GSJ: Volume 11, Issue 12, December 2023, Online: ISSN 2320-9186

www.globalscientificjournal.com

On The Forecasting Accuracy Of Time Series of Count Data Using TheInteger-Valued First Order Moving Average (INMA(1)) Model With Negative Binomial (NB) Innovation

NasiruMukailaOlakorede,Samuel OlayemiOlanrewaju, andYahayaHaruna Umar

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

Email Address: nasiru.olakorede@uniabuja.edu.ng, yahaya.umar@uniabuja.edu.ng, and olanrewaju.samuel@uniabuja.edu.ng

Abstract

In this paper, we focused our attention on the theoretical investigation of integer-valued first order moving average (INMA(1)) model with negative binomial (NB) innovation for the forecasting accuracy of time series of count data. We employed the method of the Conditional Least squares (CLS) estimator to estimate the parameter of INMA(1) model, and Maximum Likelihood Estimator (MLE) to estimate the mean (μ) and the dispersion parameter (k) of the NB distribution. The study is based on simulation experiment used to addressed the theoretical generated dataunder different parameter valuesβ=0.2, 0.6, 0.8, different sample sizesn=30, 90, 120, 600 for the class of INMA(1) model, and μ =0.85, 1.5, 2, k=1,2, 4 for the NB distribution. The Monte Carlo simulations were conducted with codes written in R, all results were based on 1000 runs.With small number of observations and high parameter value, the estimation of parameter for the class of INMA(1) model gives a better results. Similarly, theestimation of the dispersion parameter (k) of the NB distribution gives a better result when the number of observations is small and with large k values.The forecasting accuracy of the INMA(1) model with the mean of NB distribution at different lead time period l=1, 3, 5, 7, 9, 15 were investigated with codes written in R. The results showed that the minimum mean square error(MMSE) produced when the number of lead times forecasts is between one and five were less than that produced when the numbers of lead times forecast were greater than five. The MMSE increased when the number of lead time periods increases. This result indicates that forecasting with this class of model is better with short time frame of predictions. The study is applied to the number of Measles laboratory confirmed cases in Nigeria which consist of count time series data of 37 observations (monthly data), from January 2021 to December 2021. The application results corroborate the theoretical investigation.

Keywords:INMA(1) model, NB distribution, count data, CLS estimation, MLE estimation, Accuracy, Measles,Forecasting

1.Introduction

The Integer-valued Moving Average (INMA) models is a special case of Integer Autoregressive Moving Average (INARMA) models which has recently received wider attention in the literature. The INARMA models are capable of modelling and forecasting Time Series of Count data that appears in several diverse scientific especially for low frequency count with overdispersed data. A time series of count data is an integer-valued non-negative sequence of count observations observed at equidistant instants of time.

The INARMA model was originally introduced in the 1980s (Mc Kenzie, 1985, Al-Osh and Alzaid, 1987). The INARMA models have been proposed for forecasting time series of counts, and have received wider attentions in the last three decades. This model has been shown to be analogous to well-known conventional time series model namely Autoregressive Moving Average (ARMA) models (Box *et al*, 1994), for modelling continuous data.

The most distinguish feature that makes the Integer-valued Moving Average (INMA) models class different from its continuous Moving Average (MA) counterpart is that multiplication of the models with real valued parameter no longer remains a viable operation when the result is to be integer-valued.

The study and analysis of count time series poses several problems and questions. For instance, a common distribution that is used in practice to model the response time series, is the Poisson distribution. Such an assumption is sensible because the Poisson distribution is the simplest discrete distribution, yet its properties are satisfactory to cover a large class of problems (Vasiliki Christou, 2013).

Several investigations into the classes of INARMA models with the assumption that the innovation distribution are Poisson distribution have been carried out. But the Poisson distribution has a feature of equal mean-variance relationship which makes it inadequate for modeling time series count data because most of the count data have properties of overdispersion. In this paper, we investigates theINMA(1) model with the assumption that the innovation distribution is the Negative Binomial distribution. The Negative Binomial distribution is capable of taking into account the overdispersion found in time series count data. Hence, the justification.

Modeling discrete -valued time series is the most challenging and, yet, least well-developed of all areas of research in time series. The fact that variate values are integer, renders most traditional representations of dependence either impossible or impractical. In the past there have been a number of imaginative attempts to develop a suitable class of models (Eddie Mc Kenzie, 2000). In recent times, Fokianos (2012), Davis and Liu (2015) have made an effort to the development of models appropriate for discrete valued time series. Such data usually occur in the form of counts rendering the traditional ARMA-type models impractical. Steutel and Harn (1979), proposed the most popular count time series models that are based on the notion of binomial thinning. Brännäsand Quoreshi (2010), Quoreshi (2006, 2008, 2014), proposes a bivariate integer-valued moving average (BINMA) model, a vector integer-valued moving average (VINMA) model, and an integer-valued autoregressive fractionally integrated moving average (INARFIMA) model for analyzing high frequency financial data.

For the Innovation distribution ε_t, of INARMA models, many models have been proposed in the literature for the integer-valued time series count data. The Poisson distribution is often assumed as the distribution of ε_t in the INARMA models. The Poisson distribution has a characteristic of equidispersion. In practice, however, count data are overdispersed in nature relative to the Poisson distribution. For this reason, the INARMA models with Poisson innovations is not always suitable for modeling integer-valued time series, therefore, several models which describe the over-dispersion phenomena have been discussed in the statistical literature.

Changing the distribution of innovations is also used to modify the INAR(1) model. Jung et al. (2005) indicated that the INAR(1) model with negative binomial innovation (NB- INAR(1)) is appropriate for generating overdispersion. Jazi et al.(2012) defined a zero-inflated Poisson ZIP(p, λ)for innovation (ZIP- INAR(1)), because a frequent occurrence in overdispersion is that the incidence of zero counts is generated than expected from the Poisson distribution. Jazi et al. (2012) proposed a modification of INAR(1) model with Geometric innovation (G-INAR(1)) for modeling overdispersed count data. Schwer and Weiß (2014) investigated the compound Poisson INAR(1) (CP- INAR(1)) model, which is suitable for fitting data sets with overdispersion. According to Schwer and Weiß (2014) the negative binomial distribution and the geometric distribution both belonging to the compound Poisson distribution. Livio et al. (2018) presented the INAR(1) model with the Poisson-Lindely innovations, that is, PL-INAR(1) model. Bourgnignon et al. (2019) introduced the INAR(1) model with double Poisson (DP- INAR(1)) and generalized Poisson innovations (GP- INAR(1)) model. Qi et al. (2019) considered zero-order one-inflated INAR(1)-type models, and Cunha et al. (2021) introduced an INAR(1) model with Borel innovation to model zero truncated count time series. Huang and Zim (2021) introduced a new INAR(1) model with Bell innovations (BL- INAR(1)). Huang and Zim (2021) used a relative simple distribution introduced by Castellares et al. (2018) for innovation. Mahmoudiand Rostami (2020) introduced a first-order nonnegative integer-valued moving average process with power series innovations based on a Poison thinning operator (PINMAPS(1)) for modeling overdispersed and underdispersed count time series. Bouguinon and Vasconcellos (2015) introduced INAR(1) processes with power series innovations. Yu and Wang (2021) introduced a new overdispersed integer-valued moving average model with dependent counting series.Nasiru and Olanrewaju (2023) employed a simulation procedure to investigate the properties of INAR(1) model with NB innovation. In this paper we investigate the theoretical properties of INMA(1) model with NB innovation, and assess the practical validity and applicability of the main results of the study on real life data.

2. Methodology

2.1 The Binomial Thinning Operator

Before introducing the INMA(1) model, we first introduced the meaning of Binomial thinning operation and it properties.

The binomial thinning operation was defined by Steutel and Harn (1979). Suppose Y is a non-negative integer-valued random variable. Then, for any $\alpha \in [0,1]$, the thinning operation "∘" is definedby:

$\alpha \circ \mathrm{Y} = \sum_{i=1}^{y} x_i$(2.1)

Where $\{X_i\}$ is a sequence of i.i.d. Bernoulli random variables, independent of Y, and with a constant probability that the variable will take the value of unity:

$PX_i = 1 - PX_i = 0 = \alpha$ (2.2)

Some of the properties of the thinning operation can be obtained as follows:

(1) $0 \circ \mathrm{Y} = 0$

(2) $1 \circ \mathrm{Y} = Y$

(3) $\alpha \circ (\beta \circ Y) \stackrel{d}{=} (\alpha\beta) \circ Y$

(4) $(\circ Y) = \alpha E(Y)$

(5) $E\,(\alpha \circ Y)^2 = \alpha^2 (Y^2 + \alpha 1 - \alpha (Y)$

(6) $\mathrm{var} \alpha \circ Y = \alpha^2 \mathrm{var} Y + \alpha 1 - \alpha E(Y)$

2.2 Integer-Valued First Order Moving Average (INMA(1)) Model

Al-Osh and Alzaid (1988) developed a class of models for integer-valued moving average (INMA) processes. In INMA models, a stationary sequence of random variables $\{Yt\}$ is formed from a sequence $\{Zt\}$ of i.i.d. random variables which are non-negative and also integer-valued. The first order model which is the case in which adjacent members of the sequence are correlated. A process $\{Yt\}$ is called an INMA(1) process if it satisfies the equation:

$Y_t = \beta \,^{o} Z_{t-1} + Z_t$ (2.3)

where $\beta \in [0,1]$ and $\{Z_t\}$ are as before and the thinning operation is defined via:

$\beta \circ Z = \sum_{i=1}^{z} X_{\mathrm{I}}$ (2.4)

where $\{Xi\}$ is a sequence of i.i.d. Bernoulli random variables, independent of Y and satisfying:

$P(X_i = 1) = 1- p(\mathrm{X_i} = 0) = \beta$ (2.5)

The INMA(1) model defined by equation (2.3) is similar to the Gaussian MA(1) process except that scalar multiplication is replaced by the thinning operation.

Jung and Tremayne (2006a) present a physical interpretation of this model as follows. If we consider Y_t as the number of particles in a well-defined space at time point t, it can be assumed that this number is made of two components: (*i*) particles entering during (t−1,], and (*ii*) survivors of those who entered the space during (t−2,t−1]. Therefore, the thinning at time t, is applied to only immigrants at time t−1, not all particles in space, as in an INAR(1) process. Examples of this process include the number of patients staying in a hospital or the number of customers in a department storeas has been shown (Al-Osh and Alzaid, 1988).

It can be inferred from the equation (2.3) that each element stays in the system no longer than two periods. This is in contrast to the INAR(1) process in which there is no limit on the survival of elements in the system as has been shown (Mohammadipour and Brunel University, 2009).

The unconditional first and second moments of the INMA(1) process are:

$E(Y_t) = (1+\beta)\mu_z$ (2.6)

$Var(Y_t) = \beta(1-\beta)\mu_z + (1+\beta^2)\sigma_z^2$ (2.7)

Al-Osh and Alzaid (1988) shows that the autocorrelation function (ACF) of this process is given by:

$\rho_k = \left\{\frac{\beta\sigma_z^2}{\beta(1-\beta)\mu_z + (1+\beta^2)\sigma_z^2}\right.$, for k=1, and 0 for k>1 (2.8)

2.3 Method of Estimation

The Conditional Least Square (CLS) estimation method was employed in this research. Lawrence and Paul (1978) developed the Conditional Least Square (CLS) estimation procedure for stochastic processes based on the minimization of a sum of squared deviations about conditional expectation.

The conditional expected value of Yt given $Yt-1$ for an INMA(1) process is given by:

$$E(Y_t/Yt_{-1}) = \beta Z_{t-1} + \lambda \quad (2.9)$$

The prediction error is:

$$e_t = Y_t - \beta Z_{t-1} - \lambda \quad (2.10)$$

The CLS estimates of β and λ can then obtained by minimizing the following function:

$$Q_n(\boldsymbol{\theta}) = \sum_{t=1}^{n}[Y_t - (\beta Z_{t-1} + \lambda)]^2 \quad (2.11)$$

with respect to $\boldsymbol{\theta}$, where $\boldsymbol{\theta}=(\beta, \lambda)'$ is the parameter vector to be estimated. The CLS estimates for β and λ are:

$$\hat{\beta} = \frac{\sum_{t=1}^{n} Y_t Z_{t-1} - (\sum_{t=1}^{n} Y_t \sum_{t=1}^{n} Z_{t-1})/n}{\sum_{t=1}^{n} Z_{t-1}^2 - (\sum_{t=1}^{n} Z_{t-1})^2/n} \quad (2.12)$$

$$\hat{\lambda} = [\sum_{t=1}^{n} Y_t - \beta \sum_{t=1}^{n} Z_{t-1}]/n \quad (2.13)$$

2.4 Forecasting Method

One of the objectives of a time series models is to forecast the future values of a time series observations.

2.4.1 Minimum Mean Square Error (MMSE) Forecasts

The conditional expectation has been the most commonly used forecasting procedure discussed in the time series literature as has been shown (Freeland and McCabe, 2004b). The main advantage of this method, apart from being simple, is that it produces forecasts with minimum mean square error (MMSE).

Minimum mean square error (MMSE) forecasts are used to find $\hat{Y}_{T+h,} h = 1,2, \dots, H$ of the processes Y_t based on the observed series of $\{Y_1, \dots, Y_T\}$. The MMSE forecast of the process is given by:

$$\hat{Y}_{T+h,} = E(Y_{T+h} | Y_T, \dots, Y_1) \quad (2.14)$$

This method yields forecasts with minimum MSE. For an INAR(p) model, we have:

$$\hat{Y}_{T+h} = \alpha_1 Y_{T+h-1} + \alpha_2 Y_{T+h-2} + \cdots + \alpha_p Y_{T+h-p} + \mu \quad (2.15)$$

Where the Y values on the RHS of equation (2.15) may be either actual or forecast values as has been shown (Du and Li, 1991; Jung and Tremayne, 2006b).

2.4.2 Lead Time Forecasts for an INMA(1) Model

For the INMA(1) process of $Y_t = \beta \circ Z_{t-1} + Z_t$, the cumulative Y over lead time l is given

$$\sum_{j=1}^{l+1} Y_{t+j} = Y_{t+1} + Y_{t+2} + \cdots + Y_{t+l+1} = (\beta \circ Z_t + Z_{t+1}) + (\beta \circ Z_{t+1} + Z_{t+2})$$

$$+ \dots . + (\beta \circ Z_{t+1} + Z_{t+l+1}) \quad (2.16)$$

The above equation(2.16) can be written as:

$$\sum_{j=1}^{l+1} Y_{t+j} = \sum_{j=1}^{l+1} \sum_{i=1}^{n^j} \Psi_{ij} \circ Z_{t+k_{ij}} \quad (2.17)$$

Where n^j is the number of $Z_{t+k_{ij}}$ terms in each of $\{Y_{t+j}\}_{j=1}^{l+j}$ and Ψ_{ij} is the corresponding coefficient for $Z_{t+k_{ij}}$

It can be seen from equation (2.16) that because the process is an integer moving average of order one, each of $\{Y_{t+j}\}_{j=1}^{l+j}$ only has two $Z_{t+k_{ij}}$ and therefore $n^j = 2$ The corresponding coefficient for each $Z_{t+k_{ij}}$, is shown by Ψ_{ij} is :

$\{\Psi_{ij} = \beta\, \Psi_{2j}\} = 1. t + k_{ij}$ is the subscript of innovation terms in each of $\{Y_{t+j}\}_{j=1}^{l+j}$ From equation (2.16), it can be seen that $\{k_{ij} = j - 1, k_{2j} = j\}$ Therefore,

$k_{ij} = \{\begin{smallmatrix} j-1 & for\ i=1 \\ j & for\ i=2 \end{smallmatrix}\}\ for\ j = 1, \ldots, l = 1$ (2.18)

Based on equation(2.17), the conditional expected value of the aggregated process is:

$E\left(\sum_{j=1}^{l+1} Y_{T+j} \mid Y_T\right) = +\left(\sum_{j=1}^{l+1} \sum_{i=1}^{2} \Psi_{\mathrm{IJ}}\right)\mu\left(\sum_{j=1}^{l+1}(1+\beta)\right)\mu = (L+1)(1+\beta)\mu$ (2.19)

2.5 The Negative Binomial (NB) Distribution

The innovation distribution assumed in this research is the negative binomial distribution. The negative binomial distribution has two parameters: the mean $\mu$and the shape parameter or the dispersion parameter k, which is commonly considered to be fixed to measure overdispersion. For a sample of counts Xthat fits a negative binomial distribution (X~ NB(μ, k)), the variance of the distribution is

$\mu + \mu^2 / k$. The probability that the variable *X* takes the value xis:

Prb[X=x]= $\frac{\Gamma(x+k)}{x!\Gamma(k)}\left(\frac{\mu}{\mu+k}\right)^x\left(1+\frac{\mu}{k}\right)^{-k} = \frac{(x+k-1)(x+k-2)\ldots(k+1)k}{x!}\left(\frac{\mu}{\mu+k}\right)^x\left(1+\frac{\mu}{k}\right)^{-k}$, $\mu, k>0$,

x=0,1,2,… (2.20)

Where $\Gamma(.)$ denotes the gamma function defined by:

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt. \quad (2.21)$$

From the probability density function of the negative binomial distribution, it can be seen that k is an essential part of the model. Estimation of k is thus important given a sample of counts.

In this research, the method of maximum likelihood estimator (MLE) is adopted to estimate the mean and the dispersion parameter of the NB. According to Fisher, the log-likelihood function from a sample of independent identically distributed (i.i.d.) variate $(x_i^{'}, s)$ is proportional to:

$$l(\text{k},\mu) = \frac{1}{n}\sum_{i=1}^{n} log\left(\frac{\Gamma(x_i+k)}{\Gamma(k)}\right)+\bar{x}\log(\text{u})-(\bar{x}+k)\log(1+\frac{\mu}{\text{k}}) \qquad (2.22)$$

Where μ is again the mean of the negative binomial distribution. The sample variate are integers in practice, which yields:

$$\frac{\Gamma(x+k)}{\Gamma(k)}= (\text{x+k-1})(\text{x+k-2})\ldots(\text{k+1})\text{k. the term } log\left(\frac{\Gamma(x_i+k)}{\Gamma(k)}\right) \qquad (2.23)$$

then can be written as:

$$log\left(\frac{\Gamma(x_i+k)}{\Gamma(k)}\right)=\sum_{v=0}^{x_i-1} klog\left(1+\frac{v}{k}\right) \qquad (2.24)$$

Without call to the gamma function.

Thus, the log-likelihood function can finally be expressed by:

$$l(\text{k},\mu)=\frac{1}{n}\sum_{i=1}^{n}\sum_{v=0}^{x_i-1} klog\left(1+\frac{\mu}{k}\right)+\bar{x}\log(\mu)-(\bar{x}+k)\log(1+\frac{\mu}{\text{k}}) \qquad (2.25)$$

With gradient elements

$\nabla_\mu \text{l}=\frac{\bar{x}}{\mu}-\frac{1+\bar{x}/k}{1+\mu/k}$ and

$$\nabla_k \text{l}=\frac{1}{n}\sum_{i=1}^{n}\sum_{v=0}^{x_i-1}\left(\frac{v}{1+v/k}\right)+k^2\log(1+\frac{\text{u}}{\text{k}}) - \frac{\mu(\bar{x}+k)}{1+\mu/k}. \qquad (2.26)$$

From the gradient element, setting ∇_μ l=0 yields $\hat{\mu}=\bar{x}$. Then the MLE of k can be obtained via a nonlinear root-finder by setting ∇_k l=0 and given $\mu = \hat{\mu}$.

3 Results and Interpretations

This section focuses on the result which is based on the simulation study of theoretical investigation of the class ofINMA(1) model with NB innovation. The study makes use of the Conditional Least squares (CLS) estimate to estimate the parameter of INMA(1) model, and

Maximum Likelihood Estimate (MLE) to estimate the mean and the dispersion parameter of the NB distribution.

3.1 Estimation of Parameters ForINMA(1) Model and NB Distribution

A simulation experiment based on theoretical generated data were addressedunder different parameter values and different sample sizes. The Monte Carlo simulations were conducted with a code written in R, all results were based on 1000 runs.

Equation (2.12) is used to estimate the parameter of INMA(1) model, with parameter values β =0.2, 0.6, and 0.8, different sample sizes n=30, 90, 120, and 600, and with number of replication r= 1000.

In estimating the mean and dispersion parameter of the NB distribution, equation (2.25) is used with the following parameter values: μ =0.85, 1.5 and 2 respectively. K:1,2, and 4 for each μ, with number of replication r= 1000 times. The results and interpretation are shown below:

Table3.1: Parameter Estimate of CLS Estimator for INMA(1) Series

Replication=1000

Estimator	Parameter setting (β)	Parameter Estimate and S.E	Sample Size (n)			
			30	90	120	600
Conditional Least Square	0.2	$\hat{\beta}$	0.1250	0.1247	0.1251	0.1244
		S.E	0.2341	0.2311	0.2354	0.2295
	0.6	$\hat{\beta}$	0.9286	0.6505	1.0222	0.6056
		S.E	0.0874	0.0212	0.0629	0.1985
	0.8	$\hat{\beta}$	1.0565	0.8875	0.9710	0.7861
		S.E	0.0294	0.1063	0.0491	0.1625

Table3.1 presents the results of the parameter estimates of the Integer Moving Average of order 1 (INMA(1)) model. The first row reports the parameter estimates of the model, while the second row reports the standard errors (S.E) of each of the estimates obtained by simulation. Results are based on 1000 replication.

The results show that, the standard error (SE) produced by the Conditional Least Squares (CLS) increases as the number of samples increases. The SE reduces as the parameter values increases. This means that estimating the parameter of INMA(1) model is better when the number of observations is small and the value of the parameter is high.

Table3.2: Maximum Likelihood Estimation of K and μ of Negative Binomial Distribution for n=30,90,120, and 600

Sample Size (n)	μ	K=1	K=2	K=4
30	0.85	$\hat{k}$=0.4965 $\hat{\mu}$=1.0666 AIC=87.8352	$\hat{k}$=0.5387 $\hat{\mu}$=0.600 AIC=66.8343	$\hat{k}$=0.4967 $\hat{\mu}$=0.5334 AIC=62.7172
	1.5	$\hat{k}$=0.6576 $\hat{\mu}$=1.7331 AIC=87.8352	$\hat{k}$=0.5936 $\hat{\mu}$=1.000 AIC=84.8273	$\hat{k}$=0.9652 $\hat{\mu}$=0.9667 AIC=85.7712
	2.0	$\hat{k}$=0.6942 $\hat{\mu}$=2.2664 AIC=123.8003	$\hat{k}$=0.4612 $\hat{\mu}$=1.5340 AIC=102.8366	$\hat{k}$=0.9775 $\hat{\mu}$=1.3001 AIC=98.4756
90	0.85	$\hat{k}$=0.7585 $\hat{\mu}$=0.9332 AIC=244.4199	$\hat{k}$=3.1373 $\hat{\mu}$=0.7999 AIC=222.3790	$\hat{k}$=2.2996 $\hat{\mu}$=0.8667 AIC=232.7188

	1.5	$\hat{k}$=0.8171 $\hat{\mu}$=1.5669 AIC=312.3411	$\hat{k}$=2.2563 $\hat{\mu}$=1.4444 AIC=296.6447	$\hat{k}$=2.7858 $\hat{\mu}$=1.5556 AIC=303.4139
	2.0	$\hat{k}$=1.2249 $\hat{\mu}$=1.9217 AIC=341.3869	$\hat{k}$=1.5143 $\hat{\mu}$=2.0444 AIC=348.6402	$\hat{k}$=4.5232 $\hat{\mu}$=2.1221 AIC=335.6395
120	0.85	$\hat{k}$=0.8268 $\hat{\mu}$=0.9501 AIC=327.8408	$\hat{k}$=2.9590 $\hat{\mu}$=1.008 AIC=331.9965	$\hat{k}$=2.5928 $\hat{\mu}$=0.9418 AIC=321.632
	1.5	$\hat{k}$=0.8957 $\hat{\mu}$=1.6167 AIC=421.4639	$\hat{k}$=2.1894 $\hat{\mu}$=1.6752 AIC=421.6663	$\hat{k}$=2.6640 $\hat{\mu}$=1.5831 AIC=406.9943
	2.0	$\hat{k}$=1.1353 $\hat{\mu}$=1.9331 AIC=455.4091	$\hat{k}$=1.9922 $\hat{\mu}$=2.2333 AIC=476.1481	$\hat{k}$=2.8983 $\hat{\mu}$=2.1333 AIC=458.4345
600	0.85	$\hat{k}$=1.0327 $\hat{\mu}$=0.7931 AIC=1481.292	$\hat{k}$=1.8040 $\hat{\mu}$=0.8583 AIC=1532.379	$\hat{k}$=1.7264 $\hat{\mu}$=0.8099 AIC=1443.602
	1.5	$\hat{k}$=1.0882 $\hat{\mu}$=1.450 AIC=1991.472	$\hat{k}$=2.2187 $\hat{\mu}$=1.5402 AIC=2010.441	$\hat{k}$=6.5003 $\hat{\mu}$=1.4667 AIC=1901.274
	2.0	$\hat{k}$=1.0480 $\hat{\mu}$=1.9848 AIC=2287.914	$\hat{k}$=2.3424 $\hat{\mu}$=2.0767 AIC=2279.278	$\hat{k}$=5.2949 $\hat{\mu}$=1.9482 AIC=2149.624

Table 3.2 present the result of the maximum likelihood estimation results for K and μ of NB at different simple sizes. Comparing the AIC of the result at different K values and at different sample sizes, the estimation produced with minimum value of AIC andwith low sample sizes especially when n=30 is better. However, as the number of k increases the value of AIC decreases. This means that estimation of K of NB distribution is better when the number of observations is small and the more the dispersion the better for the estimation.

3.2 Forecasting in INMA(1) Model With NB Innovation

This section concentrate on the investigation of the forecasting accuracy of INMA(1) model, with NB innovation. The forecast accuracy at different lead time period l=1, 3, 5, 7, 9, and 15 were investigated with codes written in R statistical package. All results were based on 1000 runs.

At time T, when $Y_T$is observed, the lead time forecast is obtained using equation (2.19), with the following Parameter values: l=1, 3, 5, 7, 9, and 15. β= 0.81, μ=0.75, $j = 1, \ldots, l + 1$, and Y_T=30, 90, 120, 600, number of replication r=1000. The result is summarized in the table 3.3.

Table 3.3 MMSE of The Lead Time Forecasts For INMA(1) Series

Y_T	l=1	l=3	l=5	l=7	l=9	l=15
30	3.230813e-05	3.230813e-05	3.250729e-05	3.270645e-05	3.290561e-05	3.350309e-05
90	1.833389e-05	1.833389e-05	1.833389e-05	1.899000e-05	1.964611e-05	2.161444e-05
120	0.00017582	0.000175821	0.00017582	0.00019715	0.000218478	0.000282465

	12	2	12		8	2
600	0.001127962	0.001127962	0.001127962	0.001127962	0.00143882	0.00243985

Table 3.3 presents the results of the MMSE forecasts for INMA(1) series with NB innovation. The results show that, the MMSE produced when the number of lead time forecasts is between one and five were less than that produced when the numbers of lead times forecast were greater than five. The MMSE increased when the number of lead time periods increases. This result indicates that, making forecasts with this class of model would be better with short time frame of predictions.

3.3 Application to Measles Data Set Laboratory Confirmed Cases in Nigeria

The theoretical investigations result obtained in this study was applied to the number of Measles cases (Laboratory confirmed cases) in Nigeria. The count time series data consists of 37 observations (Monthly data), from January 2021 to December 2021 across 36 states and FCT. The data was obtained from the Nigeria Centre for Disease Control (NCDC), and analyzed with the aid of R statistical package. The results and the interpretation of the analysis are presented below:

Table3.4 Descriptive Statistic of Measles Cases

	Measles cases (Lab confirmed cases)
Mean	47.48
Median	36.50
Maximum	159
Minimum	7
Variance	1298.78

Observations	37

Table3.4 depicts the summary statistic of the number of Measles confirmed cases in Nigeria in the year 2021. From the table, the mean and the variance respectively are 47.48 and 1298.78 which is an evident of overdispersion in the data.

Table3.5 Preliminary Test

Test type	Test value	p-value	Decision
Overdispersion	Z=3.1074	0.0009438	Reject H_0
Autocorrelation	Chi-square=3.0651	0.0003435	Reject H_0

Table3.5 presents the Preliminary test of the number of confirmed Measles in Nigeria in the year 2021. The results suggest that the null hypotheses (H_0) (i.e. no true dispersion and autocorrelation in the series and its residuals respectively) cannot be accepted, thus there is true dispersion in the data set and presence of autocorrelation in the residuals, which corroborates the descriptive analysis. Hence, a negative binomial distribution is assumed for the innovation.

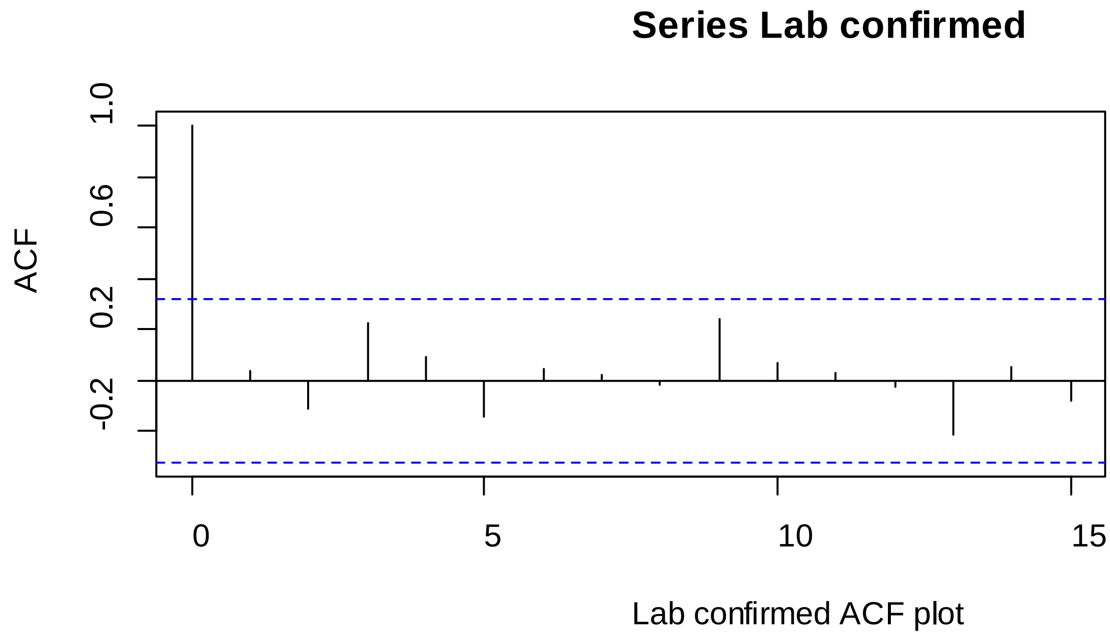

Fig3.1 ACF Plot of Measles cases

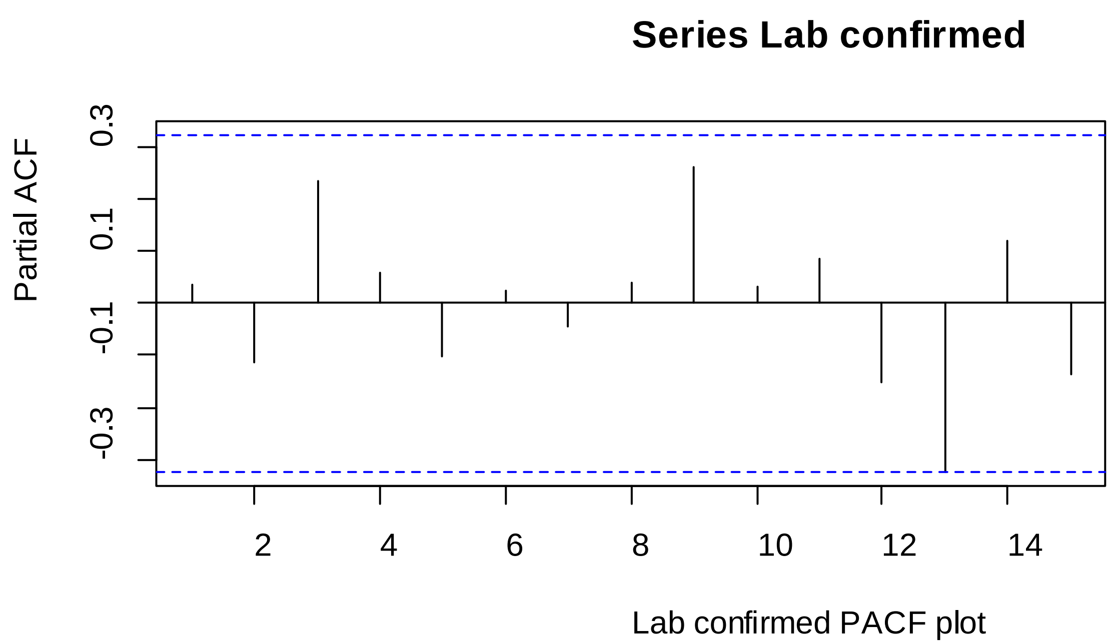

Fig3.2 PACF plot of Measles cases

Fig3.1 and Fig3.2 respectively depicts the plots of ACF and PACF respectively. Based on the information supplied by the plots, it is clear that moving average model is suggested. The candidates' models in table3.6 were compared. Comparing the AIC of the models in table3.6. the INMA(1) model gives the minimum AIC and hence an INMA(1) model best fit the data set.

Table3.6 Candidates of INARMA Models

Candidate Models	AIC
INARMA(2,2)	422
INARMA(0,1)	415
INARMA(1,0)	416
INARMA(1,1)	425

Table3.7 Parameter Estimate of INMA(1) Model

Model	Parameter estimate	Std. Error	p-value
INMA(1)	-0.30899	1.13925	0.02649

Table3.7 presents the estimate of parameter of INMA(1) model. In line with the simulation result, Conditional Least Square (CLS) estimation method was employed. The parameter of the model was tested and found to be statistically significant at 5% level of significance ($p< 0.05$).

Table3.8 Lead Time Forecast of Measles Cases Using INMA(1) Model

Lead time	1	3	5	7	9
Forecast	107.0171	107.0459	107.0486	107.0489	107.0489
MMSE	1.971362	1.972074	1.972141	1.972148	1.972159

Table3.8 depicts the lead time forecast of the Measles cases in Nigeria, using the fitted INMA(1) model. The accuracy of the forecast is measured by the MMSE. The error of forecast increases as the number of lead time increased. This result is in line with the theoretical investigations obtained in this study. This shows that forecasting with this class of model is betterwith short time prediction.The results show that, the number of Measles forecasted shows an increasingtrend, but with an approximately equal values as the number of leads time increases,

4.0 Conclusion

From our findings, the following conclusions were drawn:

The results of the estimation of parameter of the INMA(1) model confirmed that the standard error (SE) produced by the Conditional Least Squares (CLS) increases as the number of samples increases, the error reduced as the parameter values increases. This means that estimating the parameter of INMA(1) model is better when the number of observations is small and the parameter value is high.

The result of the estimation of the parameters(K and μ) of NB distribution at different simple sizes. Comparing the Akaike Information Criterial (AIC) of the result at different K values and at different sample sizes, the estimation produced less AIC with low sample sizes especially when n=30. However, as the number of k increases the result showed a decreased in the value of AIC. This means that estimation of K of NB distribution is better when the number of observations is small and the more the dispersion the better for the estimation.

The forecasting accuracy were measured by the MMSE. The results showed that, the MMSE produced when the number of lead time forecasts between one and five were less than that produced when the numbers of lead times forecast were greater than five. The MMSE increased when the number of lead time periods increases. This result indicates that, forecasts with this class of model is better with short time frame of predictions.

Lastly, the theoretical investigations were validated with a real life data using the number Measles confirmed in Nigeria. The results obtained corroborates the results from the theoretical investigations.

REFRENCES

Al-Osh, M. A. and A. A. Alzaid (1987). First order integer-valued autoregressive (INAR(1)) processes. *Journal of Time Series Analysis* **8**(3): 261-275.

Al-Osh, M. A. and A. A. Alzaid (1988). Integer-valued moving average (INMA) process. *Statistical Papers* **29**(4): 281-300.

Box, G. E. P., G. M. Jenkins and G. C. Reinsel (1994). *Time series analysis: Forecasting and control,* 3rd ed., Prentice Hall: New Jersey.

Brännäs, K. and Quoreshi A.M.M.S. (2010). Integer-value moving average modeling of the transactions in stocks. Applied Financial Economics 20, 1429-1440.

Castellares, F.; Ferrari, S.L.P.; Lemonte, A.J (2018). On the Bell distribution and its associated regression model for count data. Appl. Math. Model. 56, 172–185.

Cunha, E.T.D.; Bourguignon, M.; Vasconcellos, K.L.P. (2021). On shifted integer-valued autoregressive model for count time series showing equidispersion, underdispersion or overdispersion. Commun. Stat.

Davis, R.A. and Liu, H. (2015). Theory and inference for a class of observation-driven models with application to time series of counts. Statistical sinica. doi: 10.5705/ss.2014.145t (to appear).

Du, J. G. and Y. Li (1991). The integer-valued autoregressive (INAR(p)) model.*Journal of Time Series Analysis* **12**(2): 129-142.

Eddie Mc Kenzie (2000). Discrete Variate Time Series

EisaMahmoudi, AmenehRostami (2020). First-order integer-valued moving average process with power series innovation. Journal of Statistical Theory and applications. Vol. 19(3), pp. 415-431.

Fokianos K (2012) Count time series models. In T Subba Rao S Subba Rao and CR Rao(eds). Handbook of statistics 30: Time series – methods and applications, pages 315–47. Amsterdam: Elsevier B.V.

Freeland, R. K. and B. P. M. McCabe (2004b). Forecasting discrete valued low count time series. *International Journal of Forecasting* **20**(3): 427-434.

Jazi MA Jones G and Lai CD (2012) First-order integer valued AR processes with zero inflated Poisson innovations. Journal of Time Series Analysis, 33, 954–63.

Jie Huang and Fukang Zhu (2021). A new first-order integer-valued autoregressive model with Bell innovations. Entropy (Basel), 23(6):713

Jung, R.C. and Tremayne, A.R. (2006a). Coherent forecasting in integer time seriesmodels.*International Journal of Forecasting*, *22*, 223-238.

Jung, R. C. and A. R. Tremayne (2006b). Coherent forecasting in integer time series models. *International Journal of Forecasting* **22**(2): 223-238.

Jung, R.C., Ronning, G., Tremayne, A.R. (2005) Estimation in conditional first order autoregression with discrete support. Statistical Papers 46(2), 195–224.

Kaizhi Yu and Huiqiao Wang (2021). A new overdispersed integer-valued moving average model with dependent counting series. Entropy 23,706.

Lawrence A. Klimko and Paul I. Nelson (1978). On Conditional Least Square estimation for stochastic processes. The Anals of statistics, vol.6, No.3, 629-642

Livio, T.; Mamode Khan, N.; Bourgignon, M.; Bakouch, H.S. (2018). An INAR(1) model with Poisson–Lindley innovations. Econ. Bull. 38, 1505–1513.

Marcelo. Bourguignon, J. Rodrigues and M. Santos-Neto (2019). Extended Poisson INAR(1) processes with equidispersion, underdispersion and overdispersion. Journal of Applied Statistics vol.46-issue 1

M. Bourguinon, and K.L.P. Vasconcellos (2015). Braz.J. Probab. Stat. 29, 27-93

M. Mohammadipour and Brunnel University (2009). Intermittent Demand Forecasting with Integer Autoregressive Moving Average Models. PhD Thesis.

McKenzie, E. (1985). Some simple models for discrete variate series. *Water Resources Bulletin* **21**(4): 645-650.

Nasiru M. Olakorede and Olanrewaju S. Olayemi (2023). Integer-Valued first order autoregressive (INAR(1)) model with Negative Binomial Innovation. International Journal of statistics and probability, vol.12, No.6, November 2023. ISSN 19270-7032, E-ISSN 1927-7040

Qi. X., LiQ., Zhu F. (2019). Modeling time series of count with excess zeros and ones based on INAR(1) model with zero and one inflated Poisson innovations. J. comput. Appl. Math. 346:572-590.

Quoreshi A.M.M.S. (2006). Bivariate time series modeling of financial count data. Communication in statistics: Theory and Method, 35, 1343-1358.

Quoreshi A.M.M.S. (2008). A vector integer-valued moving average model for high frequency financial count data. Economics Letter 101, 258-261.

Quoreshi A.M.M.S. (2014). A long memory integer-valued time series model, INARFIMA, for financial application Quantitative finance Vol. 12, 2225-2235.

Schweer S and Weiß CH (2014) Compound Poisson INAR(1) processes: stochastic properties and testing for overdispersion. Computational Statistics and Data Analysis, 77, 267–84.

Sueutel, F. W. and K. van Harn (1979). Discrete analogues of self-decomposability and stability. *Annals of Probability* **7**(5): 893-899.

Vasiliki Christou (2013). Statistical Theory for mixed Poisson time series models. A PhD thesis submitted to the University of Cyprus Nicosia, Cyprus.