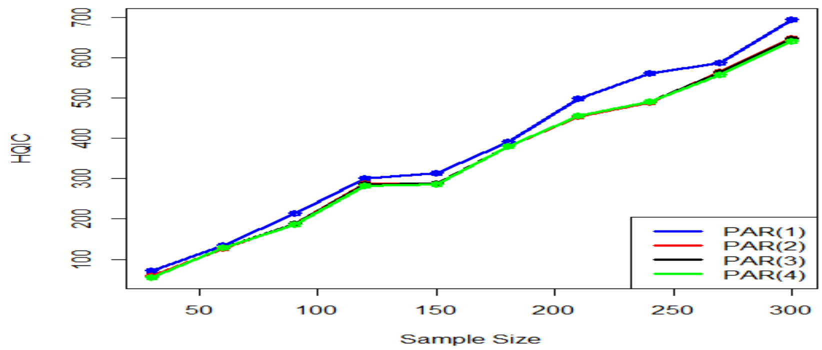

Fig. 4b. HQIC of the fitted PAR (p) models when there is low over-dispersion

The values of HQIC of each model at various sample sizes when there is low over-dispersion are presented in Table 4. ACP (2,1) performed good at low sample sizes from the ACP models especially below the sample sizes of 120 followed by ACP (2,2) become best above sample size of 270 based on the minimum values of the HQIC criteria. When the sample size increases the performances of ACP (2, 2) and other models increase. Relatively, PAR (4) display good trend pattern across sample size followed by PAR (3) in closed linear trend in Fig. 4b based on the minimum reported HQIC values.

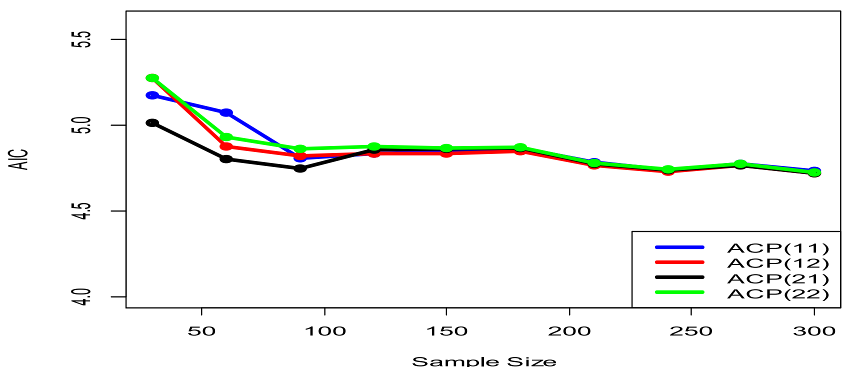

Fig. 5a. AIC of the fitted ACP (p, q) models when there is high over-dispersion

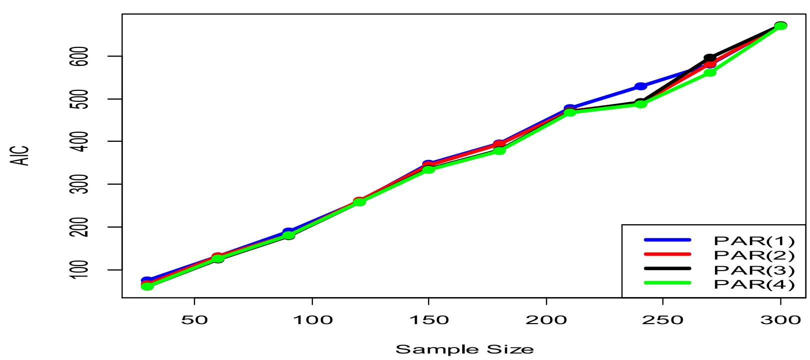

Fig. 5b. AIC of the fitted PAR (p) models when there is high over-dispersion

Table 5 shows the fitted performances of the ACP and PAR models to data simulated under high levels of over-dispersion with the average values of AIC of each model at various sample sizes. The results obtained were plotted on the graphs as shown in Fig. 5a and 5b respectively. However, the performance of the APC (2, 1) models supersedes others at the moderate sample size of 120 based on AIC criterion, whereas, ACP (1, 2) well performed above 120 sample sizes. In Fig. 5b, the PAR (3) model takes the lead at lower sample sizes below 120 while PAR (3) fitted best to the data with high over-dispersion when the sample size increases from 120.

Table 5. AIC of models' performance when there is high over-dispersion

Sample Sizes Models	30	60	90	120	150	180	210	240	270	300
ACP(1,1)	5.175	5.072	4.805	4.835	4.846	4.864	4.783	4.739	4.774	4.731
ACP(1,2)	5.274	4.874	4.820	4.834	4.834	4.847	4.766	4.727	4.763	4.719
ACP(2,1)	5.011	4.800	4.746	4.857	4.861	4.867	4.775	4.737	4.766	4.719
ACP(2,2)	5.274	4.932	4.863	4.874	4.865	4.871	4.779	4.742	4.772	4.724
PAR(1)	74.58	130.144	188.625	260.059	347.376	395.961	478.435	530.096	581.229	672.386
PAR(2)	66.098	130.388	180.997	260.824	344.699	394.287	470.679	491.140	582.200	672.578
PAR(3)	60.095	124.167	178.072	257.294	334.585	378.622	469.404	491.978	595.877	671.328
PAR(4)	61.299	126.373	181.082	258.406	332.904	377.198	468.359	487.501	560.898	670.636

Table 6. HQIC of Models' performance when there is high over-dispersion

Sample Sizes Models	30	60	90	120	150	180	210	240	270	300
ACP(1,1)	13.060	11.130	-41.367	-48.402	-50.405	-47.201	-74.308	-102.14	-114.93	-127.6
ACP(1,2)	12.827	-29.135	-42.714	-52.272	-57.181	-55.579	-84.277	-111.44	-123.24	-137.10
ACP(2,1)	1.726	-27.167	-40.049	-47.56	-53.465	-53.56	-83.409	-110.08	-122.81	-138.93
ACP(2,2)	12.827	-27.380	-40.836	-47.498	-56.079	-55.971	-85.699	-111.55	-123.43	-139.52
PAR(1)	76.373	133.42	192.657	264.587	352.269	401.140	483.848	535.706	587.009	678.315
PAR(2)	67.891	133.665	185.029	265.352	349.591	399.465	476.091	496.749	587.980	678.507
PAR(3)	61.088	126.844	182.405	260.822	339.477	383.801	474.816	497.588	601.657	677.257
PAR(4)	63.092	129.650	185.114	262.935	337.796	382.376	473.771	439.111	566.678	675.565

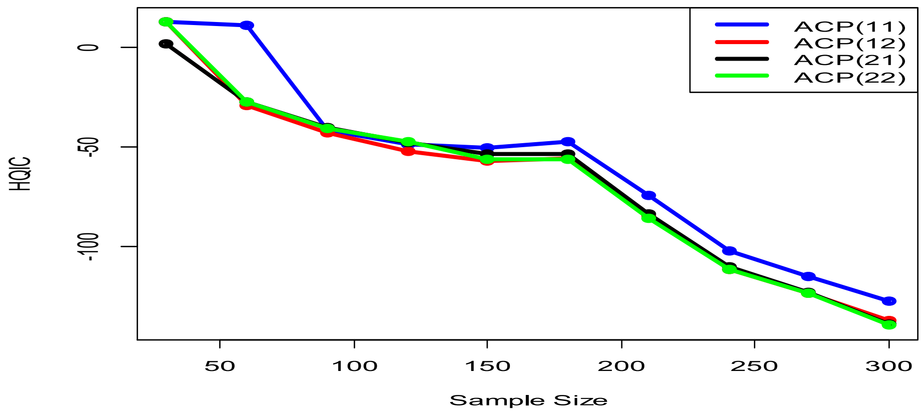

Fig. 6a. HQIC of the fitted APC (p, q) models when there is high over-dispersion

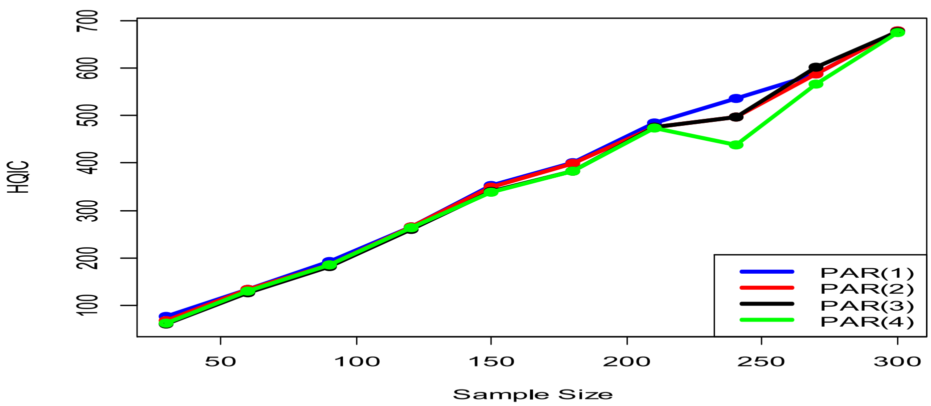

Fig. 6b. HQIC of the fitted PAR (p) models when there is high over-dispersion

Table 6 shows the fitted performances of the ACP and PAR models to data simulated under high levels of over-dispersion with the average values of HQIC of each model at various sample sizes recorded. The results obtained were plotted on the graphs as shown in Fig. 6a and 6b respectively. The ACP (2,1) model takes the lead at lower sample sizes below 60 while ACP (1,2) at less than 180, APC (2, 2) fitted best to the data with high over-dispersion when the sample size increases from 180. Indeed, the ACP (2, 1) and ACP (2, 2) are best at lower and higher sample sizes respectively. More so, in Fig. 6b, PAR (3) model takes the lead at lower sample sizes below 150 while PAR (4) fitted best to the data with high over-dispersion when the sample size increases from 150. Indeed, the PAR (3) and PAR (4) are best performed at lower and higher sample sizes respectively.

3.1 Forecast ability of the selected best models

The predictive ability of the best three models selected from ACP and PAR where 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. U>1 indicate that the forecasting technique is better than guessing, U = 1, indicate that the forecasting technique is as good as guessing, U<1 indicates that the forecasting technique is worse than guessing. The results for the Theil U test of the two best orders forecasted for at different steps ahead in the three models with different level of over-dispersions are presented in Table 4.21 and 4.22.

The relative forecast performance of the selected best models among APCs and PARs at different categories of dispersion using TheilU statistics were presented in table 7. ACP (2,2) has the best forecasting ability when there is no and low overdispersions, while ACP (2,1) considered to be better in forecasting ability at high overdispersion than PAR models.

Table 7. Forecast performance of the models without over-dispersion, with over dispersion and there is low over dispersion using theil U statistic

	No Over Dispersion		High Over Dispersion		Low Over Dispersion	
Steps Ahead	**ACP (2,2)**	**PAR (4)**	**ACP (2,1)**	**PAR (3)**	**ACP (2, 2)**	**PAR (4)**
5	2.8411	2.5721	2.7249	1.9439	2.2642	1.3301
10	2.8142	2.5452	2.7065	1.9239	2.2439	1.3032
15	2.7873	2.5184	2.6843	1.9038	2.2246	0.2734
20	2.7604	1.4915	2.665	1.8838	2.2043	0.2044
25	2.7335	1.4646	2.6490	1.1086	2.1843	0.2255
30	2.7066	1.4377	2.6249	0.8385	2.1647	0.1565
35	2.6797	1.4108	2.6049	0.7667	2.1424	0.1376
40	2.6528	1.3839	2.5849	0.6123	2.1240	0.1218
45	2.6259	1.3570	2.5648	0.5901	2.1048	0.1149
50	2.59908	1.3301	2.5448	0.5736	2.0841	0.0880

4 Conclusion

This study discovered the highest performing model in fitting and forecasting different count time series data with different levels of over-dispersion is the ACP model based on all criteria of the assessment. The model has the speedy fitting capabilities at both high and low sample sizes. PAR models has the slowest fitting speed across sample sizes. Specifically, the ACP (2,2) has the highest performance followed by ACP (2,1) among all the models in fitting any time series count data with the underlying features reported in this research. The research focused on the simulated data and recommend for further application on real life data.

Disclaimer

The products used for this research are commonly and predominantly use products in our area of research and country. There is absolutely no conflict of interest between the authors and producers of the products because we do not intend to use these products as an avenue for any litigation but for the advancement of knowledge. Also, the research was not funded by the producing company rather it was funded by personal efforts of the authors.

Competing interests

Authors have declared that no competing interests exist.

References

[1] Breslow NE. Tests of hypotheses in Overdispersed Poisson regression and other quasilikelihood models. Journal of American Statistical Association. 1990;85:565-571.

[2] Hilbe JM. Negative Binomial Regression. Cambridge University Press, United Kingdom. 2008;51-61.

[3] Cox D. Some remarks on overdispersion, Biometrics. 1983;10:269–274.

[4] Dean CB. Testing for overdispersion in Poisson and binomial regression models. J. Am. Stat. Assoc. 1992;87:451–457.

[5] Prentice RL. Binary regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors. J. Am. Stat. Assoc. 1986;81:321–327.

[6] Pounds SCG, Zhang H. Empirical bayesian selection of hypothesis testing procedures for analysis of sequence count expression data, Statistical Applications in Genetics and Molecular Biology. 2012;11(7).

[7] Auer P, Doerge R. A two-stage poisson model for testing rna-seq, Statistical Applications in Genetics and Molecular Biology. 2012;10:26.

[8] McCarthy DJ MR, Smyth G. A bioconductor package for differential expression analysis of digital gene expression data, Bioinformatics. 2010;26:139–140.

[9] Ndwiga AM, Oscar Ngesa, Anthony Wanjoya, Damaris Felistus Mulwa. Comparison of Statistical Models in Modeling OverDispersed Count Data with Excess Zeros. International Journal of Research and Innovation in Applied Science (IJRIAS). 2019;IV(V):80-90.

[10] Qian L, Li Q, Zhu F. Modelling heavy-tailedness in count time series Applied Mathematical Modelling. 2020;82(2020):766–784.

[11] Cameron CA, Trivedi PK. Regression analysis of count data, econometric society monographs No. 30, Reprinted edition, Cambridge University Press, New York. 2008;1(96-97):221.

[12] Lawal HB. On zero-truncated generalized poisson count regression model, Pioneer Journal of Theoretical and Applied Statistics. 2011;1(1):16.

[13] Engle Robert, Russell Jeffrey R. Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data, Econometrica. 1998;66(5):1127.

[14] Bollerslev Tim. Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics. 1986;31(3):307-327.

[15] Brandt Patrick, Williams John. A linear poisson autoregressive model: The Poisson AR(p) Model. Political Analysis. 2001;9:164-184.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle4.com/review-history/72702