# Coherent forecasting for over-dispersed time series of count data

## Raju Maiti and Atanu Biswas

*Indian Statistical Institute*

**Abstract.** In the context of an over-dispersed count time series data on disease incidences, we consider the Geometric integer-valued autoregressive process of order 1 or GINAR(1), which was first introduced by McKenzie (*Adv. Appl. Probab.* **18** (1986) 679–705) as an analogue of continuous AR(1) process with exponential margin (*Adv. Appl. Probab.* **12** (1980) 727–745) on the positive support ($\mathbb{R}^+$). A strong enthusiasm still persists as it is apparent from Ristić et al. (*J. Stat. Plann. Inf.* **139** (2009) 2218–2226). Coherent forecasting of Poisson INAR(1) process due to Al-Osh and Alzaid (*J. Time Ser. Anal.* **8** (1987) 261–275) was studied by Freeland and McCabe (*Int. J. Forecast.* **20** (2004) 427–434). Here, we study the *h*-step ahead forecasting distribution corresponding to GINAR(1) process in details using probability generating function. Large sample distributions of the conditional least squares estimates of the model parameters are derived. Some numerical study is performed to illustrate the theoretical results.

## 1 Introduction

### 1.1 Motivating data examples

1.1.1 *Poliomyelitis data.* Consider a disease incidence data over a period of fourteen years (see Zeger, 1988). The data consisting of 168 observations, lists the monthly number of cases of poliomyelitis reported by the US Centers for Disease Control for the years 1970 to 1983. Figure 1 displays the data with its autocorrelation function (a.c.f.) and partial a.c.f. (p.a.c.f.) plots. From the a.c.f. and p.a.c.f. plots, it seems that the data has a good fit for the AR(1) process. The data seems to be over-dispersed since its marginal mean is 1.333 whereas its variance is 3.505. Also from the observed frequency distribution of the data presented in Table 4, it is clearly seen that the data has some observations with moderate frequency toward the tail part. This fact motivates us to consider some suitable time series process for count over-dispersed data.

1.1.2 *Skin-lesions data.* Consider a data set giving the monthly number of submissions with skin-lesions from a region in New Zealand reported in animal health
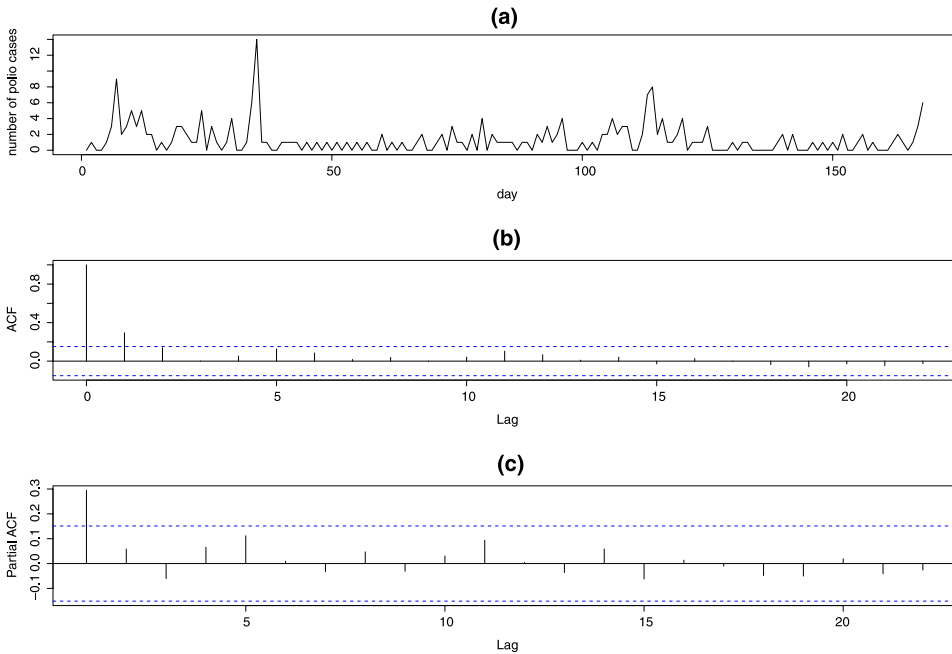
---

**Figure 1**  *Poliomyelitis data with its a.c.f. and p.a.c.f.* (a) *Polio incidence data.* (b) *Acf plot of polio incidence data.* (c) *Pacf plot of polio incidence data.*

laboratories during the period 2003 to 2009 (see Jazi et al., 2012). The marginal mean and the marginal variance are derived as 1.256 and 2.168, respectively. Hence, it is another example of over-dispersed time series data. The data with its a.c.f. and p.a.c.f. plots are displayed in Figure 2. In p.a.c.f. plot, the first lag appears to be significant, and hence an AR(1) process would be a good fit to the data also.

## 1.2  Models

Integer-valued times series process based on binomial thinning operator (due to Steutel and van Harn, 1979) was first introduced by McKenzie (1985). Thereafter, it was well developed and took a general shape called integer-valued ARMA($p$, $q$) process or INARMA($p$, $q$) process. Poisson integer-valued AR(1) process or PINAR(1), introduced by McKenzie (1985) and Al-Osh and Alzaid (1987), has a wide application in modeling integer-valued time series data. However, when an integer-valued time series is over dispersed, Poisson time series model may not be a good choice. McKenzie (1986) proposed the INAR(1) processes with geometric and negative binomial distributions as the marginals, which are discrete analogues of Box-Jenkins AR(1) processes with exponential and gamma margins (Gaver and Lewis, 1980), respectively, on $\mathbb{R}^+$. When the count time series data has some large
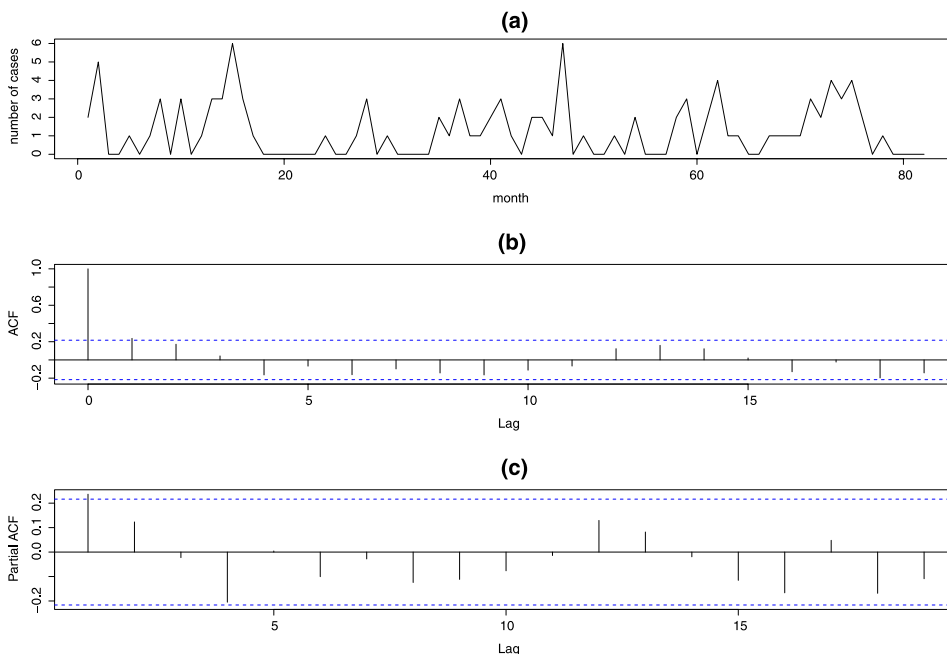
**Figure 2** *Monthly cases of submission of skin lesions data with its a.c.f. and p.a.c.f. (a) Skin lesions data. (b) Acf plot of skin lesions data. (c) Pacf plot of skin lesions data.*

observations in the tail part, the geometric INAR(1) [denoted by GINAR(1)] process and negative binomial INAR(1) [denoted by NBINAR(1)] process may be some suitable alternatives. Ristic et al. (2009) introduced a new GINAR(1) process based on negative binomial thinning operator which added a new way in modeling low count over-dispersed time series. Estimation of the model parameters for the process is also discussed by Ristic et al. (2009).

Coherent forecasting, which is an integral part of count time series analysis, has got very little attention in the context of integer-valued time series analysis. Here, the coherent forecasting means forecasting values are to be integer. So far very few works on coherent forecasting have been done in the count time series context. Freeland and McCabe (2004) possibly be the first authors who used the concept of $h$-step ahead coherent forecasting of $Y_{n+h}$ given the available data $Y_n, Y_{n-1}, \ldots, Y_1$ of the time series process $\{Y_t\}$ by using the median and mode of the $h$-step ahead forecasting distribution. Note that, although the mean of a discrete distribution may not be an integer, median and mode are always so. Moreover, median has optimizing properties like it minimizes the expected absolute error $E\{|Y_{n+h} - \hat{Y}_{n+h}||Y_n\}$ (discussed more details in Section 3). On the other hand, mode has properties like $h$-step ahead forecasting distribution attains its maximum value at it. Silva et al. (2009) carried out Bayesian forecasting in the context of Poisson INAR(1) process. But a major problem with the Poisson INAR(1) pro-

cess is that its mean and variance are equal. However, in practice some data may be over-dispersed. See, for example, the Poliomyelitis data and Skin lesions data discussed in Sections 1.1.1 and 1.1.2 where marginal variances are much higher than the marginal means.

In this article, we consider the geometric INAR(1) process and study its coherent forecasting with some extensive simulation study. In addition, we derive the asymptotic distribution of the conditional least squares (CLS) estimators which is used in finding the asymptotic distribution of the $h$-step ahead forecasting distribution using the delta method. This result helps us in understanding the point that when we incorporate the CLS estimates into the forecasting distribution, the estimated forecasting distribution becomes a consistent estimate of the true forecasting distribution. A detailed simulation study along with the analysis of two data sets are carried out to illustrate the results.

### 1.3 Structure of the paper

In this present article, in Section 2, we discuss the GINAR(1) process, originally introduced by McKenzie (1986), with its conditional distribution. In Section 3, we investigate the $h$-step ahead coherent forecasting corresponding to the GINAR(1) process using the probability generating function (PGF). In practice when a low count time series is over-dispersed due to its fat tail part, the GINAR(1) process appears to be a viable alternative to the PINAR(1) process. Coherent forecasting due to the PINAR(1) process was well studied by Freeland and McCabe (2004). In this present paper, we study the coherent forecasting related to the GINAR(1) process. In Section 4, we use the conditional least squares (CLS) estimates for model parameters and derive the explicit large sample distributions of the CLS estimators of the GINAR(1) process and the $100(1 - \gamma)\%$ confidence intervals. Some simulation studies are presented in Section 5. In Section 6, two over-dispersed disease incidence data sets are analyzed. Section 7 concludes.

## 2 GINAR(1) process

### 2.1 Model

Integer-valued auto-regressive process of order 1 with geometric margin [denoted by GINAR(1)] was introduced by McKenzie (1986), and is defined as

$$Y_t = \alpha \circ Y_{t-1} + \varepsilon_t, \qquad \alpha \in (0, 1), \tag{2.1}$$

where $\{Y_s, s < t\}$ is independent of $\varepsilon_t$. Here, $\alpha \circ Y_{t-1} = \sum_{i=1}^{Y_{t-1}} B_{ti}$, where $B_{ti}$'s are independent Bernoulli($\alpha$) random variables, and consequently $\alpha \circ Y_{t-1}|Y_{t-1} \sim$ Bin($Y_{t-1}, \alpha$). Thus, if the marginal distribution of $Y_t$ is

$$P_{Y_t}(y) = (1 - \theta)\theta^y, \qquad \theta \in (0, 1); y = 0, 1, 2, \dots,$$

the PGF of the distribution is given by

$$\Phi_{Y_t}(s) = E((1-s)^Y) = \frac{1-\theta}{1-\theta+s\theta} = \frac{1}{1+\mu s},$$

where $\mu = \frac{\theta}{1-\theta}$.

If this is to be the marginal distribution of the above process, then the innovation process $\varepsilon_t$ must have the PGF

$$\Phi_{\varepsilon_t}(s) = \frac{1+\mu\alpha s}{1+\mu s} = \alpha + (1-\alpha)\frac{1}{1+\mu s},$$

that is,

$$\varepsilon_t = \begin{cases} 0, & \text{with probability } \alpha, \\ Y_t, & \text{with probability } 1-\alpha. \end{cases}$$

Consequently, the marginal distribution is given by

$$P(\varepsilon_t = i) = \begin{cases} \alpha + (1-\alpha)(1-\theta), & i = 0, \\ (1-\alpha)\theta^i(1-\theta), & i = 1, 2, \ldots \end{cases}$$

and hence $\mu_\varepsilon = (1-\alpha)\mu$ and $\sigma_\varepsilon^2 = (1-\alpha)\frac{\theta}{(1-\theta)^2} + \alpha(1-\alpha)(\frac{\theta}{1-\theta})^2$. Thus, the discrete AR(1) process can be rewritten in this case in the form

$$Y_t = \alpha \circ Y_{t-1} + U_t W_t,$$

where $\{U_t\}$ are independently and identically distributed (i.i.d.) binary random variables with $P(U_t = 0) = \alpha$, independent of $\{W_t\}$, which is a geometric process with mean $\mu$.

## 2.2 Conditional distribution

Under the above set up, conditional mean and conditional variance can be derived as

$$E(Y_t|Y_{t-1}) = \alpha Y_{t-1} + (1-\alpha)\mu$$

and

$$\text{Var}(Y_t|Y_{t-1}) = \alpha(1-\alpha)Y_{t-1} + (1-\alpha)\frac{\theta}{(1-\theta)^2} + \alpha(1-\alpha)\frac{\theta^2}{(1-\theta)^2}.$$

Joint PGF of two consecutive observations in this process is given by

$$\Phi_{Y_t, Y_{t-1}}(u, v) = \frac{1+\alpha\mu u}{(1+\mu u)(1+\alpha\mu u+\mu v)}.$$

Therefore, the conditional distribution of the above process is given by

$$
p(i|j) =
\begin{cases}
(1-\alpha)(1-\theta)\theta^{i-j} \displaystyle\sum_{k=0}^{i} \binom{j}{k}\alpha^k\{(1-\alpha)\theta\}^{j-k} \\
\qquad + \binom{j}{i}\alpha^{i+1}(1-\alpha)^{j-i}; \qquad\qquad i = 0, 1, \ldots, j, \\
(1-\alpha)(1-\theta)\theta^{i-j}\{\alpha + (1-\alpha)\theta\}^j; \qquad i = j+1, j = 2, \ldots,
\end{cases}
\tag{2.2}
$$

where $p(i|j) = P(Y_t = i|Y_{t-1} = j)$. Again we can rewrite (2.2) as

$$
p(i|j) = \sum_{k=0}^{\min(j,i)} \binom{j}{k}\alpha^k(1-\alpha)^{j-k}\{(\alpha + (1-\alpha)(1-\theta))I_{(i-k=0)}
$$
$$
+ (1-\alpha)(1-\theta)\theta^{i-k}I_{(i-k\neq 0)}\},
\tag{2.3}
$$

where $I_{(.)}$ is the indicator function.

## 3 Coherent forecasting

The study of the $h$-step ahead coherent forecasting is essential from the point of low count time series. We know that $h$-step ahead forecasting mean, $E(Y_{n+h}|Y_n)$, is widely used in making the $h$-step ahead forecasting in continuous time series since it has some optimal properties like it minimizes the prediction mean squared error, that is, $E(Y_{n+h}|Y_n) = \arg\min_{c\in\mathbb{R}^+} E\{(Y_{n+h} - c)^2|Y_n\}$. But it may not be an integer for count time series and hence is not coherent. One may of course use some rounding operators to make the forecasting mean an integer, and hence make it coherent. Other ways of making the $h$-step ahead coherent forecasting are used the median and mode predictors of the forecasting distribution $Y_{n+h}|Y_n$ which always belong to the support space, and hence are coherent. In addition median has some optimal properties like it minimizes the prediction mean absolute error $E\{|Y_{n+h} - c||Y_n\}$. However, the predictive probability attains the maximum value at its mode. It is important to note that through some simulated results (discussed in Section 5) we observe that the median and mode predictors have much higher accuracy over mean predictor in making the $h$-step ahead forecasting for low count time series with respect to a descriptive measure defined in (3.6) (see Table 2). One possible explanation of such observation is that the forecasting distributions are skewed to the right and unimodal. Therefore, the mean predictor has lower accuracy compared to other two predictors, namely the median predictor and the mode predictor.

For the process $\{Y_t\}$ in (2.1), the $h$-step ahead conditional mean can easily be figured out (which is done in the next section) without knowing the exact expression of the probability distribution. However, to find the median or mode, one needs to have the exact expression of the probability distribution. Therefore, the problem is to find the $h$-step ahead forecasting distribution of the process given by (2.1).

### 3.1 *h*-step ahead forecasting distribution

Now to find the *h*-step ahead forecasting distribution, we use the following recursive method:

$$
\begin{aligned}
Y_{n+h} &= \alpha \circ Y_{n+h-1} + \varepsilon_{n+h} \\
&= \alpha \circ \{\alpha \circ Y_{n+h-2} + \varepsilon_{n+h-1}\} + \varepsilon_{n+h} \\
&\vdots \\
&= \alpha^h \circ Y_n + \sum_{i=1}^{h} \alpha^{h-i} \circ \varepsilon_{n+i}.
\end{aligned}
$$

Hence, the *h*-step ahead conditional mean is given as

$$
E(Y_{n+h}|Y_n) = \alpha^h Y_n + \frac{1-\alpha^h}{1-\alpha}(1-\alpha)\mu,
$$

and $\lim_{h\to\infty} E(Y_{n+h}|Y_n) = E(Y_n) = \mu$. On the other hand, the *h*-step ahead conditional variance is given by

$$
\mathrm{Var}(Y_{n+h}|Y_n) = \alpha^h(1-\alpha^h)Y_n + \sum_{i=1}^{h} \mathrm{Var}(\alpha^{h-i} \circ \varepsilon_{n+i})
$$

$$
= \alpha^h(1-\alpha^h)Y_n + \sigma_\varepsilon^2 \sum_{i=1}^{h} \alpha^{2(h-i)} + \mu_\varepsilon \sum_{i=1}^{h} \alpha^{h-i}(1-\alpha^{h-i})
$$

$$
= \alpha^h(1-\alpha^h)Y_n + \frac{1-\alpha^{2h}}{1-\alpha^2}(\sigma_\varepsilon^2 - \mu_\varepsilon) + \frac{1-\alpha^h}{1-\alpha}\mu_\varepsilon
$$

and $\lim_{h\to\infty} V(Y_{n+h}|Y_n) = V(Y_n) = \frac{\theta}{(1-\theta)^2}$.

In case of Poisson INAR(1) process, the *h*-step ahead forecasting distribution was derived by Freeland and McCabe (2004) using the properties of binomial thinning operator discussed by Al-Osh and Alzaid (1987). It came out as a convolution of binomial and Poisson distributions. But the method discussed in Freeland and McCabe (2004) cannot be used to derive the *h*-step ahead forecasting distribution for the geometric INAR(1) process. On the other hand, using the property of probability generating function (PGF), which determines the probability distribution uniquely, we can derive the exact forecasting distribution. It is nice to see that the method not only helps us in finding the exact forecasting distribution but also helps us in interpreting the model.

**Theorem 1.** *The conditional PGF of $Y_{n+h}$ given $Y_n$ is given by*

$$
\Phi_{Y_{n+h}|Y_n}(s) = (1 - s\alpha^h)^{Y_n}\left(\frac{1-\theta+\theta\alpha^h s}{1-\theta+\theta s}\right).
$$

**Proof.**

$$\Phi_{Y_{n+h}|Y_n}(s) = E\big((1-s)^{Y_{n+h}}|Y_n\big)$$

$$= E\big(((1-s)^{\alpha^h \circ Y_n + \sum_{i=0}^{h-1} \alpha^i \circ \varepsilon_{n+h-i}})|Y_n\big)$$

$$= E\big((1-s)^{\alpha^h \circ Y_n}|Y_n\big) \prod_{i=0}^{h-1} E\big((1-s)^{\alpha^i \circ \varepsilon_{n+h-i}}\big)$$

since $\alpha^h \circ Y_n|Y_n \sim \mathrm{Bin}(Y_n, \alpha^h)$, $E((1-s)^{\alpha^h \circ Y_n}|Y_n) = (1-s\alpha^h)^{Y_n}$ and

$$\prod_{i=0}^{h-1} E\big((1-s)^{\alpha^i \circ \varepsilon_n}\big) = \prod_{i=0}^{h-1} E_{\varepsilon_n} E\big((1-s)^{\alpha^i \circ \varepsilon_n}|\varepsilon_n\big)$$

$$= \prod_{i=0}^{h-1} E\big(1-s\alpha^i\big)^{\varepsilon_n}$$

$$= \prod_{i=0}^{h-1} \left\{ \alpha + \frac{(1-\alpha)(1-\theta)}{(1-\theta+\theta\alpha^i s)} \right\}$$

$$= \prod_{i=0}^{h-1} \left( \frac{1-\theta+\theta\alpha^{i+1}s}{1-\theta+\theta\alpha^i s} \right)$$

$$= \left( \frac{1-\theta+\theta\alpha^h s}{1-\theta+\theta s} \right).$$

Hence, the result.                                                                                    □

**Corollary 1.** *As $h \to \infty$, the h-step ahead conditional PGF $\Phi_{Y_{n+h}|Y_n}(s)$ converges to the marginal PGF of the process $\{Y_t\}$, that is, $\lim_{h \to \infty} \Phi_{Y_{n+h}|Y_n}(s) = \frac{1-\theta}{1-\theta+\theta s} = \Phi_{Y_n}(s)$. Since the PGF uniquely determines the probability mass function, the h-step ahead forecasting distribution of $Y_{n+h}$ given $Y_n$ converges to the marginal distribution of $Y_n$.*

**Corollary 2.** *From Theorem 1, the h-step ahead prediction distribution of $Y_{n+h}$ given $Y_n$ is nothing but convolution of $\mathrm{Bin}(Y_n, \alpha^h)$ and some random variable $W_{n+h}$ having the PGF of the form $(\frac{1-\theta+\theta\alpha^h s}{1-\theta+\theta s}) = \{\alpha^h + (1-\alpha^h)\frac{1-\theta}{1-\theta+\theta s}\}$. Therefore, $W_{n+h}$ can be interpreted as $W_{n+h} = 0$ with probability $\alpha^h$ and geometric distribution with mean $\mu$ with probability $(1-\alpha^h)$, that is,*

$$P(W_{n+h} = i) = \begin{cases} \alpha^h + (1-\alpha^h)(1-\theta), & i = 0, \\ (1-\alpha^h)\theta^i(1-\theta), & i = 1, 2, \ldots. \end{cases}$$

*In short, the prediction distribution can be presented as*

$$Y_{n+h}|Y_n \overset{d}{=} \mathrm{Bin}(Y_n, \alpha^h) * W_{n+h},$$

*where "∗" is called the convolution between two distributions.*

**Corollary 3.** *Using Theorem* 1 *and Corollary* 2, *the h-step ahead forecasting distribution of $Y_{n+h}$ given $Y_n$ is given by*

$$
p_h(i|j) =
\begin{cases}
(1-\alpha^h)(1-\theta)\theta^{i-j}\displaystyle\sum_{k=0}^{i}\binom{j}{k}\alpha^{hk}\{(1-\alpha^h)\theta\}^{j-k} \\
\qquad + \binom{j}{i}\alpha^{h(i+1)}(1-\alpha^h)^{j-i}, \qquad\qquad i = 0, 1, \ldots, j, \\
(1-\alpha^h)(1-\theta)\theta^{i-j}\{\alpha^h + (1-\alpha^h)\theta\}^j, \qquad i = j+1, j+2, \ldots,
\end{cases}
\tag{3.1}
$$

*where $p_h(i|j) = P(Y_{n+h} = i|Y_n = j)$. An alternate way of representation of the above forecasting distribution* (3.1) *is*

$$
p_h(i|j) = \sum_{k=0}^{\min(j,i)}\binom{j}{k}\alpha^{hk}(1-\alpha^h)^{j-k}\{(\alpha^h + (1-\alpha^h)(1-\theta))I_{(i-k=0)}
$$
$$
+ (1-\alpha^h)(1-\theta)\theta^{i-k}I_{(i-k\neq0)}\}.
\tag{3.2}
$$

*Again, it can be written after reparameterizing $\mu = \frac{\theta}{1-\theta}$ in the following form:*

$$
p_h(i|j) = \sum_{k=0}^{\min(j,i)}\binom{j}{k}\alpha^{hk}(1-\alpha^h)^{j-k}\left\{\left(\alpha^h + \frac{1-\alpha^h}{1+\mu}\right)I_{(i-k=0)}\right.
$$
$$
\left. + (1-\alpha^h)\frac{\mu^{i-k}}{(1+\mu)^{i-k+1}}I_{(i-k\neq0)}\right\}.
\tag{3.3}
$$

### 3.2  $100(1-\gamma)\%$ prediction interval

Since we have observed the true form of the $h$-step ahead predictive probability distribution given by (3.3), we can easily find the $100(1-\gamma)\%$ prediction interval for $Y_{n+h}$. The problem with the standard prediction intervals is that it assumes the predictive probability distribution to be symmetrically distributed. But simulation study shows that the derived predictive distributions in this case are positively skewed and unimodal. See Figure 3 in this regard. Hence, we find the $100(1-\gamma)\%$ highest predictive probability (HPP) interval for $Y_{n+h}$ which is defined as the following.

**Definition.** A $100(1-\gamma)\%$ highest predictive probability (HPP) interval $\mathcal{C}_h = (Y_L, Y_U)$ for $Y_{n+h}$ given $Y_n$ is defined as

$$
\mathcal{C}_h = \{i : p_h(i|y_n) \geq k_\gamma\},
$$

where $k_\gamma$ is the largest number such that

$$
P(Y_L \leq Y_{n+h} \leq Y_U|Y_n = y_n) = \sum_{i=Y_L}^{Y_U} p_h(i|y_n) \geq (1-\gamma).
$$

$$\alpha = 0.6, \mu = 1.5 \qquad\qquad\qquad \alpha = 0.6, \mu = 1$$



$$\alpha = 0.4, \mu = 1.5 \qquad\qquad\qquad \alpha = 0.4, \mu = 1$$
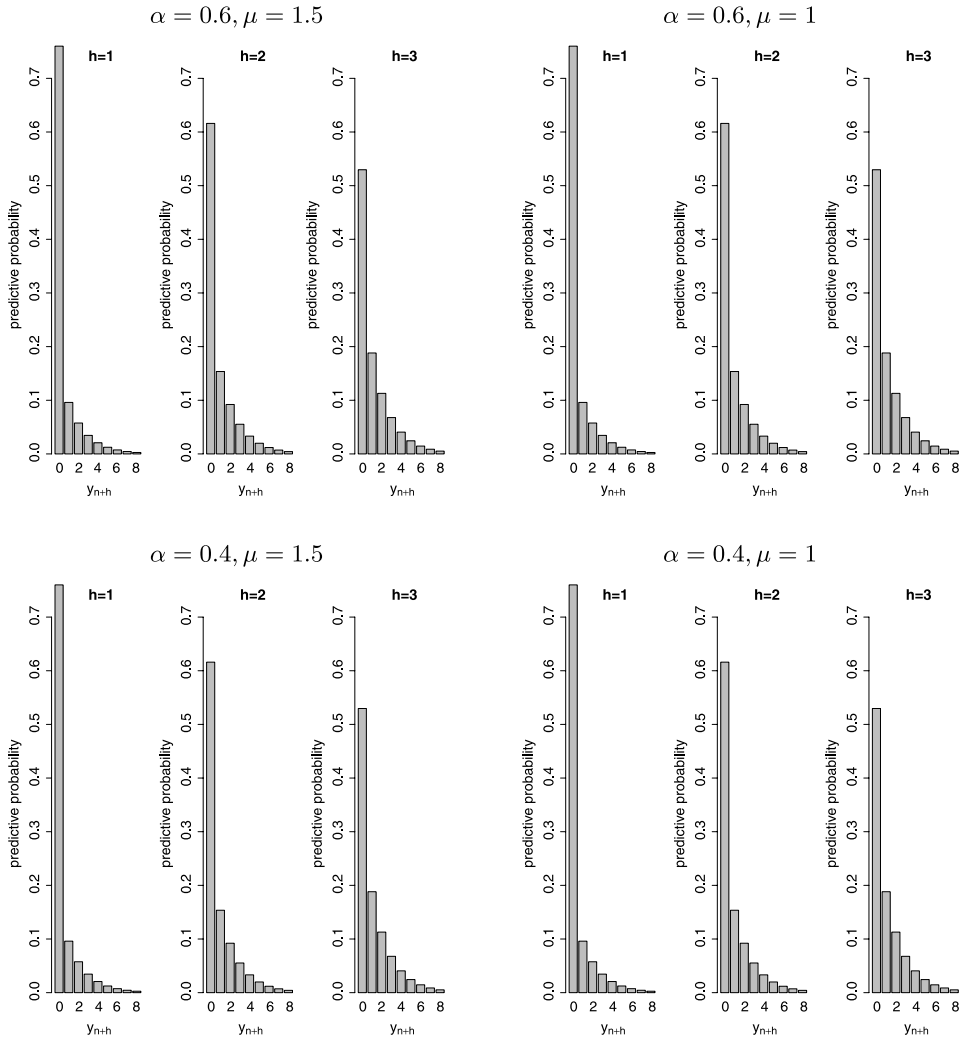


**Figure 3** *h-step ahead predictive probability distribution for GINAR(1) process for different sets of parameters.*

Here, $p_h(i|y_n)$ is defined in (3.3).

## 3.3 Some descriptive measures of forecasting accuracy

Given an observed data set $Y_1, \ldots, Y_n, Y_{n+1}, \ldots, Y_{n+m}$ of size $(n + m)$, we partition the data into two sets. The training set containing first $n$ observations is used to estimate the parameters of the process and based on the rest $m$ observations called the test set, we define the following three descriptive measures of forecast-

ing accuracy. First measure is the prediction root mean squared error [denoted by PRMSE], and defined as

$$\text{PRMSE} = \sqrt{\frac{1}{m} \sum_{h=1}^{m} (Y_{t+h} - \hat{Y}_{t+h}^{\text{me}})^2}, \tag{3.4}$$

where predictions are made through predictive mean. Second measure is the prediction mean absolute error [denoted by PMAE], defined as

$$\text{PMAE} = \frac{1}{m} \sum_{h=1}^{m} |Y_{t+h} - \hat{Y}_{t+h}^{\text{med}}|, \tag{3.5}$$

where predictions are made through predictive median. The third measure is the percentage of true predictions [denoted by PTP], which is defined as

$$\text{PTP} = \frac{1}{m} \sum_{h=1}^{m} I(Y_{t+h} = \hat{Y}_{t+h}) 100\%, \tag{3.6}$$

where $I(\cdot)$ is an indicator function and here predictions can be made through predictive mean, median and mode. This measure gives us the number of true forecasts among 100 forecasts.

All these measures can be used to investigate the forecasting accuracy of the models considered for comparison study. Some simulation results based on these forecasting measures are presented in Section 5 where we take average of these measures obtained in 100 simulations.

## 4 Estimation

### 4.1 Conditional least squares estimation

Conditional least squares estimation is usually used for estimating model parameters in the context of time series process. Freeland and McCabe (2004, 2005) used this approach for Poisson INAR(1) process.

In conditional least squares estimation, we minimize $Q^*(\alpha, \mu) = \sum_{t=2}^{n} \{Y_t - E(Y_t|Y_{t-1})\}^2$ instead of minimizing $Q(\alpha, \mu) = \sum_{t=1}^{n} (Y_t - E(Y_t))^2$ with respect to model parameters $\alpha$ and $\mu$. The conditional mean of the above process is $E(Y_t|Y_{t-1}) = \alpha Y_{t-1} + (1-\alpha)\mu$. After taking the derivative of $Q^*(\alpha, \mu)$, we get the following estimating equation:

$$\begin{pmatrix} \sum_{t=2}^{n} (Y_t - \alpha Y_{t-1} - (1-\alpha)\mu) Y_{t-1} \\ \sum_{t=2}^{n} (Y_t - \alpha Y_{t-1} - (1-\alpha)\mu) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{4.1}$$

Solving the above estimating equation (4.1), we get the parameter estimators as

$$\hat{\alpha}_{\text{cls}} = \frac{(n-1)\sum_{t=2}^{n} Y_{t-1}Y_t - \sum_{t=2}^{n} Y_{t-1}\sum_{t=2}^{n} Y_t}{(n-1)\sum_{t=2}^{n} Y_{t-1}^2 - (\sum_{t=2}^{n} Y_{t-1})^2} \tag{4.2}$$

and

$$\hat{\mu}_{\text{cls}} = (1 - \hat{\alpha}_{\text{cls}})^{-1}\frac{1}{(n-1)}\left(\sum_{t=2}^{n} Y_t - \hat{\alpha}_{\text{cls}}\sum_{t=2}^{n} Y_{t-1}\right). \tag{4.3}$$

**Theorem 2.** *Under the above set up, the conditional least squares estimators have asymptotic normal distribution, that is, $\sqrt{n}\binom{\hat{\alpha}_{\text{cls}}-\alpha}{\hat{\mu}_{\text{cls}}-\mu} \xrightarrow{d} N(0, \Sigma_{\alpha,\mu})$ where $\Sigma_{\alpha,\mu} = (S^{-1})V(S^{-1})^T$ with*

$$S^{-1} = \begin{pmatrix} -\dfrac{1}{\mu(1+\mu)} & \dfrac{1}{1+\mu} \\ 0 & -\dfrac{1}{1-\alpha} \end{pmatrix} \tag{4.4}$$

*and*

$$V = ((v_{ij})), \qquad i, j = 1, 2, \tag{4.5}$$

*where*

$$v_{11} = 2(1-\alpha)(1+3\alpha)\mu^3 + (1-\alpha)(1+6\alpha)\mu^2 + \alpha(1-\alpha)\mu,$$
$$v_{12} = (1-\alpha^2)\mu^3 + (1-\alpha)(1+2\alpha)\mu^2 + \alpha(1-\alpha)\mu,$$
$$v_{22} = (1-\alpha^2)\mu.$$

**Proof.** Al-Osh and Alzaid (1987) showed that INAR(1) process with marginal distribution belonging to discrete self-decomposable (DSD) class is stationary and ergodic. All distributions which are discrete self-decomposable in the sense of Steutel and Van Harn (1979), have marginal distributions for the stationary solution to equation (2.1). This includes many of the commonest distributions on the non-negative integers, for example, Poisson, geometric and negative binomial, but none defined on bounded sets, so that the binomial distribution is excluded and alternative model forms must be sought. The conditions of the asymptotic normality of conditional least squares estimators are due to Klimko and Nelson (1978) for such class. Using this result, Freeland and McCabe (2005) explicitly derived the asymptotic normal distribution of the conditional least squares estimators of Poisson INAR(1) process. The result for GINAR(1) process (2.1) with variance covariance matrix $\Sigma_{\alpha,\mu}$ is similar.                                    □

**Corollary 4.** *Since the process (2.1) is a stationary and ergodic, $\binom{\hat{\alpha}_{\text{cls}}-\hat{\alpha}_{yw}}{\hat{\mu}_{\text{cls}}-\hat{\mu}_{yw}} = o_p(n^{-1/2})$ (see the proof in Freeland and McCabe, 2005) and hence as a consequence YW and CLS estimators have the same asymptotic distribution.*

**Corollary 5.** *Using the delta method and Theorem* 2, *we have the following asymptotic normal distribution of the estimated h-step ahead conditional probability $p_h(y|Y_n, \hat{\alpha}_{cls}, \hat{\mu}_{cls})$ defined in* (3.3):

$$\sqrt{n}\big(p_h(y|Y_n, \hat{\alpha}_{cls}, \hat{\mu}_{cls}) - p_h(y|Y_n, \alpha, \mu)\big) \xrightarrow{d} N\big(0, (\nabla p_h)^T \Sigma_{\alpha,\mu}(\nabla p_h)\big),$$

*where*

$$\nabla p_h = \begin{pmatrix} \dfrac{\partial}{\partial \alpha} p_h \\ \dfrac{\partial}{\partial \mu} p_h \end{pmatrix}$$

*with $p_h$ refers to $p_h(y|Y_n, \alpha, \mu)$ and $\hat{p}_h$ to $p_h(y|Y_n, \hat{\alpha}_{cls}, \hat{\mu}_{cls})$. Consequently the 95% asymptotic confidence interval of the conditional probability $p_h(y|Y_n, \alpha, \mu)$ is given by $(p_h(y|Y_n, \hat{\alpha}_{cls}, \hat{\mu}_{cls}) \mp 1.96 \frac{\hat{\sigma}_h}{\sqrt{n}})$, where $\hat{\sigma}_h^2 = (\nabla \hat{p}_h)^T \hat{\Sigma}_{\alpha,\mu}(\nabla \hat{p}_h)$.*

The requirement of Corollary 5 is that some times median forecasts are not very informative for the entire distribution. Consider the two cases—(a) $P(X = 0) = 0.5 = 1 - P(X = 1)$ and (b) $P(X = 0) = 0.1 = 1 - P(X = 1)$. For both the cases median is 0. But the probability distributions are entirely different on the support $\{0, 1\}$ (see Freeland and McCabe, 2004). Hence, it is suggested that point mass with its confidence interval helps one to visualize the real scenario.

## 5 Simulation study

In this section, we perform a simulation study to investigate the $h$-step ahead coherent forecasting pattern of the GINAR(1) process. For that, we simulated data from the GINAR(1) process with four sets of parameter values—(a) $\alpha = 0.6, \mu = 1.5$, (b) $\alpha = 0.6, \mu = 1$, (c) $\alpha = 0.4, \mu = 1.5$ and (d) $\alpha = 0.4, \mu = 1$. It has been observed that when a count time series data is over-dispersed, that is, the variance is larger than the mean or a time series having some observations toward the tail part resulting in an heavy tailed or fat tailed pattern, the geometric marginal is an alternative to the Poisson marginal. Hence, here we take the PINAR(1) process as a competitor of the GINAR(1) process to visualize the difference in $h$-step ahead forecasting with respect to the measures PRMSE and PMAE defined in (3.4) and (3.5), respectively.

Simulations were done with data of size 700 with 100 trials for the sets of parameter combinations mentioned in the above paragraph. We divided the data into two parts. The training set containing first 400 observations was used for parameter estimation and the test set consisting rest 300 observations was used to find all the descriptive measures. Table 1 presents the average PRMSE and the average PMAE for $h = 1, 2, \ldots, 5$. As one can see, both the measures increase with $h$ which means the simulated results behave as it is expected. That is as one goes

**Table 1**  *Some estimated PRMSE and PMAE values for four sets of parameters for simulated GI-NAR*(1) *process*

| | GINAR(1) | | PINAR(1) | | GINAR(1) | | PINAR(1) | |
|---|---|---|---|---|---|---|---|---|
| $h$-step | PRMSE | PMAE | PRMSE | PMAE | PRMSE | PMAE | PRMSE | PMAE |
| | (a) $\alpha = 0.6, \mu = 1.5$ | | | | (b) $\alpha = 0.6, \mu = 1$ | | | |
| 1 | 1.5763 | 0.8508 | 1.5887 | 1.0011 | 1.1188 | 0.5813 | 1.1229 | 0.5861 |
| 2 | 1.8395 | 1.1378 | 1.8534 | 1.2430 | 1.3101 | 0.7809 | 1.3132 | 0.8758 |
| 3 | 1.9231 | 1.2661 | 1.9328 | 1.3113 | 1.3669 | 0.8778 | 1.3682 | 0.9652 |
| 4 | 1.9544 | 1.3126 | 1.9595 | 1.3399 | 1.3823 | 0.9217 | 1.3831 | 0.9773 |
| 5 | 1.9663 | 1.3184 | 1.9693 | 1.3577 | 1.3874 | 0.9461 | 1.3885 | 0.9780 |
| | (c) $\alpha = 0.4, \mu = 1.5$ | | | | (d) $\alpha = 0.4, \mu = 1$ | | | |
| 1 | 1.7972 | 1.0784 | 1.8111 | 1.2084 | 1.3207 | 0.7691 | 1.3270 | 0.8530 |
| 2 | 1.9373 | 1.2927 | 1.9452 | 1.3119 | 1.4282 | 0.9283 | 1.4314 | 1.0138 |
| 3 | 1.9637 | 1.3058 | 1.9668 | 1.3316 | 1.4458 | 0.9842 | 1.4470 | 1.0151 |
| 4 | 1.9703 | 1.3069 | 1.9716 | 1.3381 | 1.4481 | 1.0035 | 1.4486 | 1.0148 |
| 5 | 1.9726 | 1.3075 | 1.9736 | 1.3425 | 1.4488 | 1.0086 | 1.4492 | 1.0149 |

far ahead from the present, the chance of making true forecasting based on the present observations decreases. On the other hand, both the measures have lower values corresponding to the GINAR(1) process as compared to the PINAR(1) process. Thus, when the actual process is the GINAR(1), it provides better fit than the PINAR(1) process.

Another simulation has been performed to illustrate the advantage of using the median predictor and mode predictor over the mean predictor based on the descriptive measure PTP defined in (3.6). Here also, we repeated the same simulation exercise discussed in the preceding paragraph for all the four sets of parameters, and used the training and test samples of size 400 and 300, respectively, and the results are reported in Table 2 based on a total of 100 trials. We observe that the median and mode predictors outperform the mean predictor in predicting the $h$-step ahead true observations for $h = 1, 2, \ldots, 5$.

In an another simulation study, we used training and test samples of size 200 and 5, respectively, for different sets of parameters. Based on the training samples we observed the $100(1 - \gamma)\%$ HPP intervals. The test sets are used to observe the prediction values through the mean, median and mode predictors over varying $h$ and the results are reported in Table 3 based on the total of 500 replications. As one can see, the length of the HPP intervals increases over varying $h$ which indicates that to retain the same percentage of true forecasting values over $h$, the interval lengths have to be increased.

**Table 2** *Estimated PTP values through mean, median and mode for four sets of parameters for simulated GINAR*(1) *process*

| h-step | PTP | | | PTP | | |
|---|---|---|---|---|---|---|
| | Mean | Median | Mode | Mean | Median | Mode |
| | (a) $\alpha = 0.6, \mu = 1.5$ | | | (b) $\alpha = 0.6, \mu = 1$ | | |
| 1 | 28.35 | 53.77 | 53.76 | 59.48 | 62.25 | 62.39 |
| 2 | 26.76 | 52.45 | 46.52 | 29.02 | 61.40 | 55.20 |
| 3 | 24.65 | 48.09 | 41.78 | 26.87 | 60.16 | 51.43 |
| 4 | 22.42 | 28.08 | 39.73 | 26.15 | 58.65 | 50.05 |
| 5 | 20.55 | 23.97 | 39.24 | 25.93 | 55.81 | 49.72 |
| | (c) $\alpha = 0.4, \mu = 1.5$ | | | (d) $\alpha = 0.4, \mu = 1$ | | |
| 1 | 22.16 | 44.72 | 46.19 | 25.15 | 52.97 | 53.96 |
| 2 | 21.47 | 32.77 | 42.49 | 25.05 | 52.93 | 51.35 |
| 3 | 20.25 | 23.94 | 40.46 | 24.76 | 51.08 | 50.26 |
| 4 | 19.43 | 23.92 | 40.25 | 24.73 | 44.75 | 50.21 |
| 5 | 19.06 | 23.91 | 40.26 | 24.72 | 40.99 | 50.18 |

**Table 3** $100(1 - \gamma)\%$ *HPP intervals for the GINAR*(1) *simulated data with average forecasting values for different sets of parameters where* $\gamma = 0.2$

| h | $(Y_L, Y_U)$ | Mean | Median | Mode | $(Y_L, Y_U)$ | Mean | Median | Mode |
|---|---|---|---|---|---|---|---|---|
| | $\alpha = 0.4, \mu = 1$ | | | | $\alpha = 0.4, \mu = 1.5$ | | | |
| 1 | (0.01, 1.73) | 0.966 | 0.552 | 0.342 | (0.04, 2.63) | 1.501 | 0.867 | 0.621 |
| 2 | (0, 2.04) | 0.990 | 0.423 | 0.068 | (0, 2.86) | 1.486 | 0.881 | 0.171 |
| 3 | (0, 2.05) | 1.012 | 0.438 | 0 | (0, 2.89) | 1.485 | 0.982 | 0.033 |
| 4 | (0, 2.05) | 1.014 | 0.454 | 0 | (0, 2.91) | 1.487 | 0.988 | 0.010 |
| 5 | (0, 2.05) | 1.023 | 0.462 | 0 | (0, 2.92) | 1.489 | 0.984 | 0 |
| | $\alpha = 0.6, \mu = 1$ | | | | $\alpha = 0.6, \mu = 1.5$ | | | |
| 1 | (0.05, 1.42) | 0.991 | 0.744 | 0.724 | (0.19, 2.27) | 1.588 | 1.160 | 1.100 |
| 2 | (0, 1.80) | 0.987 | 0.594 | 0.306 | (0.03, 2.76) | 1.547 | 0.912 | 0.566 |
| 3 | (0, 1.93) | 0.987 | 0.476 | 0.122 | (0, 2.87) | 1.525 | 0.834 | 0.302 |
| 4 | (0, 1.99) | 0.988 | 0.432 | 0.040 | (0, 2.90) | 1.514 | 0.948 | 0.138 |
| 5 | (0, 1.99) | 0.989 | 0.428 | 0.010 | (0, 2.92) | 1.508 | 0.952 | 0.060 |

# 6 Data analysis

## 6.1 Poliomyelitis data

Consider the poliomyelitis incidence data described in Section 1. We fitted the above two processes, PINAR(1) and GINAR(1), to the poliomyelitis data set. The estimates of the parameters of the GINAR(1) process by CLS method are $\hat{\alpha}_{cls} = 0.294$ and $\hat{\theta}_{cls} = 0.587$, and hence $\hat{\mu}_{cls} = \frac{\hat{\theta}_{cls}}{1 - \hat{\theta}_{cls}} = 1.333$ with a 95% asymp-

**Table 4**  *Frequency distribution of observed and expected number of the monthly cases of Poliomyelitis during the year* 1970 *to* 1983

| Number of cases | Observed | GINAR(1) | PINAR(1) |
|---|---|---|---|
| 0 | 64 | 69 | 41 |
| 1 | 55 | 40 | 56 |
| 2 | 22 | 28 | 47 |
| 3 | 12 | 14 | 17 |
| 4 | 6 | 7 | 5 |
| ≥ 5 | 9 | 10 | 3 |
| AIC | | 471.71 | 504.43 |

totic confidence intervals $(0.269, 0.333)$ and $(1.095, 1.572)$ for $\alpha$ and $\mu$, respectively (using Theorem 2). Since the confidence interval corresponding to $\alpha$ does not include "0," we may conclude that there is significant dependence in the data. After fitting the two models to the data set, we presented the observed frequency with the expected frequency determined by the above fitted models in Table 4. As one can see, the expected frequencies by the GINAR(1) process are much closer to the observed frequency compared to the PINAR(1) process, not only in the tail part but in lower part as well. Also, the AIC corresponding to the GINAR(1) process for the data set is 471.71, which is much lower compared to that of the PINAR(1) process, which is 504.43.

In this part, we divided the data into two sets for carrying out a $h$-step ahead forecasting analysis based on the three measures defined in (3.4), (3.5) and (3.6). The training set consisting first 138 observations is used for fitting the GINAR(1) and PINAR(1) processes. The test set consisting rest 30 observations is used to determine all the three measures. The results are reported in Table 5. There was no visible difference in the performance of the processes considered for comparison based on the PRMSE and PMAE measures. Here, we have reported only the PTP based on the mean and mode predictors. As one can see, there is a remarkable difference in the performance of the GINAR(1) process over PINAR(1) process in capturing the true forecasting values based on the mode predictor. On the other hand, the PTP values for the GINAR(1) process through mean predictor for varying $h$ are reported which clearly indicates that the mode predictor preforms much better than the mean predictor. In Table 6, the forecasting values for various values of $h$ by mean, median and mode predictors are reported with the HPP intervals. A residual plot is presented in Figure 4.

## 6.2 Skin-lesions data

Consider the skin-lesions data. In Table 7, we presented the observed frequency with the expected frequency determined by the two fitted models—PINAR(1) and GINAR(1). Here also, the expected frequencies by the GINAR(1) method are

**Table 5** *Estimated parameters, AIC and the PTP values for the Poliomyelitis data*

| Model | Estimated values | AIC | $h$-step | PTP[mean] | PTP[mode] |
|---|---|---|---|---|---|
| PINAR(1) | $\hat{\alpha}_{\text{cls}} = 0.294$ | 504.43 | 1 | 30.00 | 20.00 |
| | $\hat{\lambda}_{\text{cls}} = 0.969$ | | 2 | 31.58 | 31.58 |
| | | | 3 | 33.33 | 33.33 |
| GINAR(1) | $\hat{\alpha}_{\text{cls}} = 0.294$ | 471.71 | 1 | 30.00 | 40.00 |
| | $\hat{\mu}_{\text{cls}} = 1.333$ | | 2 | 31.58 | 47.37 |
| | | | 3 | 33.33 | 50.00 |

**Table 6** $100(1 - \gamma)\%$ *prediction intervals with actual values and forecasting values through mean, median and mode for the Poliomyelitis data where* $\gamma = 0.2$

| | | GINAR(1) | | | | PINAR(1) | | | |
|---|---|---|---|---|---|---|---|---|---|
| $h$ | Actual | $(Y_L, Y_U)$ | Mean | Median | Mode | $(Y_L, Y_U)$ | Mean | Median | Mode |
| 1 | 1 | (0, 3) | 1.51 | 1 | 1 | (0, 3) | 1.51 | 1 | 1 |
| 2 | 0 | (0, 3) | 1.37 | 1 | 1 | (0, 2) | 1.36 | 1 | 1 |
| 3 | 1 | (0, 3) | 1.32 | 1 | 1 | (0, 2) | 1.32 | 1 | 1 |
| 4 | 3 | (0, 3) | 1.31 | 1 | 1 | (0, 2) | 1.31 | 1 | 1 |
| 5 | 6 | (0, 3) | 1.30 | 1 | 1 | (0, 2) | 1.30 | 1 | 1 |
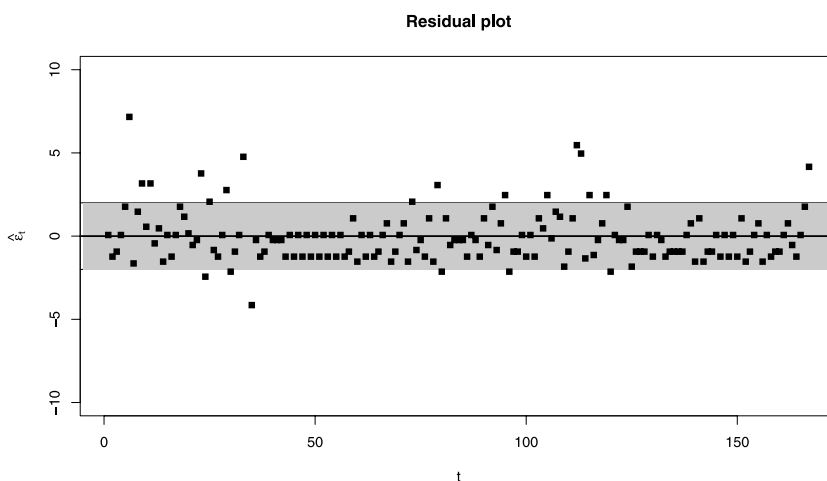


**Figure 4** *Residual plot for mean predictor for the poliomyelitis data.*

much closer to the observed frequency compared to the PINAR(1) model and the AIC corresponding to the GINAR(1) model is much lower compared to that of the PINAR(1) model. The estimates are given in Table 8. As one can see, the median

**Table 7** *Frequency distribution of observed and expected number of the monthly cases of submission of skin lesions in New Zealand during the year* 2003 *to* 2009

| Number of cases | Observed | GINAR(1) | PINAR(1) |
|---|---|---|---|
| 0 | 34 | 37.19 | 21.19 |
| 1 | 21 | 20.46 | 25.32 |
| 2 | 10 | 11.04 | 15.21 |
| 3 | 11 | 6.11 | 5.97 |
| 4 | 3 | 3.47 | 1.85 |
| $\geq 5$ | 3 | 3.01 | 0.44 |
| AIC | | 214.77 | 223.04 |

**Table 8** *Estimated parameters*, *AIC and the PTP values for the skin-lesions data*

| Model | Estimated values | AIC | $h$-step | PMAE | PTP[median] | PTP[mode] |
|---|---|---|---|---|---|---|
| PINAR(1) | $\hat{\alpha}_{\mathrm{cls}} = 0.1542$ | 223.04 | 1 | 1.3076 | 15.38 | 15.38 |
| | $\hat{\lambda}_{\mathrm{cls}} = 1.0175$ | | 2 | 1.4167 | 16.66 | 16.66 |
| | | | 3 | 1.3636 | 18.18 | 18.18 |
| GINAR(1) | $\hat{\alpha}_{\mathrm{cls}} = 0.1542$ | 214.77 | 1 | 1.1538 | 30.77 | 38.46 |
| | $\hat{\mu}_{\mathrm{cls}} = 1.2031$ | | 2 | 1.4167 | 16.67 | 33.33 |
| | | | 3 | 1.3636 | 18.18 | 27.27 |

**Table 9** $100(1 - \gamma)\%$ *prediction intervals with actual values and forecasting values through mean*, *median and mode for the skin-lesions data where* $\gamma = 0.2$

| | | GINAR(1) | | | | PINAR(1) | | | |
|---|---|---|---|---|---|---|---|---|---|
| $h$ | Actual | $(Y_L, Y_U)$ | Mean | Median | Mode | $(Y_L, Y_U)$ | Mean | Median | Mode |
| 1 | 0 | (0, 2) | 1.04 | 0 | 0 | (0, 2) | 1.04 | 1 | 1 |
| 2 | 1 | (0, 3) | 1.26 | 1 | 0 | (0, 2) | 1.26 | 1 | 1 |
| 3 | 0 | (0, 3) | 1.31 | 1 | 0 | (0, 2) | 1.31 | 1 | 1 |
| 4 | 0 | (0, 3) | 1.32 | 1 | 0 | (0, 2) | 1.32 | 1 | 1 |
| 5 | 0 | (0, 3) | 1.32 | 1 | 0 | (0, 2) | 1.32 | 1 | 1 |

and mode predictors for the GINAR(1) process give much higher accuracy in predicting the actual forecasting values compared to that of the PINAR(1) process. In Table 9, we reported the forecasting values for various $h$ by mean, median and mode predictors with the HPP intervals. A residual plot for the data is given in Figure 5.
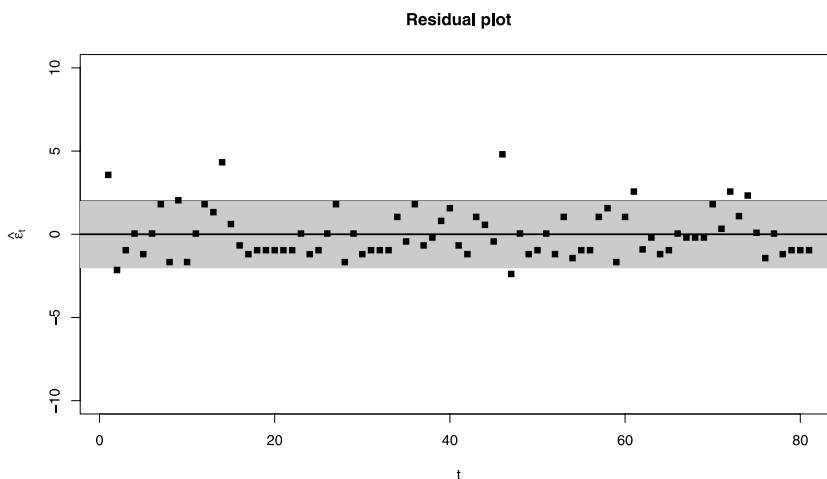
**Residual plot**



**Figure 5** *Residual plot for the mean predictor for the skin lesions data.*

## 7 Conclusion

In this article, we developed and investigated the $h$-step ahead coherent forecasting for the GINAR(1) process as an alternate of the PINAR(1) process. There are many practical incidents where data are over-dispersed. In those cases, it is not wise to apply such Poisson time series processes. It is observed that in application to the over-dispersed independent and identically distributed count data, geometric, negative binomial and some other heavy tailed distributions are widely used. The same idea was incorporated by McKenzie (1986) toward modeling low count time series.

Through some simulation studies and two practical data sets on disease incidences, we established that a detailed study of the $h$-step ahead forecasting for over-dispersed time series processes like the GINAR(1) process is required for the sake of completeness. The coherent forecasting in that context might be an useful addition in that direction.

## Acknowledgments

## References

Al-Osh, M. A. and Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *Journal of Time Series Analysis* **8**, 261–275. MR0903755

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

Freeland, R. K. and McCabe, B. (2005). Asymptotic properties of CLS estimators in the Poisson AR(1) model. *Statistics & Probability Letters* **73**, 147–153. MR2159250

Gaver, D. P. and Lewis, P. A. W. (1980). First-order autoregressive gamma sequences and point processes. *Advances in Applied Probability* **12**, 727–745. MR0578846

Jazi, M. A., Jones, G. and Lai, C.-D. (2012). First-order integer valued AR processes with zero inflated poisson innovations. *Journal of Time Series Analysis* **33**, 954–963. MR2991911

Klimko, L. A. and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *The Annals of Statistics* **6**, 629–642. MR0494770

McKenzie, E. (1985). Some simple models for discrete variate time series. *JAWRA Journal of the American Water Resources Association* **21**, 645–650.

McKenzie, E. (1986). Autoregressive moving-average processes with negative-binomial and geometric marginal distributions. *Advances in Applied Probability* **18**, 679–705. MR0857325

Ristić, M. M., Bakouch, H. S. and Nastić, A. S. (2009). A new geometric first-order integer-valued autoregressive (NGINAR (1)) process. *Journal of Statistical Planning and Inference* **139**, 2218–2226. MR2507983

Silva, N., Pereira, I. and Silva, M. E. (2009). Forecasting in INAR(1) model. *REVSTAT—Statistical Journal* **7**, 119–134. MR2501374

Steutel, F. W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Annals of Probability* **7**, 893–899. MR0542141

Zeger, S. L. (1988). A regression model for time series of counts. *Biometrika* **75**, 621–629. MR0995107

Applied Statistics Unit
Indian Statistical Institute
203 BT Road
Kolkata 700108
India
E-mail: rajumaiti@gmail.com
            atanu@isical.ac.in