JIRSS (2020)
Vol. 19, No. 02, pp 145-173
DOI:10.52547/jirss.19.2.145

Poisson-Lindley INAR(1) Processes: Some Estimation and Forecasting Methods

Roya Nasirzadeh [1]**, and Atefeh Zamani** [2]

[1] Department of Statistics, Faculty of Science, Fasa University, Fasa, IRAN.
[2] Department of Statistics, Faculty of Science, Shiraz University, Shiraz, IRAN.

Received: 19/02/2020, Revision received: 24/09/2020, Published online: 03/12/2020

Abstract. This paper focuses on different methods of estimation and forecasting in first-order integer-valued autoregressive processes with Poisson-Lindley (PLINAR(1)) marginal distribution. For this purpose, the parameters of the model are estimated using Whittle, maximum empirical likelihood and sieve bootstrap methods. Moreover, Bayesian and sieve bootstrap forecasting methods are proposed and predicted value for *h*-step ahead of the series is obtained. Some simulations and a real data analysis are applied to compare the presented estimations and the prediction methods.

Keywords. Autoregressive, Estimation, Integer-Valued Time Series, Poisson-Lindley Distribution, Prediction.

MSC: 62G05, 62M10.

1 Introduction

Time series analysis is one of the most important statistical techniques when dealing with the study of real data sets. Although this method is applied for decades, analyzing

Corresponding Author: Roya Nasirzadeh (nasirzadeh.roya@fasau.ac.ir)
Atefeh Zamani (zamania@shirazu.ac.ir)

time series of counts or integer-valued time series has recently attracted the attention of researchers. These time series are concerned with counting certain objects or events at specific times and can be studied from two aspects. In one aspect, if an integer-valued time series has a big enough range, it can be approximated by a standard continuous model and, in the other aspect, it is necessary to fit an integer-valued model to the series.

Integer-valued time series has been widely used in different studies, such as the number of daily transactions in the stock market, the annual counts of hurricanes, the number of rainy days in successive weeks, the number of patients treated each day in an emergency department and the daily counts of swine flu cases, see Fokianos and Kedem (2003), McKenzie (2003), and Brännäs and Shahiduzzaman Quoreshi (2010), Moriña et al. (2011), Livsey et al. (2018), etc. The main characteristic of integer-valued time series is their integer-valued structure and their correlation over time. Therefore, they cannot be well approximated by continuous variables.

Since the late 1970s, various models have been introduced to model count time series with a preset marginal distribution. Many of these models are thinning operator-based and resemble the autoregressive moving-average (ARMA) models. Among these thinning operators, the binomial thinning operator, introduced by Steutel and Van Harn (1979), is very popular. Al-Osh and Alzaid (1987) were the first researchers that proposed an INAR model based on this thinning operator, called first-order integer-valued autoregressive (INAR(1)) process. Du and Li (1991) presented a pth-order INAR model. Zheng et al. (2007) proposed a first-order random coefficient INAR process. Ristić (2009) applied the geometric distribution and defined a new INAR(1) process based on the negative binomial thinning operator. Bakouch and Ristić (2010) presented an INAR(1) process with zero truncated Poisson marginal distribution. Using signed generalized power series thinning operator, Zhang et al. (2010) introduced the pth-order INAR process. In 2016, introducing a random environment in INAR process by Nastić et al. (2016) results in a significant breakthrough in INAR modeling. Mohammadpour et al. (2018) proposed a first-order INAR model with Poisson-Lindley marginal distribution. For more related research, we refer readers to Weiß (2008) and Scotto et al. (2015).

In time series analysis, estimation and forecasting are two critical topics. Various researchers apply different methods for parameter estimation in INAR models. Among these methods, the Yule-Walker, maximum likelihood, conditional least squares and quasi-likelihood estimates are more popular, see for example, Zheng et al. (2006) However, prediction methods in integer-valued time series are less developed. For instance, McCabe et al. (2011) used an efficient probabilistic forecast method. Maiti and Biswas (2015) and Awale et al. (2017) considered the coherent forecasting first-order INAR process with geometric marginals. Maiti et al. (2016) studied the forecasting for count time series using Box-Jenkins's AR(p) model.

The Poisson-Lindley distribution, which belongs to the compound Poisson family, is unimodal, overdispersed, and infinitely divisible, Ghitany and Al-Mutairi (2009). Besides, this distribution has an increasing hazard rate and can be viewed as a mixture of a geometric and a negative binomial distribution. Moreover, it has smaller skewness and kurtosis than the negative binomial distribution, Ghitany and Al-Mutairi (2009). The theoretical advances of this distribution, along with its performance in real data analysis, are the motivations of Mohammadpour et al. (2018) to introduce and established several statistical properties the Poisson-Lindley INAR(1) process, PLIAR(1) in abbreviation. Besides, they considered conditional least square, Yule-Walker and maximum likelihood methods for estimating the unknown parameters of this model. Moreover, in 2018, Wang and Zhang proposed the Quasi-Likelihood estimation for the parameters of PLINAR(1) process. In this paper, we are going to study Whittle, maximum empirical likelihood and sieve bootstrap methods for parameter estimation and Bayesian and sieve bootstrap forecasting methods for predicting PLINAR(1) process.

The remainder of this article is organized as follows. Section 2 is devoted to preliminary notations and definitions. Besides, we present some statistical properties of the PLINAR(1) process, which will be used throughout the paper. In Section 3, we propose the Whittle, maximum empirical likelihood and sieve bootstrap estimation methods. The Bayesian and sieve bootstrap prediction methods for PLINAR(1) process are studied in Section 4. The simulation results and an application are presented in Sections 5 and 6.

2 Preliminaries and some Basic Properties of PLINAR(1)

In this section, we are going to have a quick review on the definition of PLINAR(1) model along with some of its properties, which are presented in Mohammadpour et al. (2018) and Wang and Zhang (2018) and will be applied in the rest of the paper.

Consider the binomial thinning operator "$\circ$" defined as $\alpha \circ X := \sum_{j=1}^{X} B_j$, where B_j is a sequence of i.i.d. Bernoulli random variables with $P\left(B_j = 1\right) = 1 - P\left(B_j = 0\right) = \alpha$, Steutel and Van Harn (1979). Let $\{X_t\}_{t\in\mathbb{Z}}$ be a sequence of scalar time series satisfying the INAR(1) equation

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \qquad t \geq 1. \tag{2.1}$$

If, additionally, X_t is stationary with Poisson-Lindley distribution, PL(θ), with probability mass function (PMF)

$$f(x, \theta) = \frac{\theta^2 (x + \theta + 2)}{(1 + \theta)^{x+3}}, \qquad x = 0, 1, \ldots, \quad \theta > 0, \tag{2.2}$$

the resulting process is called PLINAR(1). Mohammadpour et al. (2018) demonstrated that ϵ_t, which is a non-negative integer-valued random variable, possesses the PMF $f_\epsilon(x) = \alpha h(x) + (1 - \alpha) g(x)$, where $h(x)$ is a degenerate distribution at zero, and $g(x)$ is a PMF presented by

$$\begin{aligned} g(x) = {} & \frac{\theta^2 (1-\alpha)^2 + \theta\left(1-\alpha^2\right) + 2\alpha}{(\theta(1-\alpha)+1)^2} \frac{\theta}{\theta+1}\left(1 - \frac{\theta}{1+\theta}\right)^x \\ & + \frac{(1-\alpha)(x+1)}{\theta(1-\alpha)+1}\left(\frac{\theta}{1+\theta}\right)^2\left(1 - \frac{\theta}{1+\theta}\right)^x \\ & - \frac{\alpha}{(\theta(1-\alpha)+1)^2} \frac{1+\theta}{1+\theta+\alpha}\left(1 - \frac{\theta+1}{1+\theta+\alpha}\right)^x. \end{aligned} \tag{2.3}$$

Here, ϵ_t is independent of X_m and its associated counting series for all $m \leq t$. Besides, they provide two other representations for PLINAR(1) process:

$$X_t = \begin{cases} \alpha \circ X_{t-1} & w.\ p.\ \alpha \\ \alpha \circ X_{t-1} + \epsilon_t & w.\ p.\ 1 - \alpha, \end{cases} \tag{2.4}$$

or equivalently,

$$X_t = \alpha \circ X_{t-1} + I_t H_t, \tag{2.5}$$

where $P(I_t = 0) = 1 - P(I_t = 1) = \alpha$, H_t has the PMF g, presented in equation (2.3) and X_{t-k} is independent of $I_t H_t$, for $k \geq 1$.

It can easily be shown that, for $k \geq 0$, the autocovariance and autocorrelation functions of the PLINAR(1) process are obtained as $\gamma(k) = Cov(X_t, X_{t-k}) = \alpha^k \gamma_0$, and $\rho(k) = \alpha^k$, respectively. Consequently, the spectral density function of the PLINAR(1) process is formulated as

$$
\begin{aligned}
f(\lambda) &= \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} Cov\left(X_t, X_{t-k}\right) \exp\left(-i\lambda k\right) \\
&= \frac{\left(\theta^3 + 4\theta^2 + 6\theta + 2\right)\left(1-\alpha^2\right)}{2\pi\theta^2\left(\theta+1\right)^2\left(1+\alpha^2 - 2\alpha\cos\left(\lambda\right)\right)},
\end{aligned}
\tag{2.6}
$$

where $\lambda \in (-\pi, \pi]$, Mohammadpour et al. (2018).

The PLINAR(1) process possesses the Markov property and, consequently, for $\theta \geq 1$, if X_0 is PL(θ), then the process $\{X_t\}$ is Poisson-Lindley for every $t \geq 1$ and is strict stationarity. The 1-step ahead transition probabilities for this process can be calculated as

$$
f_{X_{t+1}|X_t}\left(i|j\right) = P\left(X_{t+1} = j|X_t = i\right) = \sum_{k=0}^{\min(i,j)} \binom{i}{k} \alpha^k \left(1-\alpha\right)^{i-k} P\left(I_t H_t = j-k\right), \tag{2.7}
$$

Mohammadpour et al. (2018), and similarly, the h-step ahead transition probabilities can be obtained as

$$
\begin{aligned}
f_{X_{t+h}|X_t}\left(i|j\right) &= P\left(X_{t+h} = i|X_t = j\right) \\
&= \sum_{k=0}^{\min(i,j)} \binom{j}{k} \alpha^{hk}\left(1-\alpha^h\right)^{j-k} P\left(W_{t+h} = i-k\right),
\end{aligned}
\tag{2.8}
$$

where

$$
P\left(W_{t+h} = i\right) = \begin{cases} \alpha^h + \left(1-\alpha^h\right)\left(A_h \frac{\theta}{1+\theta} + B_h\left(\frac{\theta}{1+\theta}\right)^2 + C_h \frac{1+\theta}{1+\theta+\alpha^h}\right), & i = 0 \\ \left(1-\alpha^h\right)\Big[A_h \frac{\theta}{1+\theta}\left(\frac{1}{1+\theta}\right)^i + B_h(i+1)\left(\frac{\theta}{1+\theta}\right)^2\left(\frac{1}{1+\theta}\right)^i & \\ \qquad + C_h \frac{1+\theta}{1+\theta+\alpha^h}\left(\frac{\alpha^h}{1+\theta+\alpha^h}\right)^i\Big], & i = 1, 2, \ldots, \end{cases} \tag{2.9}
$$

and

$$A_h = \frac{\theta^2 \left(1-\alpha^h\right)^2 + \theta\left(1-\alpha^{2h}\right) + 2\alpha^h}{\left(\theta\left(1-\alpha^h\right)+1\right)^2},$$
$$B_h = \frac{1-\alpha^h}{\theta\left(1-\alpha^h\right)+1},$$
$$C_h = \frac{-\alpha^h}{\left(\theta\left(1-\alpha^h\right)+1\right)^2},$$

Wang and Zhang (2018). Additionally, using the induction method, the k-step ahead conditional mean of PLINAR(1) process is obtained as

$$E\left(X_{t+k}|X_t = x\right) = \alpha^k x + \left(1-\alpha^k\right)\frac{\theta+2}{\theta\left(\theta+1\right)}, \tag{2.10}$$

Mohammadpour et al. (2018). In the following sections, we will apply these definitions and properties to provide different estimators and prediction methods in PLINAR(1) process.

3 Estimation Methods

In time series analysis, parameter estimation is a key component in developing the appropriate model for a set of data. To estimate the unknown parameters from the realization $x_1, \cdots, x_n$ of the PLINAR(1) process, different methods, such as conditional least square, Yule-Walker, maximum likelihood (Mohammadpour et al. (2018)) and Quasi-Likelihood estimation methods (Wang and Zhang (2018)) are proposed. In this section, we will present the Whittle, empirical likelihood ratio (ELR) and sieve bootstrap (SB) estimators for the unknown parameters of PLINAR(1) process.

3.1 Whittle Estimation

Generally speaking, the estimation of the parameters in finite-parameter time series models is based on the time-domain approach. However, in 1953, Whittle proposed an estimator based on the frequency domain to estimate the parameters in Gaussian processes. The motivation of introducing Whittle's estimator is that sometimes it is easier to obtain the spectral density function of a model than the exact likelihood. This estimator, which is investigated by several authors (.Walker (1964); Hannan (1973);

Rice (1979), is used in different applications and situations (Fox and Taqqu (1986); Sesay and Rao (1992); Rao and Chandler (1996)).

Let $\boldsymbol{\omega}$ be the vector of the parameters of the considered model and $f(\lambda, \boldsymbol{\omega})$ denotes the non-normalized spectral density function of the process. The periodogram, $I_N(\lambda)$, of the considered process is defined as:

$$\begin{aligned} I_N(\lambda) &= (2\pi N)^{-1} \left| \sum_{t=1}^{N} X_t \exp(i\lambda t) \right|^2 \\ &= (2\pi N)^{-1} \left\{ \left(\sum_{t=1}^{N} X_t \cos(\lambda t) \right)^2 + \left(\sum_{t=1}^{N} X_t \sin(\lambda t) \right)^2 \right\}^2 . \end{aligned}$$

Whittle (1953) demonstrated that, for Gaussian time series, the maximization of the log likelihood function of the sample is asymptotically equivalent to minimization of Whittle's criterion given by

$$\hat{L}_N(\boldsymbol{\omega}) = \frac{N}{4\pi} \int_{-\pi}^{\pi} \left\{ \log f(\lambda, \boldsymbol{\omega}) + \frac{I_N(\lambda)}{f(\lambda, \boldsymbol{\omega})} \right\} d\lambda. \tag{3.1}$$

Consider $\hat{\boldsymbol{\omega}}$ to represent the estimator obtained by minimizing (3.1). Under the Gaussianity assumption, this estimator is weakly consistent and has an asymptotic normal distribution, Whittle (1953) and Walker (1964). In non-Gaussian time series, the Whittle's criterion can be applied and the obtained estimator is still consistent and has an asymptotic normal distribution, Rao and Chandler (1996) and Rice (1979). However, difficulty in obtaining expressions for the asymptotic variance of $(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_0)$ makes these estimators impractical.

To estimate the parameters using the minimization of Whittle criterion, (3.1) is replaced by

$$\hat{l}_N(\boldsymbol{\omega}) = \frac{1}{N} \sum_{j=1}^{[N/2]} \left\{ \log f\left(\lambda_j, \boldsymbol{\omega}\right) + \frac{I_N\left(\lambda_j\right)}{f\left(\lambda_j, \boldsymbol{\omega}\right)} \right\}, \tag{3.2}$$

where $f\left(\lambda_j, \boldsymbol{\omega}\right)$ is the spectral density function at the frequency point $\lambda_j = 2\pi j/N$, and $\boldsymbol{\omega}$ is the vector of the parameters. For PLINAR(1) process, considering the spectral density function, presented in (2.6), and $\boldsymbol{\omega} = (\alpha,\ \theta)^T$, we have

$$\hat{l}_N(\boldsymbol{\omega}) \propto \frac{[N/2]}{N}\left\{\log\left(\theta^3+4\theta^2+6\theta+2\right)+\log\left(1-\alpha^2\right)-2\log\left(\theta\left(\theta+1\right)\right)\right\}$$
$$+\frac{2\pi\theta^2(\theta+1)^2}{N(\theta^3+4\theta^2+6\theta+2)(1-\alpha^2)}\sum_{j=1}^{[N/2]}\left(1+\alpha^2-2\alpha\cos\left(\lambda_j\right)\right)I_N\left(\lambda_j\right)$$
$$-\frac{1}{N}\sum_{j=1}^{[N/2]}\log\left(1+\alpha^2-2\alpha\cos\left(\lambda_j\right)\right). \tag{3.3}$$

The numerical minimization of equation (3.3) is achieved using numerical algorithms.

3.2 Maximum Empirical Likelihood Estimation

The empirical likelihood (EL) method, which is introduced by Owen (1988, 1990), is a way to expand the likelihood-based inference ideas to certain nonparametric situations, Chuang and Chan (2002). For a sequence of i.i.d. observations, $x_1,\cdots,x_n$, of a discrete-valued random variable X with some unknown distribution functions F, the EL function is defined as

$$L(F)=\prod_{t=1}^{n}dF(x_t)=\prod_{t=1}^{n}p_t, \tag{3.4}$$

and it can be shown that the maximizer of EL function, subject to the constraints $p_t \geq 0$ and $\sum_{t=1}^{n}p_t=1$, is the empirical distribution function $F_n(x)=(1/n)\sum_{t=1}^{n}I(x_t<x)$. The empirical likelihood ratio (ELR) is introduced as $R(F)=L(F)/L(F_n)$ and equivalently, it can be written as $R(F)=\prod_{t=1}^{n}np_t$.

Let $\boldsymbol{\omega}\in\mathbb{R}^d$ be the vector of parameters of interest in the distribution F. Presume that, for some function g, the constrain $E_F(g(X,\boldsymbol{\omega}))=0$ holds. The sample version of this constraint, denoted by $\sum_{t=1}^{n}p_tg(x_t,\boldsymbol{\omega})=0$, should be imposed in the estimation of $\boldsymbol{\omega}$. For this purpose, the profile ELR function is defined as

$$R(\boldsymbol{\omega})=\max\left\{\prod_{t=1}^{n}np_t|\ p_t\geq 0,\ \sum_{t=1}^{n}p_t=1,\ \sum_{t=1}^{n}p_tg(x_t,\boldsymbol{\omega})=0\right\}. \tag{3.5}$$

The maximizer of $R(\boldsymbol{\omega})$ can be obtained using the Lagrange multiplier (LM) method. Let $\phi\in\mathbb{R}$ and $\boldsymbol{\varphi}\in\mathbb{R}^d$ be the Lagrange multipliers and consider $\mathcal{L}(\boldsymbol{\omega})$ as

$$\mathcal{L}(\boldsymbol{\omega})=\sum_{t=1}^{n}\log(p_t)+\phi\left(1-\sum_{t=1}^{n}p_t\right)-n\boldsymbol{\varphi}'\sum_{t=1}^{n}p_tg(x_t,\boldsymbol{\omega}).$$

It can be demonstrated that the maximum is attained for $p_t = 1/\left[n\left(1+\boldsymbol{\varphi}' g\left(x_t, \boldsymbol{\omega}\right)\right)\right]$ and $\phi = n$. Besides, $\boldsymbol{\varphi}$, which is a function of the unknown parameter $\boldsymbol{\omega}$, is a solution to

$$\sum_{t=1}^{n} \frac{g\left(x_t, \boldsymbol{\omega}\right)}{1+\boldsymbol{\varphi}' g\left(x_t, \boldsymbol{\omega}\right)} = 0.$$

Therefore, the minus log profile ELR function can be written as

$$\mathcal{L}\left(\boldsymbol{\omega}\right) = -\log R\left(\boldsymbol{\omega}\right) = \sum_{t=1}^{n} \log\left(1+\boldsymbol{\varphi}' g\left(x_t, \boldsymbol{\omega}\right)\right),$$

and the minimizer of $\mathcal{L}\left(\boldsymbol{\omega}\right)$ is called the maximum empirical likelihood estimator (MELE) of $\boldsymbol{\omega}$.

This method is extended to statistical models with a martingale structure by Mykland (1995). He proved that the derivative of the objective function with respect to the unknown parameter is a martingale under the true parameter. Therefore, the score function is used to construct the ELR statistic. Here, this method is applied to obtain the profile ELR function.

Consider the PLINAR(1) process given by (2.1) and let the conditional least squares (CLS) criterion function be defined as

$$S\left(\boldsymbol{\omega}\right) = \sum_{t=1}^{n} S_t\left(\boldsymbol{\omega}\right) = \sum_{t=1}^{n} \left(X_t - \alpha X_{t-1} - (1-\alpha)\frac{\theta+2}{\theta\left(\theta+1\right)}\right)^2, \tag{3.6}$$

where $\boldsymbol{\omega} = (\alpha, \theta)'$. Let $D_t\left(\boldsymbol{\omega}\right) = -\left(1/2\right) \partial S_t\left(\boldsymbol{\omega}\right)/\partial \boldsymbol{\omega}$. The solution of equation $\sum_{t=1}^{n} D_t\left(\boldsymbol{\omega}\right) = 0$ yields CLS estimation of $\boldsymbol{\omega}$. It is easy to verify that $D_t\left(\boldsymbol{\omega}\right) = \left(D_{t1}\left(\boldsymbol{\omega}\right), D_{t2}\left(\boldsymbol{\omega}\right)\right)'$, with

$$D_{t1}\left(\boldsymbol{\omega}\right) = \left(X_t - \alpha X_{t-1} - (1-\alpha)\frac{\theta+2}{\theta\left(\theta+1\right)}\right)\left(X_{t-1} + \alpha\frac{\theta+2}{\theta\left(\theta+1\right)}\right), \tag{3.7}$$

$$D_{t2}\left(\boldsymbol{\omega}\right) = \left(X_t - \alpha X_{t-1} - (1-\alpha)\frac{\theta+2}{\theta\left(\theta+1\right)}\right)\left((1-\alpha)\frac{\theta^2+\theta+2}{\theta^2\left(\theta+1\right)^2}\right). \tag{3.8}$$

Let $\mathcal{F}_t = \sigma\left(X_0, X_1, \cdots, X_t\right)$. It can be shown that $\{D_t\left(\boldsymbol{\omega}\right), \mathcal{F}_t, t \geq 1\}$ is a martingale difference sequence, Zhang et al. (2011). Based on the score function $D_t\left(\boldsymbol{\omega}\right)$ and following Mykland (1995) and Chuang and Chan (2002), the profile ELR function can be constructed as

$$R\left(\boldsymbol{\omega}\right) = \max\left\{\prod_{t=1}^{n} np_t \middle| p_t > 0, \sum_{t=1}^{n} p_t = 1, \sum_{t=1}^{n} p_t D_t\left(\boldsymbol{\omega}\right) = 0\right\}, \tag{3.9}$$

and the maximizer of (3.9) can be found by the method of Lagrange multipliers. Following the same steps as in the i.i.d case, It can be verified that the optimal value of p_t is $1/\left(n\left(1+\boldsymbol{\varphi}'\left(\boldsymbol{\omega}\right)D_t\left(\boldsymbol{\omega}\right)\right)\right)$ and $\boldsymbol{\varphi}'\left(\boldsymbol{\omega}\right)$ satisfying

$$\sum_{t=1}^{n}\frac{D_t\left(\boldsymbol{\omega}\right)}{1+\boldsymbol{\varphi}'\left(\boldsymbol{\omega}\right)D_t\left(\boldsymbol{\omega}\right)}=0. \tag{3.10}$$

The solution to Equation (3.10) can be obtained by numerical algorithms. Thus, the log profile ELR statistic has the following form

$$\mathcal{L}\left(\boldsymbol{\omega}\right)=-2\log\left(R\left(\boldsymbol{\omega}\right)\right)=2\sum_{t=1}^{n}\log\left(1+\boldsymbol{\varphi}'\left(\boldsymbol{\omega}\right)D_t\left(\boldsymbol{\omega}\right)\right). \tag{3.11}$$

The minimizer $\tilde{\boldsymbol{\omega}}$ of (3.11) is the "MELE of $\boldsymbol{\omega}$.

3.3 Sieve Bootstrap Parameter Estimation

The sieve bootstrap (SB) method, which is applied by Bühlmann (1997) in the analysis of time series data, is extended to the integer-valued time series by Cardinal et al. (1999) and Kim and Park (2008). Our proposed algorithm to obtain the parameters of PLINAR(1) using SB method is as follows:

Algorithm 1

1- Compute the residuals; $\hat{\varepsilon}_t = X_t - \hat{\alpha}X_{t-1};\ t = 2,\ldots,n$, where $\hat{\alpha}$ can be any of the estimations mentioned in Section 3.

2- Since each error $\hat{\varepsilon}_t$ may include a fractional part and even have a negative value, the considered empirical distribution is related to the modified errors $\tilde{\varepsilon}_t$ defined by

$$\tilde{\varepsilon}_t=\begin{cases}[\hat{\varepsilon}_t], & \hat{\varepsilon}_t>0,\\ 0, & \hat{\varepsilon}_t\le 0,\end{cases}$$

where [.] represents the value rounded to the nearest integer. Therefore, the empirical distribution function of the modified residuals is defined as:

$$\hat{F}_{\tilde{\varepsilon}}(x)=\frac{1}{n-1}\sum_{t=2}^{n}I_{(\tilde{\varepsilon}_t\le x)}.$$

3- Draw B sets of i.i.d samples $\varepsilon_t^b,\ b = 1, \ldots, B,\ t = 1, \ldots, n$, from the empirical distribution $\hat{F}_{\tilde{\varepsilon}}(.)$.

4- Define X_t^b by the recursion:

$$X_t^b = \hat{\alpha} \circ X_{t-1}^b + \varepsilon_t^b,\ \ t = 1, \ldots, n,\ t = 1, \ldots, B.$$

5- Based on $\{X_1^b, X_2^b, \ldots, X_t^b\}, t \leq n$, compute the estimation of the $PLINAR(1)$ coefficients $\hat{\alpha}^b$ and $\hat{\theta}^b$, as in step 1.

6- Estimations of α and θ can be obtained considering the sample mean

$$\hat{\alpha} = \frac{1}{B}\sum_{b=1}^{B} \hat{\alpha}^b,\ \ \hat{\theta} = \frac{1}{B}\sum_{b=1}^{B} \hat{\theta}^b.$$

4 Prediction in the PLINAR(1) Process

In practice, the point and interval forecasts of future values are of great importance and some methods are developed to predict the future values of integer-valued time series. For example, the concept of coherent forecasting is introduced by Freeland and McCabe (2004) in the context of integer-valued time series data. Moreover, the probabilistic forecasts by estimating the forecasting distribution is presented by Freeland and McCabe (2004) and McCabe et al. (2011).

Based on the classical methodology, the h-step-ahead predictor of PLINAR(1) is defined using the conditional mean, i.e.,

$$\hat{X}_{t+h} = E\left[X_{t+h}|X_t\right] = \alpha^h X_t + \left(1 - \alpha^h\right)\frac{\theta + 2}{\theta\left(\theta + 1\right)}. \tag{4.1}$$

This prediction method is studied by Bránnás (1994) and Freeland and McCabe (2004), but it hardly produces integer-valued forecasts. In other words, the conditional mean-based prediction is not a coherent forecasting method, i.e., it is not integer-valued.

Median of transition probability-based predictor, which is proposed by Freeland and McCabe (2004), applies the minimizer of the expected absolute error given the sample, $E\left[\left|X_{t+h} - \hat{X}_{t+h}\right|\ |X_t\right]$, as a coherent prediction for X_{t+h}.

Let m_{t+h} be the median of the h-step ahead conditional distribution $f\left(x_{t+h}|x_t\right)$. In this case, m_{t+h} is the predicted value of X_{t+h}, i.e., $\hat{X}_{t+h} = m_{t+h}$, and, for PLINAR(1) process, it is defined as the smallest non-negative integer such that

$$\sum_{i=0}^{m_{t+h}} f_{X_{t+h}|X_t}(i|j) = \sum_{i=0}^{m_{t+h}} \sum_{k=0}^{\min(i,j)} \binom{i}{k} \alpha^{tk} \left(1-\alpha^h\right)^{j-k} P\left(W_{t+h} = i-k\right) \geq 0.5,$$

where the PMF of W_{t+h} is given in (2.9).

In 2018, Wang and Zhang compared the rounded conditional mean-based predictor (RCM) with the median of transition probability-based predictor (MTP) in PLINAR(1) process. Moreover, they used the highest predicted probability (HPP) interval for interval prediction of future observation in this process. In both simulation studies and real data analysis, they concluded that the MTP was much better than RCM. So, they suggested using the MTP based predictor for data prediction in the PLINAR(1) process. Note that if $f_{X_{t+h}|X_t}(0|x_t) > 0.5$, the median is not defined and median forecast method cannot be applied. (See, Simarmata et al. (2017)). As we can see in the Anorexia data set, Table (7) in Section 7, in PLINAR(1) the chance of zero is high, so the MTP method is not an appropriate method. Here, we are going to apply two other prediction methods, Baysian and sieve bootstrap, for PLINAR(1) time series.

4.1 Bayesian Forecasting Method

Consider the future observation X_{t+h} and the random vector of unknown parameters $\boldsymbol{\omega} = (\alpha, \theta)'$. The information about $\boldsymbol{\omega}$ is obtained through the observed sample $\mathbf{x}_t$ and is quantified in the posterior predictive, $\pi(\boldsymbol{\omega}|\mathbf{x}_t)$. The Bayesian predictive probability function is a weighted average over the parameter space Θ, and as a posterior distribution, it assigns a weight to every possible parameter setting. Silva et al. (2009) and Simarmata et al. (2017) suggested a Bayesian methodology for Poisson INAR(1) process. Here, we extend their method to PLINAR(1) process.

Definition 4.1. Let $\boldsymbol{\omega} \in \Theta$ be the vector of unknown parameters. The h-step ahead Bayesian posterior predictive distribution is presented by

$$\begin{aligned} f_b(x_{t+h}|\mathbf{x}_t) &= \int_\Theta f(x_{t+h}, \boldsymbol{\omega}|\mathbf{x}_t) d\boldsymbol{\omega} \\ &= \int_\Theta f(x_{t+h}|\boldsymbol{\omega}, \mathbf{x}_t)\pi(\boldsymbol{\omega}|\mathbf{x}_t) d\boldsymbol{\omega}, \end{aligned} \tag{4.2}$$

where $\pi(\boldsymbol{\omega}|\mathbf{x}_t)$ is the posterior probability function of $\boldsymbol{\omega}$.

Note 1. In Section 2, the h-step ahead transition probability is denoted by $f(x_{t+h}|x_t)$. In this section, to emphasis on the randomness of the parameters, $f(x_{t+h}|x_t)$ is substituted by $f(x_{t+h}|\boldsymbol{\omega}, \mathbf{x}_t)$.

When $f_b(x_{t+h}|\mathbf{x}_t)$ is obtained, the expected value, the median or the mode of this distribution can be considered as the Bayesian h-step-ahead predictor.

In PLINAR(1) model, since beta and gamma distributions are the conjugation of binomial and PL distributions, they can be applied as the prior distributions of the parameters α and θ. More precisely, let $\alpha \sim Beta(a, b),\ a, b > 0$ and $\theta \sim \Gamma(c, d),\ c, d > 0$. If α and θ are consider to be independent random variables, then the prior distribution of $\boldsymbol{\omega}$ would be:

$$\begin{aligned}\pi(\boldsymbol{\omega}) &= \pi(\alpha)\pi(\theta)\\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\alpha^{a-1}(1-\alpha)^{b-1}\frac{d^c}{\Gamma(c)}e^{-d\theta}\theta^{c-1}\\ &\propto \alpha^{a-1}(1-\alpha)^{b-1}e^{-d\theta}\theta^{c-1}, \qquad \theta > 0, \qquad 0 < \alpha < 1, \end{aligned} \tag{4.3}$$

where a, b, c and d are known parameters. The posterior distribution of $\boldsymbol{\omega}$ can be written as

$$\begin{aligned}\pi(\boldsymbol{\omega}|\mathbf{x}_t) &\propto L(\mathbf{x}_t|x_1, \boldsymbol{\omega})\,\pi(\boldsymbol{\omega})\\ &= f(x_t|x_{t-1}, \boldsymbol{\omega})\cdots f(x_3|x_2, \boldsymbol{\omega})\,f(x_2|x_1, \boldsymbol{\omega})\,\pi(\boldsymbol{\omega})\\ &\propto \prod_{n=2}^{t}\sum_{k=0}^{M_n}\binom{x_{n-1}}{k}\alpha^{k+a-1}(1-\alpha)^{x_{n-1}-k+b-1}P(I_nH_n = x_n - k)\,e^{-d\theta}\theta^{c-1}, \end{aligned} \tag{4.4}$$

where $L(\mathbf{x}_t|x_1, \boldsymbol{\omega})$ is the conditional likelihood function and $M_n = \min(x_n, x_{n-1})$. The complex structure of the posterior distribution of $\boldsymbol{\omega}$ makes it difficult to derive the marginal distribution, and consequently, the posterior mean value of each of the unknown parameters. Therefore, to apply the Gibbs sampling algorithm in the simulation studies, we need the full conditional posterior distribution of α and θ. Using Equation (4.4) the full conditional posterior distribution of α is given by

$$\pi(\alpha|\mathbf{x}_t, \theta) \propto \alpha^{a-1}(1-\alpha)^{b-1}\prod_{n=2}^{t}\sum_{k=0}^{M_n}\binom{x_{n-1}}{k}\alpha^{k}(1-\alpha)^{x_{n-1}-k}P(I_nH_n = x_n - k), \tag{4.5}$$

while the full conditional posterior distribution of θ is obtained as

$$\pi(\theta|\mathbf{x}_t, \alpha) \propto e^{-d\theta}\theta^{c-1}\prod_{n=2}^{t}\sum_{k=0}^{M_n}\binom{x_{n-1}}{k}\alpha^{k}(1-\alpha)^{x_{n-1}-k}P(I_nH_n = x_n - k). \tag{4.6}$$

Consequently, for the PLINAR(1) model, the Bayesian predictive function of X_{t+h} given $\mathbf{X}_t$ is

$$
\begin{aligned}
f_b(x_{t+h}|\mathbf{x}_t) &= \int_{\Theta} f(x_{t+h}|\boldsymbol{\omega}, \mathbf{x}_t)\pi(\boldsymbol{\omega}|\mathbf{x}_t)d\boldsymbol{\omega} \\
&\propto \int_0^{\infty}\int_0^1 \sum_{k=0}^{\min(x_t,x_{t+h})} \binom{x_t}{k}\alpha^{hk}\left(1-\alpha^h\right)^{x_t-k} P\left(W_{t+h}=x_{t+h}-k\right) \\
&\prod_{n=2}^{t}\sum_{k=0}^{M_n}\binom{x_{n-1}}{k}\alpha^k(1-\alpha)^{x_{n-1}-k} P\left(I_nH_t=x_n-k\right) \\
&\alpha^{a-1}(1-\alpha)^{b-1}e^{-d\theta}\theta^{c-1}d\alpha d\theta.
\end{aligned}
\tag{4.7}
$$

The complexity of $f_b(x_{t+h}|\mathbf{x}_t)$ prevents us from finding a solution using the standard Bayesian method. Therefore, the algorithm of Bayesian forecasting method is applied to determine the forecasting value. Let $(X_1,\ X_2,\ \ldots,\ X_t)$ be observed. The following algorithm is used to obtain the h step ahead sample $\left(X_{t+h;1},\ X_{t+h;2},\ \ldots,\ X_{t+h;m}\right)$ from equation (4.7).

Algorithm 2

1- A starting estimation for α and θ, called α_0 and θ_0, are calculated from the sample $(X_1,\ X_2,\ \ldots,\ X_t)$, through the classical method.

2- Obtain a sample (α_1,θ_1), (α_2,θ_2), $\ldots$, (α_m,θ_m) from the full conditional distributions of parameters, where for each pair (α_j,θ_j), $j=1,2,\ldots,m$, we simulate

$$\alpha_{j,k} \sim \pi(\alpha|\theta_{j,k-1},\mathbf{X}_t), \quad k=1,\ldots,K, \tag{4.8}$$

and

$$\theta_{j,k} \sim \pi(\theta|\alpha_{j,k-1},\mathbf{X}_t), \quad k=1,\ldots,K, \tag{4.9}$$

where $\alpha_{j,0}=\alpha_0$, $\theta_{j,0}=\theta_0$, $\alpha_j=\alpha_{j,K}$ and $\theta_j=\theta_{j,K}$.

3- For each j, $j=1,\ldots,m$, obtain $X_{t+h;j}$ from $f(x_{t+h}|x_t,\alpha_j,\theta_j)$, using the inverse transform method adapted to integer variables, i.e.,

(a) Sample u from uniform $U(0,1)$,

(b) Calculate the least integer valued s, such that $\sum_{x_{t+h}=0}^{s} f\left(x_{t+h}|x_t,\alpha_j,\theta_j\right) \geq u$,

(c) Let $X_{t+h;j} = s$.

After obtaining the sample points $X_{t+h;1}, X_{t+h;2}, \ldots, X_{t+h;m}$, the h-step ahead predictor $\hat{X}_{t+h}$ can be calculated using sample mean ($\bar{X}_{t+h)}$), median ($\tilde{m}_{t+h}$) or mode ($\tilde{M}_{t+h)}$). We can also calculate $E(X_{t+h}|X_t)$ using properties of mathematical expectation, i.e.,

$$\begin{aligned}
E\left[X_{t+h}|X_t\right] &= E\left[E\left[X_{t+h}|X_t, \boldsymbol{\omega}\right]|X_t\right] \\
&= E\left[\alpha^h X_t + \left(1-\alpha^h\right)\frac{\theta+2}{\theta(\theta+1)}\middle| X_t\right] \\
&= X_t E\left[\alpha^h\middle| X_t\right] + E\left[\left(1-\alpha^h\right)\frac{\theta+2}{\theta(\theta+1)}\middle| X_t\right].
\end{aligned} \tag{4.10}$$

Markov Chain Monte Carlo (MCMC) algorithm can be used to estimate these expected values. Here, Metropolis algorithm is performed in conjunction with Adaptive Rejection Sampling Method (ARMS) in order to sample values from full conditional distributions of α and θ, denoted by $(\alpha_1, \alpha_2, \ldots, \alpha_m)$, $(\theta_1, \theta_2, \ldots, \theta_m)$, respectively, (see Silva and Oliveira (2005) and Silva et al. (2009)). We have,

$$\hat{E}\left[\alpha^h\middle| x_t\right] = \frac{1}{m}\sum_{i=1}^{m}\alpha_i^h,$$

and

$$\hat{E}\left[\left(1-\alpha^h\right)\frac{\theta+2}{\theta(\theta+1)}\middle| x_n\right] = \frac{1}{m}\sum_{i=1}^{m}\left(1-\alpha_i^h\right)\frac{\theta_i+2}{\theta(\theta_i+1)}.$$

Consequently, the predictor can be written as

$$\hat{X}_{n+h} = X_n\left(\frac{1}{m}\sum_{i=1}^{m}\alpha_i^h\right) + \left(\frac{1}{m}\sum_{i=1}^{m}\left(1-\alpha_i^h\right)\frac{\theta_i+2}{\theta_i(\theta_i+1)}\right).$$

4.2 Sieve Bootstrap Forecast

As mentioned previously, when the time series is integer-valued, the conditional expectation, the Bayesian approach, the mean and MCMC methods are not coherent forecasting methods. To preserve the integer-valued nature of data, we suggest an alternative method, namely the bootstrap approach. A bootstrap approach is a distribution-free alternative method. In the following, we employ the bootstrap method proposed by Pascual et al. (2004) after some modifications to PLINAR(1) model using the next algorithm.

Algorithm 3

1- Estimate α, θ by SB method as $\hat{\alpha}$ and $\hat{\theta}$ mentioned in Section 3.3.

2- Compute future bootstrap observations by the recursion:

$$X^*_{n+h} = \hat{\alpha} \circ X^*_{n+h-1} + \varepsilon^*_{n+h}, \qquad h > 0,$$

where $h > 0$, and $X^*_t = X_t$, $t \leq n$ and ε^*_{n+h} is a sample from the empirical distribution computed in Step 2.

5 Numerical Simulations

In this section, we generate observations from PLINAR(1) model with four different sets of parameters, $(\alpha, \theta) = (0.3, 2)$, $(0.3, 3)$, $(0.5, 2)$, and $(0.5, 3)$. In each case, we do the simulation study for three sample sizes, $n = 100$, 300, and 500 and all the sample observations are repeated $N = 1000$ times.

To check the efficiency of the proposed estimation methods, we compare them using the conditional least square (CLS) and maximum likelihood estimation (MLE) methods (Mohammadpour et al. (2018)). In Table (1), we report the estimated bias (BIAS) and sample standard error (SSE) for $\left(\hat{\alpha}, \hat{\theta}\right)$. For each case, the first line and the second line, which is in bold-face-type, show BIAS and SSE for $\left(\hat{\alpha}, \hat{\theta}\right)$, respectively. It can be seen from the results that the BIAS and SSE of all the estimations are decreasing as sample size n increases and all the suggested methods are work well.

The aim of the second simulation is to compare the conditional mean-based predictor with conditional least square estimation suggested by Mohammadpour et al. (2018) with the Bayesian, median, bootstrap and conditional mean prediction methods, based on our parameter estimation methods, MELE, Whittle and bootstrap, toward h-step prediction.

We generate $n + 5$ observations from the PLINAR(1) process, $n = 100$, 300, 500, where the first n observations are used to estimate the parameters and the remaining observations are used to calculate the prediction mean absolute error (PMAE) as

$$PMAE(h) = \frac{1}{N} \sum_{k=1}^{N} |X^{(k)}_{N+h} - \hat{X}^{(k)}_{N+h}|,$$

where $X_{N+h}^{(k)}$ is the $(n+h)$-th observed data, $\hat{X}_{N+h}^{(k)}$ is the corresponding h-step prediction and k is the repetition times.

Table 1: Bias and SSE of $\left(\hat{\alpha},\hat{\theta}\right)$ for the PLINAR(1) process

(α,θ)	n	CLS	MELE	Whittle	Bootstrap	MLE
(0.3,2)	100	(-0.0216,0.0790)	(-0.0205,0.0754)	(-0.0091, 0.1399)	(-0.0205, 0.0754)	(-0.0107, 0.0834)
		(0.0123,0.1699)	**(0.01200,0.1638)**	**(0.0117, 0.2282)**	**(0.0120, 0.1638)**	**(0.0078, 0.1708)**
	300	(-0.0092,0.0345)	(-0.0055,0.0295)	(-0.0058,0.0538)	(-0.0055,0.0295)	(-0.0091, 0.0598)
		(0.0044,0.0517)	**(0.0046,0.0472)**	**(0.0042,0.0670)**	**(0.0046,0.0472)**	**(0.0074, 0.1241)**
	500	(-0.0055,0.0194)	(-0.0030,0.0157)	(-0.0036,0.0288)	(-0.0030,0.0158)	(-0.0088, 0.0555)
		(0.0027,0.0243)	**(0.0027,0.0277)**	**(0.0027,0.0360)**	**(0.0027,0.0277)**	**(0.0069, 0.1116)**
(0.3,3)	100	(-0.0205, -1.9727)	(-0.0273, 0.2178)	(-0.0129, 0.2312)	(-0.0273, 0.2178)	(0.1457,-0.0088)
		(0.0110, 0.2800)	**(0.0126, 0.6505)**	**(0.0126, 0.8287)**	**(0.0126, 0.6505)**	**(0.0403, 0.1192)**
	300	(-0.0094,0.0562)	(-0.0139,0.0707)	(-0.0045,0.0987)	(-0.0139,0.0707)	(-0.0243, -0.2847)
		(0.0049,0.1453)	**(0.0046,0.1808)**	**(0.0053,0.2064)**	**(0.0046,0.1808)**	**(0.0067, 0.0863)**
	500	(-0.0048, 0.0325)	(-0.0045,0.0216)	(-0.0028, 0.0555)	(-0.0045,0.0216)	(-0.0025, 0.2810)
		(0.0030,0.0811)	**(0.0029, 0.0816)**	**(1.956e-17,1.323e-16)**	**(0.0029,0.0816)**	**(0.0065, 0.0872)**
(0.5,2)	100	(-0.0304, 0.1452)	(-0.02592, 0.1214)	(-0.0202, 0.2114)	(-0.02590, 0.1214)	(-0.0157, 0.1473)
		(0.01135, 0.2639)	**(0.0113, 0.2861)**	**(0.0117,0.3488)**	**(0.0113, 0.2861)**	**(0.0067, 0.2622)**
	300	(0.0100,0.0522)	(-0.0114,0.0598)	(-0.0065,0.0763)	(-0.0114,0.0598)	(-0.0097, 0.0569)
		(0.0039,0.0828)	**(0.0042,0.0818)**	**(0.0039,0.1092)**	**(0.0042,0.0818)**	**(0.0059, 0.1242)**
	500	(-0.0070,0.0260)	(-0.0072,0.0183)	(-0.0051,0.0381)	(-0.0072,0.0183)	(-0.0065, 0.0387)
		(0.0025,0.0400)	**(0.0026,0.0462)**	**(0.0025,0.0514)**	**(0.0026,0.0462)**	**(0.0053, 0.1110)**
(0.5,3)	100	(-0.0363, 0.3333)	(-0.0420, 0.3832)	(-0.0164, 0.3594)	(-0.0420, 0.3832)	(-0.0156, -0.5000)
		(0.01407, 1.2658)	**(0.0139, 1.4615)**	**(0.0300, 1.2788)**	**(0.0139, 1.4615)**	**(0.0030, 0.2502)**
	300	(-0.0089,0.0660)	(-0.0124,0.0433)	(-0.0096,0.0376)	(-0.0124,0.0433)	(-0.0068, 0.0496)
		(0.0046,0.2457)	**(0.0046,0.2161)**	**(-0.0397,0.4317)**	**(0.0046,0.2161)**	**(0.0070, 0.1690)**
	500	(-0.0065,0.0417)	(-0.0038,0.0533)	(-0.0059,0.0682)	(-0.0038,0.0353)	(-0.0069, 0.0416)
		(0.0029,0.1543)	**(0.0029,0.1626)**	**(0.0027,0.1843)**	**(0.0029,0.1626)**	**(0.0078, 0.1659)**

The results can be seen in Tables (2). The PMAE for Bayesian method and, after that, Bayesian, the median and classic prediction when the parameters are estimated with Whittle method, were less than other methods, which indicate that these three prediction method are better than the other methods.

A standard interval in autoregressive models for h-step interval forecasting is base on asymptotic normality property of $E\left(X_{n+h}|X_n\right)$ (Bhansali (1974)). Maiti and Biswas (2015) suggested that we can use the $100(1-\alpha)\%$ highest predicted probability (HPP) interval where $\alpha\in(0,1)$. Based on the definition of HPP interval and the unimodality of the forecasting distribution, Wang and Zhang (2018) suggested an algorithm for

$100(1-\alpha)\%$ HPP interval for PLINAR(1).

Table 2: PMAE for simulated PLINAR(1) process

(α,θ)	n	h	conditional mean (CLS)	conditional mean (MELE)	conditional mean (Whittle)	Median	Bootstrap	Bayesian
(0.3,2)	100	1	0.7212	1.0703	0.7214	0.7870	0.9610	0.6431
		2	0.8103	1.2722	0.8089	0.7830	1.0360	0.6807
		3	0.7618	0.9059	0.7480	0.7510	0.9660	0.6333
		4	0.7832	0.8997	0.7507	0.7210	0.9620	0.6348
		5	0.7850	1.0816	0.7817	0.7850	1.0030	0.6959
	300	1	0.7130	0.7617	0.6860	0.7230	0.9320	0.5466
		2	0.7686	0.7962	0.7807	0.7530	1.0040	0.6572
		3	0.7612	0.7873	0.7586	0.7030	0.9661	0.6305
		4	0.7753	0.7969	0.7508	0.7070	0.9614	0.6340
		5	0.7823	0.7520	0.7806	0.7240	0.9690	0.5769
	500	1	0.7105	0.7499	0.6853	0.7150	0.8680	0.5398
		2	0.7630	0.7747	0.7616	0.6480	0.9660	0.6560
		3	0.7607	0.7219	0.7776	0.6460	0.9540	0.5644
		4	0.7749	0.7495	0.7492	0.6470	0.9600	0.6289
		5	0.7741	0.7524	0.7734	0.6710	0.9140	0.4907
(0.3,3)	100	1	1.3261	0.5148	0.5318	1.4200	0.6854	0.4163
		2	1.3927	0.5914	0.5857	1.4280	0.7392	0.4359
		3	1.3054	0.5924	0.5901	1.3460	0.6623	0.4719
		4	1.2968	0.5981	0.5826	1.3600	0.6886	0.4713
		5	1.4050	0.5835	0.5799	1.5100	0.7355	0.4863
	300	1	0.5187	0.5127	0.5347	0.5560	0.6450	0.4110
		2	0.5862	0.6000	0.5761	0.5640	0.6360	0.4302
		3	0.5844	0.6046	0.5676	0.5970	0.5990	0.4552
		4	0.5808	0.5726	0.5821	0.5920	0.6370	0.4702
		5	0.6012	0.5535	0.5793	0.5610	0.6120	0.4400
	500	1	0.5478	0.5175	0.5367	0.5460	0.6260	0.3829
		2	0.5770	0.5677	0.5759	0.5470	0.6230	0.4309
		3	0.5801	0.5787	0.5786	0.5630	0.6140	0.4544
		4	0.5801	0.5686	0.5793	0.5720	0.6000	0.4659
		5	0.5726	0.5454	0.5708	0.5410	0.6290	0.4404
(0.5,2)	100	1	0.7198	0.6059	0.6080	0.7160	0.7650	0.5001
		2	0.8312	0.7412	0.6626	0.7220	0.7880	0.6201
		3	1.0555	0.7797	0.7106	0.7120	0.7989	0.7274
		4	0.9874	0.7500	0.7129	0.7160	0.8250	0.6641
		5	1.1123	0.7728	0.7214	0.7120	0.8990	0.6631
	300	1	0.7067	0.5804	0.6000	0.6770	0.7650	0.5005
		2	0.8187	0.7256	0.7009	0.6800	0.7820	0.6142
		3	1.1373	0.7582	0.7005	0.6810	0.7940	0.7284
		4	0.9816	0.7759	0.7137	0.6950	0.8070	0.6341
		5	1.0283	0.7625	0.7012	0.6810	0.8750	0.6677
	500	1	0.6300	0.5605	0.5985	0.6570	0.7650	0.5001
		2	0.7401	0.7252	0.6390	0.6530	0.7560	0.6451
		3	0.7733	0.7244	0.6721	0.6500	0.7920	0.6083
		4	0.7761	0.7666	0.7125	0.6750	0.8080	0.6217
		5	0.7649	0.7584	0.7012	0.6800	0.8670	0.6627
(0.5,3)	100	1	0.4608	0.4358	0.4432	0.4230	0.5870	0.3565
		2	0.5177	0.5244	0.5463	0.4370	0.5380	0.4350
		3	0.5462	0.5616	0.5463	0.4300	0.5980	0.4808
		4	0.5873	0.5844	0.5949	0.4790	0.5860	0.4187
		5	0.5891	0.6185	0.5977	0.4880	0.5960	0.4545
	300	1	0.4407	0.4280	0.4417	0.3920	0.5800	0.3540
		2	0.5110	0.5201	0.5304	0.4140	0.5340	0.4235
		3	0.5419	0.5609	0.5451	0.4280	0.5790	0.4070
		4	0.5749	0.5785	0.5778	0.4610	0.5870	0.3933
		5	0.5765	0.5971	0.5878	0.4690	0.5470	0.4103
	500	1	0.4219	0.4033	0.4343	0.3800	0.5480	0.3270
		2	0.5030	0.5011	0.5150	0.4050	0.5060	0.4138
		3	0.5154	0.5600	0.5290	0.4010	0.5460	0.4029
		4	0.5480	0.5618	0.5726	0.4490	0.5830	0.3853
		5	0.5618	0.5955	0.5710	0.4630	0.5270	0.3675

Table 3: 95% HPP intervals for the prediction of PLINAR(1) simulated data

(α,θ)	h		conditional mean (MELE)	conditional mean (Whittle)	Bootstrap	Bayesian
(0.3,2)				$n=100$		
	1	HPPI	(0,2.624)	(0,2.519)	(0,2.446)	(0.001,2.572)
		CP%	0.964	0.968	0.969	0.968
		LPI	2.624	2.519	2.446	2.570
	2	HPPI	(0,2.896)	(0,2.808)	(0, 2.534)	(0.001,2.828)
		CP%	0.964	0.968	0.969	0.968
		LPI	2.896	2.808	2.534	2.827
	3	HPPI	(0,2.966)	(0,2.871)	(0,2.542)	(0.001,2.887)
		CP%	0.964	0.968	0.969	0.968
		LPI	2.966	2.871	2.542	2.886
	4	HPPI	(0,2.984)	(0,2.894)	(0, 2.543)	(0.001,2.918)
		CP%	0.964	0.968	0.969	0.968
		LPI	2.984	2.894	2.543	2.917
	5	HPPI	(0,2.992)	(0,2.903)	(0, 2.543)	(0.001,2.939)
		CP%	0.964	0.968	0.969	0.968
		LPI	2.992	2.903	2.543	2.938
				$n=300$		
	1	HPPI	(0,2.544)	(0,2.537)	(0,2.722)	(0,2.430)
		CP%	0.967	0.967	0.968	0.967
		LPI	2.544	2.537	2.544	2.430
	2	HPPI	(0.001,2.871)	(0,2.852)	(0,2.833)	(0,2.748)
		CP%	0.968	0.968	0.968	0.968
		LPI	2.870	2.852	2.870	2.748
	3	HPPI	(0.001, 2.965)	(0,2.931)	(0,2.838)	(0,2.817)
		CP%	0.969	0.968	0.968	0.968
		LPI	2.964	2.931	2.964	2.817
	4	HPPI	(0.001,2.989)	(0,2.963)	(0,2.839)	(0, 2.840)
		CP%	0.969	0.968	0.968	0.968
		LPI	2.988	2.963	2.988	2.840
	5	HPPI	(0.001,2.993)	(0,2.972)	(0,2.839)	(0, 2.884)
		CP%	0.969	0.968	0.968	0.969
		LPI	2.992	2.972	2.992	2.884
				$n=500$		
	1	HPPI	(0,2.512)	(0.001,2.528)	(0,2.854)	(0,2.371)
		CP%	0.965	0.966	0.969	0.968
		LPI	2.512	2.527	2.854	2.370
	2	HPPI	(0,2.922)	(0.001,2.916)	(0, 2.950)	(0,2.704)
		CP%	0.969	0.969	0.969	0.970
		LPI	2.922	2.915	2.950	2.704
	3	HPPI	(0,3.014)	(0.001,3.003)	(0, 2.960)	(0, 2.759)
		CP%	0.970	0.969	0.969	0.970
		LPI	3.014	3.002	2.960	2.759
	4	HPPI	(0,3.046)	(0.001,3.024)	(0, 2.962)	(0, 2.775)
		CP%	0.970	0.969	0.969	0.967
		LPI	3.046	3.023	2.962	2.775
	5	HPPI	(0,3.052)	(0.001,3.030)	(0, 2.962)	(0, 2.807)
		CP%	0.970	0.969	0.969	0.971
		LPI	3.052	3.029	2.962	2.807

Table 4: 95% HPP intervals for the prediction of PLINAR(1) simulated data

(α, θ)	h		conditional mean (MELE)	conditional mean (Whittle)	Bootstrap	Bayesian
(0.5,2)				$n = 100$		
	1	HPPI	(0.020,2.368)	(0.005,2.356)	(0.014,2.377)	(0.009,2.363)
		CP%	0.968	0.968	0.968	0.969
		LPI	2.348	2.351	2.363	2.353
	2	HPPI	(0.029,2.741)	(0.008, 2.756)	(0.014,2.758)	(0.009,2.753)
		CP%	0.967	0.968	0.967	0.968
		LPI	2.712	2.748	2.744	2.743
	3	HPPI	(0.028, 2.915)	(0.006, 2.913)	(0.014,2.933)	(0.010,2.915)
		CP%	0.968	0.968	0.968	0.968
		LPI	2.887	2.907	2.919	2.904
	4	HPPI	(0.030, 2.995)	(0.009, 2.996)	(0.016,3.014)	(0.011,2.982)
		CP%	0.968	0.968	0.968	0.968
		LPI	2.965	2.987	2.998	2.971
	5	HPPI	(0.031, 3.027)	(0.009,3.026)	(0.017,3.046)	(0.011,3.017)
		CP%	0.968	0.968	0.968	0.968
		LPI	2.996	3.017	3.029	3.006
				$n = 300$		
	1	HPPI	(0.007,2.389)	(0.011,2.398)	(0,2.491)	(0,2.332)
		CP%	0.969	0.969	0.968	0.969
		LPI	2.382	2.387	2.491	2.331
	2	HPPI	(0.007,2.739)	(0.012,2.774)	(0,2.601)	(0,2.723)
		CP%	0.967	0.967	0.968	0.968
		LPI	2.732	2.762	2.601	2.723
	3	HPPI	(0.009,2.943)	(0.011,2.949)	(0,2.613)	(0, 2.879)
		CP%	0.968	0.967	0.968	0.968
		LPI	2.934	2.938	2.613	2.878
	4	HPPI	(0.01,3.021)	(0.013,3.038)	(0,2.616)	(0.001,2.950)
		CP%	0.968	0.968	0.968	0.968
		LPI	3.011	3.025	2.616	2.949
	5	HPPI	(0.01,3.053)	(0.013,3.072)	(0,2.616)	(0.001,2.974)
		CP%	0.968	0.968	0.968	0.968
		LPI	3.043	3.059	2.616	2.972
				$n = 500$		
	1	HPPI	(0.006, 2.426)	(0.011,2.42)	(0,2.683)	(0,2.257)
		CP%	0.970	0.969	0.968	0.970
		LPI	2.420	2.409	2.683	2.256
	2	HPPI	(0.008, 2.748)	(0.011,2.762)	(0, 2.79)	(0, 2.541)
		CP%	0.966	0.966	0.968	0.967
		LPI	2.740	2.751	2.790	2.541
	3	HPPI	(0.012, 2.986)	(0.012,2.985)	(0, 2.798)	(0 ,2.709)
		CP%	0.968	0.967	0.968	0.967
		LPI	2.974	2.973	2.798	2.709
	4	HPPI	(0.012,3.1)	(0.014,3.086)	(0,2.798)	(0, 2.816)
		CP%	0.969	0.968	0.968	0.968
		LPI	3.088	3.072	2.798	2.815
	5	HPPI	(0.014, 3.136)	(0.014,3.129)	(0,2.798)	(0, 2.831)
		CP%	0.969	0.969	0.968	0.967
		LPI	3.122	3.115	2.798	2.831

Table 5: 95% HPP intervals for the prediction of PLINAR(1) simulated data

(α, θ)	h		conditional mean (MELE)	conditional mean (Whittle)	Bootstrap	Bayesian
(0.3,3)				n = 100		
	1	HPPI	(0,1.827)	(0,1.827)	(0,1.780)	(0,1.835)
		CP%	0.972	0.972	0.972	0.971
		LPI	1.827	1.827	1.780	1.834
	2	HPPI	(0, 2.031)	(0, 2.045)	(0,1.864)	(0, 2.039)
		CP%	0.972	0.972	0.972	0.972
		LPI	2.031895	2.045	1.864	2.039
	3	HPPI	(0, 2.061914)	(0, 2.102)	(0,1.873)	(0, 2.093)
		CP%	0.972	0.972	0.972	0.972
		LPI	2.061	2.102	1.873	2.093
	4	HPPI	(0, 2.082)	(0,2.121)	(0,1.874)	(0,2.117)
		CP%	0.972	0.972	0.972	0.972
		LPI	2.082	2.121	1.874	2.116
	5	HPPI	(0, 2.08818)	(0, 2.126)	(0, 1.876)	(0,2.127)
		CP%	0.972	0.972	0.973	0.972
		LPI	2.088	2.126	1.876	2.127
				n = 300		
	1	HPPI	(0,1.906)	(0,1.872)	(0,1.906)	(0 ,1.858)
		CP%	0.974	0.973	0.974	0.972
		LPI	1.906	1.872	1.906	1.858
	2	HPPI	(0, 2.064)	(0,2.088)	(0,2.064)	(0, 2.063)
		CP%	0.974	0.974	0.974	0.973
		LPI	2.064	2.088	2.064	2.063
	3	HPPI	(0, 2.102)	(0, 2.124)	(0,2.102)	(0,2.105)
		CP%	0.973	0.974	0.973	0.973
		LPI	2.102	2.124	2.102	2.105
	4	HPPI	(0,2.110)	(0, 2.134)	(0,2.110)	(0, 2.117)
		CP%	0.973	0.974	0.973	0.973
		LPI	2.110	2.134	2.110	2.117
	5	HPPI	(0,2.112)	(0,2.136)	(0, 2.112)	(0,2.122)
		CP%	0.973	0.974	0.973	0.973
		LPI	2.112	2.136	2.112	2.121
				n = 500		
	1	HPPI	(0,1.964)	(0,1.927)	(0,1.964)	(0,1.885)
		CP%	0.975	0.974	0.975	0.932
		LPI	1.964	1.927	1.964	1.885
	2	HPPI	(0,2.120)	(0,2.110)	(0, 2.120)	(0, 2.098)
		CP%	0.975	0.974	0.975	0.951
		LPI	2.120	2.110	2.120	2.098
	3	HPPI	(0, 2.140)	(0,2.136)	(0, 2.140)	(0,2.127)
		CP%	0.974	0.974	0.974	0.945
		LPI	2.140	2.136	2.140	2.127
	4	HPPI	(0,2.144)	(0,2.148)	(0, 2.144)	(0,2.138)
		CP%	0.973	0.973	0.973	0.952
		LPI	2.144	2.148	2.144	2.138
	5	HPPI	(0,2.148)	(0,2.150)	(0,2.148)	(0,2.142)
		CP%	0.973	0.973	0.973	0.934
		LPI	2.148	2.150	2.148	2.142

Table 6: 95% HPP intervals for the prediction of PLINAR(1) simulated data

(α,θ)	h		conditional mean (MELE)	conditional mean (Whittle)	Bootstrap	Bayesian
(0.5,3)				$n=100$		
	1	HPPI	(0.002, 1.64)	(0.001, 1.616)	(0.002,1.64)	(0,1.594)
		CP%	0.9712865	0.970	0.971	0.971
		LPI	1.638	1.615	1.638	1.594
	2	HPPI	(0, 1.98)	(0.001,1.944)	(0, 1.980)	(0, 1.923)
		CP%	0.972	0.971	0.972	0.971
		LPI	1.980	1.943	1.980	1.923
	3	HPPI	(0,2.086)	(0.001,2.082)	(0,2.086)	(0,2.053)
		CP%	0.972	0.971	0.972	0.971
		LPI	2.086	2.081	2.086	2.053
	4	HPPI	(0.002,2.128)	(0.001,2.149)	(0.002,2.128)	(0,2.110)
		CP%	0.972	0.971	0.972	0.971
		LPI	2.126	2.148	2.126	2.110
	5	HPPI	(0.002,2.152)	(0.001,2.182)	(0.002,2.152)	(0,2.135)
		CP%	0.972	0.971	0.972	0.971
		LPI	2.150	2.181	2.150	2.135
				$n=300$		
	1	HPPI	(0.002,1.572)	(0.001,1.525)	(0.002,1.572)	(0.004 ,1.594)
		CP%	0.968	0.968	0.968	0.969
		LPI	1.570	1.524	1.570	1.590
	2	HPPI	(0.002, 2.046)	(0.002,1.989)	(0.002, 2.046)	(0.003,1.992)
		CP%	0.973	0.972	0.973	0.972
		LPI	2.044	1.987	2.044	1.989
	3	HPPI	(0.002,2.168)	(0.002,2.123)	(0.002, 2.168)	(0.004,2.122)
		CP%	0.974	0.973	0.974	0.972
		LPI	2.166	2.121	2.166	2.118
	4	HPPI	(0.002,2.202)	(0.002,2.176)	(0.002,2.202)	(0.003,2.175)
		CP%	0.973	0.973	0.973	0.972
		LPI	2.200	2.174	2.200	2.171
	5	HPPI	(0.002,2.216)	(0.002, 2.205)	(0.002, 2.216)	(0.004,0.202)
		CP%	0.973	0.974	0.973	0.972
		LPI	2.214	2.203	2.214	2.198
				$n=500$		
	1	HPPI	(0,1.514)	(0.002 ,1.532)	(0,1.514)	(0,1.537)
		CP%	0.967	0.967	0.967	0.968
		LPI	1.514	1.530	1.514	1.536
	2	HPPI	(0,2.044)	(0.004 , 2.039)	(0,2.044)	(0, 2.003)
		CP%	0.974	0.973	0.974	0.973
		LPI	2.044	2.035	2.044	2.002
	3	HPPI	(0,2.166)	(0.004,2.174)	(0,2.166)	(0, 2.143)
		CP%	0.974	0.974	0.974	0.974
		LPI	2.166	2.170	2.166	2.143
	4	HPPI	(0,2.21)	(0.004,2.224)	(0,2.210)	(0,2.184)
		CP%	0.974	0.974	0.974	0.974
		LPI	2.210	2.220	2.210	2.184
	5	HPPI	(0,2.227)	(0.004 , 2.240)	(0,2.227)	(0, 2.205)
		CP%	0.974	0.974	0.974	0.973
		LPI	2.227	2.236	2.227	2.205

Finally, in Tables (3), (4) and (5), we present the h-step prediction interval of $X_{N+h}^{(k)}$ in the form (X_L, X_U), where $X_L = \frac{1}{N}\sum_{k=1}^{N} X_L^{(k)}$ and $X_U = \frac{1}{N}\sum_{k=1}^{N} X_U^{(k)}$ with $X_L^{(k)}$ and $X_U^{(k)}$ being respectively the left and right interval values of k-th repetition simulation, $k = 1, \ldots, N$. In the Bayesian method X_L and X_U are based on Nm sample observations, i.e., $X_L = \frac{1}{Nm}\sum_{k=1}^{N}\sum_{i=1}^{m} X_L^{(i,k)}$ and $X_U = \frac{1}{Nm}\sum_{k=1}^{N}\sum_{i=1}^{m} X_U^{(i,k)}$, where $X_L^{(i,k)}$ and $X_U^{(i,k)}$ are respectively the left and right interval value of i-th, $i = 1, \ldots, m$, sample in Step 2 of Algorithm 2 for k-th repetition simulation. In this table, the 95% coverage probability (CP) and the length of prediction interval (LPI) are also presented.

Based on the results, we conclude that CP is close to 95% and LPI decreases as n increases, which indicate that the HPP method can produce reliable prediction interval for the PLINAR(1).

6 Real Data Analysis

In this section, we discuss the possible application of the PLINAR(1) model for anorexia real count time series data. The data gives monthly numbers of submissions to animal health laboratories from January 2003 to December 2009, from a region in New Zealand (Aghababaei Jazi, Jones and Lai (2012)), see Table (7). The sample path, autocorrelation function (ACF) and partial autocorrelation function (PACF) are illustrated in Figure (1), suggesting that, the AR(1)-type model is a suitable one to model the proposed data set. As it can be seen, the data is empirically overdispersed with mean and variance equal to 0.8214 and 2.8954. Mohammadpour et al. (2018) applied the overdispersion test described in Schweer and Weiß(2014) with the significance level $\alpha = 0.05$. They showed that the observed value of the index of dispersion exceeds the critical value; hence, the data series do not stem from an equidispersed Poisson INAR(1) process. Therefore, the Poisson-Lindley or negative binomial could appear to be more appropriate than the Poisson model for this series. Besides, they show that the skewness and kurtosis of the Poisson-Lindley distribution are smaller than those of the negative binomial distribution. The sample skewness and sample kurtosis of the data are 3.38 and 17.7, respectively. Hence, the Poisson-Lindley distribution seems to be more flexibility to model the data than the negative binomial.

The performance of the prediction methods is checked using the first 79 observations to estimate the parameters, and predicting the last 5 observations. In Table (8), we report the point predictions for the last 5 observations using the conditional mean prediction method, when the parameters are estimated using MLE, MELE and

Table 7: Anorexia data set

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2003	0	1	3	1	4	1	1	4	11	2	1	1
2004	2	2	0	0	0	0	0	0	0	0	0	0
2005	0	0	0	0	0	0	0	0	0	0	0	1
2006	0	0	0	0	0	0	3	5	6	3	2	1
2007	0	0	0	0	0	0	1	0	2	0	0	0
2008	0	0	0	0	0	0	0	2	4	0	1	0
2009	1	0	0	0	2	1	0	0	0	0	0	0

Whittle methods. The bootstrap prediction method was employed as well, when the parameters are estimated using Bootstrap and MLE. In this method, the third step was repeated 1000 times, the mean of the prediction was calculated and assumed as the prediction of data. Moreover, prediction methods based on the Bayesian method, i.e., mean, median, mode and MCMC methods, are given.

The 95% HPP intervals are calculated for all prediction methods. As we see, all intervals covered the observed data. The predictions based on Bayesian methods are nearest to the real observed values. So, we suggest using Bayesian predictors for data prediction in the PLINAR(1) process.

Table 8: Prediction analysis of Anorexia data

observed value		0	0	0	0	0
conditional mean (MELE estimation)	prediction	0.45737	0.67827	0.78496	0.83648	0.86137
	lower limit	0	0	0	0	0
	upper limit	3	3	3	3	3
conditional mean (Whittle estimation)	prediction	0.72207	1.0633	1.27008	1.36278	1.40954
	lower limit	0	0	0	0	0
	upper limit	4	4	5	5	5
conditional mean (MLE estimation)	prediction	0.58979	0.81395	0.89914	0.93152	0.94383
	lower limit	0	0	0	0	0
	upper limit	2	2	2	3	3
Bayesian	Mean Method	0.16	0.53	0.44	0.45	0.41
	Median Method	0	0	0	0	0
	Mode Method	0	0	0	0	0
	MCMC	0.00769	0.01002	0.01072	0.01094	0.01100
	lower limit	0	0	0	0	0
	upper limit	2.85	3.32	3.5	3.51	3.53
Bootstrap (Bootstrap estimation)	prediction	0.620	0.814	1.042	1.194	1.180
	lower limit	0	0	0	0	0
	upper limit	3	3	3	3	3
Bootstrap (MLE estimation)	prediction	0.706	0.873	0.972	1.030	0.984
	lower limit	0	0	0	0	0
	upper limit	3	3	3	4	4

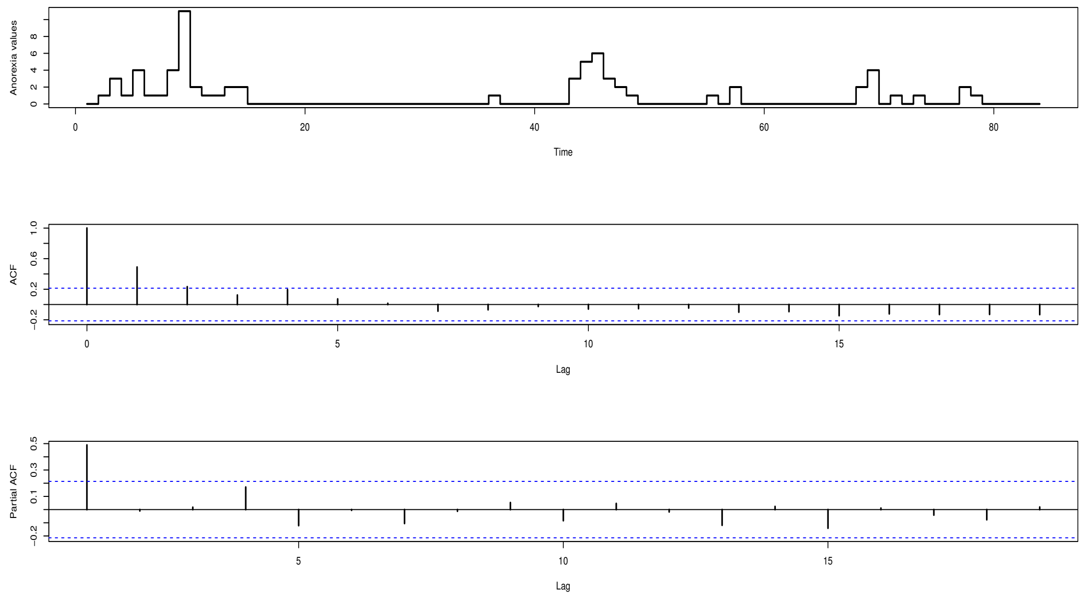

Figure 1: Sample path; ACF and PACF of Anorexia data set

References

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

Awale, M., Ramanathan, T. V., and Kale, M. (2017), Coherent forecasting in integer-valued AR (1) models with geometric marginals. *Journal of Data Science*, **15**(1), 95-114.

Bakouch, H. S., and Ristić, M. M. (2010), Zero truncated Poisson integer-valued AR (1) model. *Metrika*, **72**(2), 265-280.

Bhansali, R. (1974), Asymptotic mean-square error of predicting more than one-step ahead using the regression method. *Journal of the Royal Statistical Society: Series C*, **23**, 35–42.

Brännäs, K., and Shahiduzzaman Quoreshi, A. M. M. (2010), Integer-valued moving average modelling of the number of transactions in stocks. *Applied Financial Economics*, **20**(18), 1429-1440.

Bühlmann, P. (1997), Sieve bootstrap for time series. *Bernoulli*, **3**(2), 123-148.

Cardinal, M., Roy, R., and Lambert, J. (1999), On the application of integer-valued time series models for the analysis of disease incidence. *Statistics in Medicine*, **18**(15), 2025-2039.

Chuang, C. S., and Chan, N. H. (2002), Empirical likelihood for autoregressive models, with applications to unstable time series. *Statistica Sinica*, 387-407.

Fokianos, K., and Kedem, B. (2003), Regression theory for categorical time series. *Statistical science*, **18**(3), 357-376.

Fox, R., and Taqqu, M. S. (1986), Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, 517-532.

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

Ghitany, M. E., and Al-Mutairi, D. K. (2009), Estimation methods for the discrete Poisson-Lindley distribution. *Journal of Statistical Computation and Simulation*, **79**(1), 1-9.

Hannan, E. J. (1973), The asymptotic theory of linear time-series models. *Journal of Applied Probability*, **10**(1), 130-145.

Kim, H. Y., and Park, Y. (2008), A non-stationary integer-valued autoregressive model. *Statistical papers*, **49**(3), 485.

Livsey, J., Lund, R., Kechagias, S., and Pipiras, V. (2018), Multivariate integer-valued time series with flexible autocovariances and their application to major hurricane counts. *The Annals of Applied Statistics*, **12**(1), 408-431.

Maiti, R., and Biswas, A. (2015), Coherent forecasting for over-dispersed time series of count data. *Brazilian Journal of Probability and Statistics*, **29**(4), 747-766.

Maiti, R., Biswas, A., and Das, S. (2016), Coherent forecasting for count time series using Box-Jenkins's AR (p) model. *Statistica Neerlandica*, **70**(2), 123-145.

McCabe, B. P., Martin, G. M., and Harris, D. (2011), Efficient probabilistic forecasts for counts. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **73**(2), 253-272.

McKenzie, E. (2003), *Discrete variate time series*. In C. Rap. and D. Shanbhag, editors, Stochastic Processes: Modelling and Simulation, Handbook of Statistics, volume 21, pages 573-606. Elsevier Science B.V., North-Holland, Amsterdam.

Mohammadpour, M., Bakouch, H. S., and Shirozhan, M. (2018), Poisson-Lindley INAR (1) model with applications. *Brazilian Journal of Probability and Statistics*, **32**(2), 262-280.

Moriña, D., Puig, P., Ríos, J., Vilella, A., and Trilla, A. (2011), A statistical model for hospital admissions caused by seasonal diseases. *Statistics in medicine*, **30**(26), 3125-3136.

Mykland, P. A. (1995), Dual likelihood. *The Annals of Statistics*, 396-421.

Nastić, A. S., Laketa, P. N., and Ristić, M. M. (2016), Random environment integer-valued autoregressive process. *Journal of Time Series Analysis*, **37**(2), 267-287.

Owen, A. B. (1988), Empirical likelihood ratio confidence intervals for a single functional. *Biometrika*, **75**(2), 237-249.

Owen, A. (1990), Empirical likelihood ratio confidence regions. *The Annals of Statistics*, **18**(1), 90-120.

Pascual, L., Romo, J., and Ruiz, E. (2004), Bootstrap predictive inference for ARIMA processes. *Journal of Time Series Analysis*, **25**(4), 449-465.

Rice, J. (1979), On the estimation of the parameters of a power spectrum. *Journal of Multivariate Analysis*, **9**(3), 378-392.

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**(7), 2218-2226.

Sesay, S. A. O., and Rao, T. S. (1992), Frequency-domain estimation of bilinear time series models. *Journal of Time Series Analysis*, **13**(6), 521-545.

Silva, M. E., and Oliveira, V. L. (2005), Difference equations for the higher order moments and cumulants of the INAR (p) model. *Journal of Time Series Analysis*, **26**(1), 17-36.

Scotto, M. G., Weiß, C. H., and Gouveia, S. (2015), Thinning-based models in the analysis of integer-valued time series: a review. *Statistical Modelling*, **15**(6), 590-618.

Silva, N., Pereira, I. and Eduarda Silva, E. (2009), Forecasting in INAR(1) model. *Statistical journal*. **7**(1), 119-134.

Simarmata, D. M., Novkaniza, F. and Widyaningsih, Y. (2017), *A time series model: First-order integer-valued autoregressive (INAR(1))*. AIP Conference Proceedings 1862, 030157 (2017); doi: 10.1063/1.4991261.

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

Rao, T. S., and Chandler, R. E. (1996), *A frequency domain approach for estimating parameters in point process models*. In Athens Conference on Applied Probability and Time Series Analysis (pp. 392-405). Springer, New York, NY.

Walker, A. M. (1964), Asymptotic properties of least-squares estimates of parameters of the spectrum of a stationary non-deterministic time-series. *Journal of the Australian Mathematical Society*, **4**(3), 363-384.

Wang, Y., and Zhang, H. (2018), Some estimation and forecasting procedures in Possion-Lindley INAR (1) process. *Communications in Statistics-Simulation and Computation*, 1-14.

Weiß, C. H. (2008), Thinning operations for modeling time series of counts- a survey. *AStA Advances in Statistical Analysis*, **92**(3), 319.

Whittle, P. (1953), The analysis of multiple stationary time series. *Journal of the Royal Statistical Society: Series B (Methodological)*, **15**(1), 125-139.

Zhang, H., Wang, D., and Zhu, F. (2010), Inference for INAR (p) processes with signed generalized power series thinning operator. *Journal of Statistical Planning and Inference*, **140**(3), 667-683.

Zhang, H., Wang, D., and Zhu, F. (2011), The empirical likelihood for first-order random coefficient integer-valued autoregressive processes. *Communications in Statistics-Theory and Methods*, **40**(3), 492-509.

Zheng, H., Basawa, I. V., and Datta, S. (2006), Inference for pth-order random coefficient integer-valued autoregressive processes. *Journal of Time Series Analysis*, **27**(3), 411-440.

Zheng, H., Basawa, I. V., and Datta, S. (2007), First-order random coefficient integer-valued autoregressive processes. *Journal of Statistical Planning and Inference*, **137**(1), 212-229.