Replicated INAR(1) processes

Isabel Silva[1,4], M. Eduarda Silva[*,1,3], Isabel Pereira[2,3], Nélia Silva[2,3]

(imsilva@fc.up.pt, mesilva@fc.up.pt, isabelp@mat.ua.pt)

[1] Departamento de Matemática Aplicada, Faculdade de Ciências, Universidade do Porto, Portugal

[2] Departamento de Matemática, Universidade de Aveiro, Portugal

[3] UI&D Matemática e Aplicações, Universidade de Aveiro, Portugal

[4] Departamento de Engenharia Civil, Faculdade de Engenharia, Universidade do Porto, Portugal

Abstract

Replicated time series are a particular type of repeated measures, which consist of time-sequences of measurements taken from several subjects (experimental units). We consider independent replications of count time series that are modelled by first-order integer-valued autoregressive processes, INAR(1). In this work, we propose several estimation methods using the classical and the Bayesian approaches and both in time and frequency domains. Furthermore, we study the asymptotic properties of the estimators. The methods are illustrated and their performance is compared in a simulation study. Finally, the methods are applied to a set of observations concerning sunspot data.

Keywords: INAR Process, Replicated Time Series, Time Series Estimation, Whittle Criterion, Bayesian Estimation.

1 Introduction

Usually in time series analysis the inference is based on a single, long (or not so) time series. However, time series methodology is becoming more widely used in many areas of application where the available data consists of replicated series $\{X_{k,t}: \ k = 1, \ldots, r; \ t = 1, \ldots, n\}$ and the emphasis is on the estimation of population characteristics rather than on the behaviour of the individual series. Examples occur in experimental biology, environmental sciences and economy, with independent replicates of the same process appearing through the observation of a single series in a number of locals (the locals being sufficiently apart to be treated as independent) or the application of a treatment to a number of individuals (behaving independently of the others).

The analysis of replicated time series when the inferential focus is on the dependence of the mean response on time, experimental treatment or other explanatory variables is well documented in the literature and usually referred to as longitudinal data analysis (Diggle and al-Wasel, 1997). However, statistical analysis of replicated time series when the mean response is constant (or not of interest) and the inferential focus is on the stochastic variation about the mean has received little attention,

*Addressing author: Departamento de Matemática Aplicada, Faculdade de Ciências Universidade do Porto, Rua do Campo Alegre, 687, 4169-007 Porto, Portugal

in particular for time series of counts. In the context of replicated Gaussian time series the works of Azzalini (1981, 1984) and Dégerine (1987) should be mentioned.
The usual linear models for time series, the well known ARMA models, are suitable for modelling stationary dependent sequences under the assumption of Gaussianity, which is inappropriate for modelling counting processes. Thus, motivated by the need of modelling correlated series of counts, several models for integer valued time series were proposed in the literature, in particular the INteger-valued AutoRegressive (INAR) processes proposed by Al-Osh and Alzaid (1987) and Du and Li (1991). These processes have been considered as the discrete counter part of AR processes, but their highly nonlinear characteristics lead to some statistical challenging problems, namely in parameter estimation.
In this paper, the replicated INAR process, denoted by RINAR process, consisting of independent replications of INAR time series is considered. We address the problem of parameter estimation using several methods that can be classified into two main approaches: the estimating functions framework and Bayesian methods. The theory of estimating functions[1] proposed by Godambe (1960) provides a unified approach to the usual estimation methods in time series analysis, such as Yule-Walker equations, Conditional Least Squares, Conditional Maximum Likelihood in the time domain and the Whittle criterion in the frequency domain. Among these, the Conditional Least Squares estimators with a particular set of weights lead to optimal estimators within the class of linear estimating functions. Expressions for the asymptotic standard errors of the estimates are obtained whenever is possible and in particular, the information matrix for Conditional Maximum Likelihood is computed.
Alternatively, we consider Bayesian methods, which have been widely applied in the time series context and have played a significant role in recent developments. However, these methods have not yet been successfully applied to the INAR (and other related) processes, although Congdon (2003) refers the possibility of using the WinBugs Bayesian package for these models.
This work is organized as follows: in Section 2 we define the replicated INAR, RINAR, processes. In Section 3 we propose several estimation methods from both the classical and the Bayesian approaches and in the time and frequency domain and study the asymptotic properties of the estimators. In Section 4 we conduct a simulation study to assess and compare the performance of the small sample properties of the proposed estimators. Finally, in Section 5 we apply the RINAR model to a set of data concerning sunspot data.

2 Replicated INAR process

Consider a non negative integer-valued random variable X and $\alpha \in [0, 1]$, and define the generalized thinning operation, hereafter denoted by '$*$', as

$$\alpha * X = \sum_{j=1}^{X} Y_j, \tag{1}$$

where $\{Y_j\}, \quad j = 1, \ldots, X$, is a sequence of independent and identically distributed non-negative integer-valued random variables, independent of X, with finite mean α and variance σ^2. This sequence

[1] An estimating function, $g(y, \theta)$, is a function of the data, y, as well of the parameter, θ. An estimator is obtained by equating the estimating function to zero and solving with respect to the parameter.

is called the counting series of $\alpha * X$. Note that Steutel and Van Harn (1979) firstly defined the binomial thinning operation, in which $\{Y_j\}$ is a sequence of Bernoulli random variables. For an account of the properties of the thinning operation see Gauthier and Latour (1994) and Silva and Oliveira (2004, 2005).

A discrete time positive integer valued stochastic process, $\{X_t\}$, is said to be an INAR(p) process if it satisfies the following equation

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + \cdots + \alpha_p * X_{t-p} + e_t, \tag{2}$$

where

1. $\{e_t\}$ is a sequence of independent and identically distributed integer-valued random variables, with $\mathrm{E}[e_t] = \mu_e$, $\mathrm{Var}[e_t] = \sigma_e^2$ and $\mathrm{E}[e_t^3] = \gamma_e$;

2. all counting series of $\alpha_i * X_{t-i}$, $i = 1, \ldots, p$, $\{Y_{i,j}\}$, $j = 1, \ldots, X_{t-i}$, are mutually independent, and independent of $\{e_t\}$, and such that $\mathrm{E}[Y_{i,j}] = \alpha_i$, $\mathrm{Var}[Y_{i,j}] = \sigma_i^2$ and $\mathrm{E}[Y_{i,j}^3] = \gamma_i$;

3. $0 \leq \alpha_i < 1$, $i = 1, \ldots, p-1$, and $0 < \alpha_p < 1$.

The existence and stationarity conditions for the INAR(p) processes is that the roots of $z^p - \alpha_1 z^{p-1} - \cdots - \alpha_{p-1} z - \alpha_p = 0$ lie inside the unit circle (Du and Li, 1991) or equivalently that $\sum_{j=1}^{p} \alpha_j < 1$, (Latour, 1997, 1998). Probabilistic characteristics of the INAR models, in terms of second and third order moments and cumulants, have been obtained by Silva and Oliveira (2004, 2005).

Now, consider a replicated time series data set $\{X_{k,t} : k = 1, \ldots, r; t = 1, \ldots, n\}$, where $X_{k,t}$ denotes the kth time series observed at $t = 1, 2, \ldots, n$. We assume that all the replicates have the same length, since this seems the most common case in practice. We define a RINAR(p) model for the replicated time series $\{X_{k,t}\}$ as

$$X_{k,t} = \alpha_1 * X_{k,t-1} + \alpha_2 * X_{k,t-2} + \cdots + \alpha_p * X_{k,t-p} + e_{k,t}, \tag{3}$$

where $*$ is the (generalized) thinning operation and $\{e_{k,t}\}$ is a set of independent, integer-valued random variables with means $\mathrm{E}[e_{k,t}] = \mu_{e,k}$ and variances $\mathrm{Var}[e_{k,t}] = \sigma_{e,k}^2$.

Here, we consider only Poisson RINAR(1) processes, with $p = 1$, $\alpha_1 = \alpha \in]0, 1[$, $*$ the binomial thinning operation where the counting series, $\{Y_{i,j}^{(k)}\}$, are a set of Bernoulli random variables with $P(Y_{i,j}^{(k)} = 1) = 1 - P(Y_{i,j}^{(k)} = 0) = \alpha$, and $\{e_{k,t}\}$ is a sequence of independent Poisson distributed variables with parameter λ, independent of all counting series.

The replicated RINAR(1) process thus defined has mean and autocovariance function given by

$$\mu_X = \mathrm{E}[X_{k,t}] = \frac{\lambda}{1-\alpha}, \quad \gamma(j) = \mathrm{E}[(X_{k,t} - \mu_X)(X_{k,t+j} - \mu_X)] = \begin{cases} \dfrac{\lambda}{1-\alpha}, & \text{if } j = 0 \\ \alpha^j \dfrac{\lambda}{1-\alpha}, & \text{if } j \neq 0 \end{cases},$$

respectively. The spectral density function can be written as

$$f(\omega) = \frac{1}{2\pi} \frac{\lambda(1+\alpha)}{1 - 2\alpha\cos\omega + \alpha^2}, \qquad -\pi \leq \omega \leq \pi. \tag{4}$$

3 Estimation of the parameters

In this section we consider the estimation of the unknown parameters, $\boldsymbol{\theta} = [\alpha, \lambda]^T$, in the Poisson RINAR(1), from the observation matrix $\mathbf{X}_{r,n}$ defined as follows:

$$\mathbf{X}_{r,n} = [\mathbf{x}_{1,n}, \mathbf{x}_{2,n}, \ldots, \mathbf{x}_{r,n}]^T = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,n} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{r,1} & X_{r,2} & \cdots & X_{r,n} \end{bmatrix}. \tag{5}$$

The methods under study are the method of moments (Yule-Walker equations), Conditional Least Squares (weighted and unweighted), Conditional Maximum Likelihood and Whittle criterion, which may be included in the unifying estimating functions framework and Bayesian methodology.

3.1 Yule-Walker Estimation

Let $\hat{\gamma}_k(j) = \frac{1}{n}\sum_{t=1}^{n-j}(X_{k,t} - \overline{X}_{r,n})(X_{k,t+j} - \overline{X}_{r,n}), j \in \mathbb{Z}$, be the sample autocovariance function of the kth replicate, $\mathbf{x}_{k,n}$, where $\overline{X}_{r,n} = \frac{1}{nr}\sum_{k=1}^{r}\sum_{t=1}^{n} X_{k,t}$ is the overall sample mean, and let $\hat{\rho}_k(j) = \hat{\gamma}_k(j)/\hat{\gamma}_k(0)$ be the corresponding sample autocorrelation function.
Under our hypothesis, we incorporate the extra information brought on by the replicates, averaging over the replicates the sample functions, obtaining

$$\overline{\gamma}(j) = \frac{1}{r}\sum_{k=1}^{r}\hat{\gamma}_k(j) = \frac{1}{nr}\sum_{k=1}^{r}\sum_{t=1}^{n-j}(X_{k,t} - \overline{X}_{r,n})(X_{k,t+j} - \overline{X}_{r,n}), \qquad \overline{\rho}(j) = \frac{\overline{\gamma}(j)}{\overline{\gamma}(0)}.$$

Thus, for r replicated Poisson RINAR(1) process, the Yule-Walker (method of moments) estimate of α can be written as

$$\hat{\alpha}_{YW} = \frac{\overline{\gamma}(1)}{\overline{\gamma}(0)} = \overline{\rho}(1) = \frac{\sum_{k=1}^{r}\sum_{t=1}^{n-1}(X_{k,t} - \overline{X}_{r,n})(X_{k,t+1} - \overline{X}_{r,n})}{\sum_{k=1}^{r}\sum_{t=1}^{n}(X_{k,t} - \overline{X}_{r,n})^2}, \tag{6}$$

and an estimator of λ is given by

$$\hat{\lambda}_{YW} = \overline{X}_{r,n}(1 - \hat{\alpha}_{YW}). \tag{7}$$

According to Du and Li (1991), $\hat{\gamma}_k(j)$ and $\hat{\rho}_k(j)$ are strongly consistent. Therefore, $\overline{\gamma}(j)$ and $\overline{\rho}(j)$ and consequently $\hat{\alpha}_{YW}$ and $\hat{\lambda}_{YW}$ are also strongly consistent estimators. The estimators $\hat{\alpha}_{YW}$ and $\hat{\lambda}_{YW}$ are asymptotically unbiased normally distributed, with respect to n, with variances given by (I. Silva, 2005)

$$\mathrm{Var}[\sqrt{nr}\,\hat{\alpha}_{YW}] = \frac{\alpha(1-\alpha)}{\mu_X} + (1-\alpha)^2, \tag{8}$$

$$\mathrm{Var}[\sqrt{nr}\,\hat{\lambda}_{YW}] = \mu_X(1-\alpha)\left((1+\alpha)(1+\mu_X) + \alpha\right). \tag{9}$$

This result generalizes the work of Park and Oh (1997). Thus, as expected, the replicated observations lead to a variance reduction of the estimators of order $1/r$.

3.2 Conditional Least Squares Estimation

The Conditional Least Squares (CLS) method, proposed by Klimko and Nelson (1978), has been widely used in the time series context and, in particular, for estimating the parameters of INAR processes (Du and Li, 1991). Its application to the estimation of the parameters of the RINAR(1) model is straightforward and is described in Section 3.2.1. However, the fact that the conditional variance of the RINAR(1) process given by

$$V(\boldsymbol{\theta}, X_{k,t-1}) = \mathrm{Var}[X_{k,t}|X_{k,t-1}] = \alpha(1-\alpha)X_{k,t-1} + \lambda. \tag{10}$$

is not constant over time, suggests that we also consider Iterative Weighted Conditional Least Squares estimation, IWCLS. Moreover, since

$$g(\boldsymbol{\theta}, F_{k,t-1}) = \mathrm{E}[X_{k,t}|X_{k,t-1}] = \alpha X_{k,t-1} + \lambda, \tag{11}$$

there is a linear relationship between the conditional mean and variance of the process and IWCLS is a quasi-likelihood estimation method in the sense of Wedderburn (1974). IWCLS estimators are discussed in Section 3.2.2.

3.2.1 CLS

Let $\mathbf{x}_{k,n}$ be the kth replicate INAR(1) process with parameter vector $\boldsymbol{\theta}$ and let $F_{k,t} = \mathcal{F}(X_{k,1}, \ldots, X_{k,t})$ be the σ-algebra generated by $\{X_{k,1}, \ldots, X_{k,t}\}$. As we have seen, the conditional mean of $X_{k,t}$ given $F_{k,t-1}$, is defined in (11) by

$$g(\boldsymbol{\theta}, F_{k,t-1}) = \mathrm{E}[X_{k,t}|F_{k,t-1}] = \alpha X_{k,t-1} + \lambda.$$

The, the CLS estimator of the parameter vector $\boldsymbol{\theta}$ is obtained minimizing

$$Q(\boldsymbol{\theta}) = \sum_{k=1}^{r}\sum_{t=2}^{n}(X_{k,t} - g(\boldsymbol{\theta}, F_{k,t-1}))^2 = \sum_{k=1}^{r}\sum_{t=2}^{n}(X_{k,t} - \alpha X_{k,t-1} - \lambda)^2.$$

Therefore, given r replicates of a Poisson RINAR(1) process in the matrix of observations $\mathbf{X}_{r,n}$, defined in (5), the CLS estimators of α and λ, are given by

$$\hat{\alpha}_{CLS} = \frac{(n-1)r\sum_{k=1}^{r}\sum_{t=2}^{n}X_{k,t}X_{k,t-1} - (\sum_{k=1}^{r}\sum_{t=2}^{n}X_{k,t})(\sum_{k=1}^{r}\sum_{t=2}^{n}X_{k,t-1})}{(n-1)r\sum_{k=1}^{r}\sum_{t=2}^{n}X_{k,t-1}^2 - (\sum_{k=1}^{r}\sum_{t=2}^{n}X_{k,t-1})^2}, \tag{12}$$

$$\hat{\lambda}_{CLS} = \frac{\sum_{k=1}^{r}\sum_{t=2}^{n}X_{k,t} - \hat{\alpha}_{CLS}\sum_{k=1}^{r}\sum_{t=2}^{n}X_{k,t-1}}{(n-1)r}. \tag{13}$$

It can easily be seen that the function g given by (11) is such that $\partial g/\partial\alpha, \partial g/\partial\lambda$ and $\partial^2 g/\partial\alpha\partial\lambda$ satisfy all the regularity conditions of Theorem 3.1 in Klimko and Nelson (1978). Therefore, the CLS estimators for Poisson RINAR(1) processes are strongly consistent. Moreover, since $\mathrm{E}[|e_{k,t}|^3] < \infty$ for the Poisson distribution, by Theorem 3.2 of Klimko and Nelson (1978), it follows that the CLS estimators for the Poisson RINAR(1) process are asymptotically normally distributed

$$\sqrt{nr}\left(\hat{\boldsymbol{\theta}}_{CLS} - \boldsymbol{\theta}\right) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \mathbf{V}^{-1}\mathbf{W}\mathbf{V}^{-1}), \tag{14}$$

with $\mathbf{V} = [V_{ij}]$ and $\mathbf{W} = [W_{ij}]$ defined as

$$V_{ij} = E\left[\frac{\partial g(\boldsymbol{\theta}, F_{k,t-1})}{\partial \theta_i}\frac{\partial g(\boldsymbol{\theta}, F_{k,t-1})}{\partial \theta_j}\right],\ i,j = 1,2, \tag{15}$$

$$W_{ij} = E\left[u_2^2(\boldsymbol{\theta})\frac{\partial g(\boldsymbol{\theta}, F_{k,t-1})}{\partial \theta_i}\frac{\partial g(\boldsymbol{\theta}, F_{k,t-1})}{\partial \theta_j}\right],\ i,j = 1,2, \tag{16}$$

and where $u_2(\boldsymbol{\theta}) = X_{k,2} - \alpha X_{k,1} - \lambda$ is the one-step-ahead linear prediction error (I. Silva, 2005). Also in this case, the variance of the estimators is reduced by a factor of $1/r$ due to the presence of the replicates.

3.2.2 IWCLS

The IWCLS estimator of the parameter vector $\boldsymbol{\theta}$ is obtained minimizing the sum of the squared error between each observation and its conditional mean, $(X_{k,t} - g(\boldsymbol{\theta}, X_{k,t-1}))^2$, weighted by the inverse of the conditional variance, $1/\hat{V}(\hat{\boldsymbol{\theta}}, X_{k,t-1})$.
Thus, the IWCLS estimate of $\boldsymbol{\theta}$ is obtained by minimizing iteratively $Q_W(\boldsymbol{\theta})$, defined as

$$Q_W(\boldsymbol{\theta}) = \sum_{k=1}^{r}\sum_{t=2}^{n}\frac{(X_{k,t} - \alpha X_{k,t-1} - \lambda)^2}{\hat{\alpha}(1-\hat{\alpha})X_{k,t-1} + \hat{\lambda}}, \tag{17}$$

As initial values, $\hat{\alpha}^{(0)}$ and $\hat{\lambda}^{(0)}$, for $\hat{\alpha}$ and $\hat{\lambda}$, respectively, we choose a set of consistent estimates for instance, Conditional Least Squares estimates. At iteration i the weights in (17) are updated with $\hat{\alpha}^{(i-1)}$ and $\hat{\lambda}^{(i-1)}$, and new estimates of the parameters, $\hat{\alpha}^{(i)}$ and $\hat{\lambda}^{(i)}$, are successively obtained, until convergence is achieved.
The structural relationship between INAR(1) processes and subcritical branching processes (Dion et al, 1995), suggests the use of the Weighted Conditional Least Squares method (Winnicki, 1988; Wei and Winnicki, 1989), with a set of weights given by $\sqrt{1 + X_{k,t-1}}$. However, the former approach proposed here is to be preferred since it can be proved that the associated estimating function $\Psi(\boldsymbol{\theta}) = \frac{\partial Q_W}{\partial \boldsymbol{\theta}}$, is an unbiased and regular estimating function. Therefore, $\hat{\boldsymbol{\theta}}_{IWCLS}$ is an optimal estimator within the class of linear estimating functions.
Brännäs (1995) stated that for one replicate of a stationary Poisson INAR(1) process, the consistency and asymptotic normality of the IWCLS estimator follow directly from the work of Godambe (1960); Wooldridge (1994). Moreover, Freeland and McCabe (2004) has obtained an approximate expression for the asymptotic variance of the IWCLS estimator, which is correct up to a constant. Thus, these properties are easily extended to the IWCLS estimator in the Poisson RINAR model framework.

3.3 Conditional Maximum Likelihood Estimation

The conditional likelihood function of the r replicates from a Poisson INAR(1) process is the convolution of the distribution of the innovation process and that resulting from the binomial thinning operation, $\mathcal{B}i(\alpha, X_{t-1})$ (Johnson and Kotz, 1969; Al-Osh and Alzaid, 1987; Freeland and McCabe, 2004). Thus, given an initial value $\mathbf{x}_{k,1} = [X_{1,1}, X_{2,1}, \ldots, X_{r,1}]$, the conditional likelihood function of the RINAR(1) process is given by the following expression

$$L(\mathbf{X}_{r,n}, \boldsymbol{\theta}|\mathbf{x}_{k,1}) = \prod_{k=1}^{r}\prod_{t=2}^{n} P(X_{k,t}|X_{k,t-1}), \tag{18}$$

where

$$P(X_{k,t}|X_{k,t-1}) = \mathrm{e}^{-\lambda} \sum_{i=0}^{M_{k,t}} \frac{\lambda^{(X_{k,t})-i}}{((X_{k,t})-i)!} \binom{X_{k,t-1}}{i} \alpha^i (1-\alpha)^{(X_{k,t-1})-i},$$

with $M_{k,t} = \min\{X_{k,t-1}, X_{k,t}\}$.

The Conditional Maximum Likelihood (CML) estimator, $\hat{\boldsymbol{\theta}}$, is obtained maximizing $L(\mathbf{X}_{r,n}, \boldsymbol{\theta}|\mathbf{x}_{k,1})$, or equivalently, the conditional log-likelihood function

$$\ell(\mathbf{X}_{r,n}, \boldsymbol{\theta}|\mathbf{x}_{k,1}) = \sum_{k=1}^{r}\sum_{t=2}^{n} \log P(X_{k,t}|X_{k,t-1}).$$

Let $H_k(t) = P(X_{k,t}-1|X_{k,t-1})/P(X_{k,t}|X_{k,t-1})$, then the derivatives of $\ell(\mathbf{X}_{r,n}, \boldsymbol{\theta}|\mathbf{x}_{k,1})$ are given by

$$\ell'_{\alpha} = \frac{\partial \ell(\mathbf{X}_{r,n}, \boldsymbol{\theta}|\mathbf{x}_{k,1})}{\partial \alpha} = \sum_{k=1}^{r}\sum_{t=2}^{n} \left[X_{k,t} - \alpha X_{k,t-1} - \lambda H_k(t)\right]/\alpha(1-\alpha), \tag{19}$$

$$\ell'_{\lambda} = \frac{\partial \ell(\mathbf{X}_{r,n}, \boldsymbol{\theta}|\mathbf{x}_{k,1})}{\partial \lambda} = \sum_{k=1}^{r}\sum_{t=2}^{n} H_k(t) - (n-1)r. \tag{20}$$

The CML estimates satisfy the following equation, obtained cancelling the derivatives (19) and (20).

$$\sum_{k=1}^{r}\sum_{t=2}^{n} X_{k,t} - \hat{\alpha}_{CML} \sum_{k=1}^{r}\sum_{t=2}^{n} X_{k,t-1} = (n-1)r\hat{\lambda}_{CML}. \tag{21}$$

If we eliminate one of the parameters in (21), say $\hat{\alpha}_{CML}$, then ℓ'_{λ} in (20) can be written as a function of only λ and the estimate $\hat{\lambda}_{CML}$ can be found by iterating ℓ'_{λ}.

Franke and Seligmann (1993) and Franke and Subba Rao (1995) have shown that, for stationary Poisson INAR(1) processes and under some regularity conditions (that are satisfied by the Poisson law), the CML estimates are consistent and asymptotically normal. Since these properties are easily extended to the estimators of the Poisson RINAR model parameters we may write

$$\sqrt{rn} \begin{pmatrix} \hat{\alpha} - \alpha \\ \hat{\lambda} - \lambda \end{pmatrix} \xrightarrow{d} \mathcal{N}\left(\mathbf{0}, \mathbf{i}^{-1}\right), \tag{22}$$

where $\mathbf{i}$ is the Fisher information matrix whose elements are the expectation of the second-order derivatives of the log-likelihood function of the process, given by the following expressions (N. Silva, 2005)

$$\begin{aligned}
\ell''_{\lambda\lambda} =& \sum_{k=1}^{r}\sum_{t=2}^{n} \left\{ \frac{P(X_{k,t}-2|X_{k,t-1})}{P(X_{k,t}|X_{k,t-1})} - \left(\frac{P(X_{k,t}-1|X_{k,t-1})}{P(X_{k,t}|X_{k,t-1})}\right)^2 \right\}, \\
\ell''_{\alpha\lambda} =& \sum_{k=1}^{r}\sum_{t=2}^{n} \left\{ \frac{X_{k,t-1}P(X_{k,t}-2|X_{k,t-1}-1)}{P(X_{k,t}|X_{k,t-1})} - \frac{X_{k,t-1}P(X_{k,t}-1|X_{k,t-1})}{P(X_{k,t}-1|X_{k,t-1}-1)} \right\}, \\
\ell''_{\alpha\alpha} =& \frac{1}{(1-\alpha)^2} \sum_{k=1}^{r}\sum_{t=2}^{n} \left\{ \frac{2X_{k,t-1}P(X_{k,t}-1|X_{k,t-1}-1)}{P(X_{k,t}|X_{k,t-1})} - X_{k,t-1} \right. \\
& + \frac{X_{k,t-1}(X_{k,t-1}-1)P(X_{k,t}-2|X_{k,t-1}-2)}{P(X_{k,t}|X_{k,t-1})} \\
& \left. - \left(\frac{X_{k,t-1}P(X_{k,t}-1|X_{k,t-1}-1)}{P(X_{k,t}|X_{k,t-1})}\right)^2 \right\}.
\end{aligned}$$

Noting that the second derivatives are functions of $(X_{k,t}, X_{k,t-1})$, we obtain for the elements of $\mathbf{i}$

$$
\begin{aligned}
\mathrm{E}[\ell''_{\lambda\lambda}] &= \sum_{k=1}^{r}\sum_{t=2}^{n} \mathrm{E}[h(X_{k,t}, X_{k,t-1})] \\
&= \sum_{k=1}^{r}\sum_{t=2}^{n} \sum_{x_{k,t}=0}^{+\infty} \sum_{x_{k,t-1}=0}^{+\infty} h(X_{k,t}, X_{k,t-1}) P(X_{k,t} = x_t, X_{k,t-1} = x_{t-1}) \\
&= (n-1)r \sum_{x_{k,t}=0}^{+\infty} \sum_{x_{k,t-1}=0}^{+\infty} P(X_{k,t-1} = x_{t-1}) \left\{ P(X_{k,t} - 2 | X_{k,t-1}) - \frac{P(X_{k,t} - 1 | X_{k,t-1})^2}{P(X_{k,t} | X_{k,t-1})} \right\};
\end{aligned}
\tag{23}
$$

similarly,

$$
\begin{aligned}
\mathrm{E}[\ell''_{\alpha\lambda}] =& \frac{(n-1)r}{1-\alpha} \sum_{x_{k,t}=0}^{+\infty} \sum_{x_{k,t-1}=0}^{+\infty} X_{t-1} P(X_{k,t-1} = x_{t-1}) \Big\{ P(X_{k,t} - 2 | X_{k,t-1} - 1) \\
& - \frac{P(X_{k,t} - 1 | X_{k,t-1}) P(X_{k,t} - 1 | X_{k,t-1} - 1)}{P(X_{k,t} | X_{k,t-1})} \Big\},
\end{aligned}
\tag{24}
$$

and

$$
\begin{aligned}
\mathrm{E}[\ell''_{\alpha\alpha}] =& \frac{(n-1)r}{(1-\alpha)^2} \sum_{x_{k,t}=0}^{+\infty} \sum_{x_{k,t-1}=0}^{+\infty} P(X_{k,t-1} = x_{t-1}) \Big\{ \frac{2X_{k,t-1} P(X_{k,t} - 1 | X_{k,t-1} - 1)}{P(X_{k,t} | X_{k,t-1})} - X_{k,t-1} \\
& + \frac{X_{k,t-1}(1 - X_{k,t-1}) P(X_{k,t} - 2 | X_{k,t-1} - 2)}{P(X_{k,t} | X_{k,t-1})} \\
& - \frac{X_{k,t-1} P(X_{k,t} - 1 | X_{k,t-1} - 1)}{P(X_{k,t} | X_{k,t-1})} \Big\}.
\end{aligned}
\tag{25}
$$

The elements of matrix $\mathbf{i}$ are calculated truncating the infinite sums to m, which corresponds to substituting the sample space, $\{0, 1, \ldots\}$ of $X_{k,t}$ by the sample space $\{0, 1, \ldots, m\}$. The value for m is selected so that $P(X_t > m) < 10^{-15}$. These elements may also be computed using numerical derivatives.

3.4 Whittle Estimation

In this section we consider a frequency domain estimation procedure based on the Whittle criterion. This approach was originally proposed by Whittle (1953, 1954) for Gaussian processes and further investigated by several authors (Walker, 1964; Hannan, 1973; Rice, 1979; Dzhaparidze and Yaglom, 1983). It has been used in many situations: Fox and Taqqu (1986); Sesay and Subba Rao (1992); Subba Rao and Chandler (1996) and Silva and Oliveira (2004, 2005). The main motivation for the Whittle criterion is the fact that the spectral density function of a process may be easy to obtain whereas an exact likelihood may not. Thus, Whittle proposed to represent the likelihood of a (Gaussian) stochastic process via its spectral properties.

Although the Whittle criterion is usually considered an approximation to a Gaussian likelihood, it may also be obtained as an approximation for the likelihood function of collections of sample Fourier coefficients for several classes of processes, namely for *non-Gaussian mixing processes* (Dzhaparidze

and Yaglom, 1983; Chandler, 1997). Thus, the use of Whittle estimation in the context of RINAR processes is justified by the fact that these processes belong to the class of *non-Gaussian mixing processes*, as we show in the following. It suffices to argue the proof for INAR processes since RINAR processes are independent repetitions of these processes.
A stochastic process $\{X_t\}$ belongs to the class of *non-Gaussian mixing processes* if the following conditions are satisfied:

(NGMP1) X_t is strictly stationary;

(NGMP2) X_t has finite absolute moments of all orders, i.e.

$$\mathrm{E}[|X_t|^k] < \infty, \quad t \in \mathbb{Z}, k \in \mathbb{N};$$

(NGMP3) Let $C_k(s_1, \ldots, s_{k-1})$ be the kth-order cumulant of the X_t process, then

$$\sum_{s_1=-\infty}^{\infty} \cdots \sum_{s_{k-1}=-\infty}^{\infty} |C_k(s_1, \ldots, s_{k-1})| < \infty, \quad k = 2, 3, \ldots$$

Note that (**NGMP3**) is a *mixing* condition on X_t that guarantees a fast decrease of the statistical dependence between X_t and X_{t+s} as $s \to \infty$.
Now, condition (**NGMP1**) follows from Corollary 1 of Dion et al (1995), which states that a stationary INAR(p) process is strictly stationary. To prove condition (**NGMP3**) it is sufficient to prove that an INAR process is strongly mixing. Well, the INAR(p) process defined in (2) may be written as a p-dimensional INAR(1) process and, moreover if $0 < \alpha_i < 1$ for $i = 1, \ldots, p$, and $0 < P(e_t = 0) < 1$, then any solution of the equation satisfied by the p-dimensional INAR(1) process is an irreducible and aperiodic Markov chain on ${\mathbb{N}_0}^p$ (Lemma 3 of Franke and Subba Rao (1995)). Since a Markov chain is irreducible and aperiodic if and only if it is strongly mixing (Rosenblatt, 1971, p. 207), we obtain that the INAR is strongly mixing and therefore satisfies condition (**NGMP3**). Finally, since the absolute cumulants are summable, all the cumulants of the process exist and are finite. Therefore, the moments of all orders of an INAR process exist and are finite because the existence of the cumulants is equivalent to the existence of the moments (Rosenblatt, 1983). Thus, the condition (**NGMP2**) is satisfied by INAR models.
Now, if a model is a non-Gaussian mixing process then the periodogram ordinates, $I(\cdot)$, at the Fourier frequencies, $\omega_j = 2\pi j/n$, $j = 1, \ldots, [n/2]$, are asymptotically independent random variables, distributed as $f(\omega_j)\chi_2^2/2$ variates, where $f(\cdot)$ is the spectral density function of the process (Brillinger, 2001, p. 126). Then, the probability density of the variables $I(\omega_j)$, denoted by $p_I(I(\omega_j))$, $j = 1, \ldots, [(n-1)/2]$, is asymptotically given by

$$\begin{aligned} p_I &= \prod_{j=1}^{[(n-1)/2]} \frac{1}{f(\omega_j)} \exp \frac{-I(\omega_j)}{f(\omega_j)}, \\ \ell_I &= \log(p_I) = -\sum_{j=1}^{[(n-1)/2]} \left(\log(f(\omega_j)) + \frac{I(\omega_j)}{f(\omega_j)} \right). \end{aligned} \tag{26}$$

The last equation, (26), is a discrete version of the Whittle criterion, up to a constant.

Thus, for RINAR processes, we obtain the Whittle estimate of $\boldsymbol{\theta}$ by minimizing

$$\tilde{\boldsymbol{\ell}}(\mathbf{X}_{r,n}) = \frac{1}{n}\sum_{k=1}^{r}\sum_{j=1}^{[n/2]}\left(\log f(\omega_j) + \frac{I_k(\omega_j)}{f(\omega_j)}\right) = \frac{r}{n}\sum_{j=1}^{[n/2]}\left(\log f(\omega_j) + \frac{\bar{I}(\omega_j)}{f(\omega_j)}\right), \tag{27}$$

where $f(\omega_j)$ is the value of the spectral density function at the Fourier frequency $\omega_j = 2\pi j/n$, for $j = 1, \ldots, [n/2]$ and $\bar{I}(\omega_j)$ is the sample mean periodogram ordinate at the same frequency,

$$\bar{I}(\omega) = \frac{1}{r}\sum_{k=1}^{r} I_k(\omega) = \frac{1}{2\pi n r}\sum_{k=1}^{r}\left|\sum_{t=1}^{n} X_{k,t}\mathrm{e}^{i\omega t}\right|^2.$$

Dzhaparidze and Yaglom (1983) proved the consistency and asymptotic normality of Whittle estimators for non-Gaussian mixing processes. However, the asymptotic variance of $(\hat{\boldsymbol{\theta}}_{WHT} - \boldsymbol{\theta})$ depends on the fourth-order cumulant spectral density function, that is very difficult to obtain.

3.5 Bayesian Estimation

In this section, we consider a Bayesian analysis of the parameters of the RINAR(1) model. For this analysis prior distributions of the parameters α and λ are needed. In the context of the RINAR(1) model under study, we consider the conjugates of the Binomial and Poisson distributions and thus, $\alpha \frown Beta(a, b)$, $a, b > 0$ and $\lambda \frown Gamma(c, d)$, $c, d > 0$. Assuming independence between α and λ, the prior distribution of (α, λ) is proportional to

$$p(\alpha, \lambda) \propto \lambda^{c-1}\exp(-d\lambda)\alpha^{a-1}(1-\alpha)^{b-1}, \quad \lambda > 0, 0 < \alpha < 1, \tag{28}$$

where a, b, c and d are known parameters. Note that, as $a \to 0$, $b \to 0$ $c \to 0$ and $d \to 0$ we have a vague prior distribution.
The posterior distribution of (α, λ) can be written as

$$\begin{aligned} p(\lambda, \alpha|\mathbf{X}_{r,n}) \propto & L(\mathbf{X}_{r,n}, \boldsymbol{\theta}|\mathbf{x}_{k,1})\, p(\lambda, \alpha) \\ = & \exp\left[-(d + (n-1)r)\lambda\right]\lambda^{c-1}\alpha^{a-1}(1-\alpha)^{b-1} \\ & \prod_{k=1}^{r}\prod_{t=2}^{n}\sum_{i=0}^{M_{k,t}} \frac{\lambda^{(X_{k,t})-i}}{((X_{k,t})-i)!}\binom{X_{k,t-1}}{i}\alpha^i(1-\alpha)^{(X_{k,t-1})-i}, \end{aligned} \tag{29}$$

where $L(\mathbf{X}_{r,n}|\mathbf{x}_{k,1})$ is given by (18) and $p(\alpha, \lambda)$ by (28). The complexity of $p(\alpha, \lambda|\mathbf{X}_{r,n})$ does not allow us to get the marginal distribution of each of the unknown parameters and thus we cannot calculate the posterior mean value of α and λ. Thus, we use a Markov Chain Monte Carlo (MCMC) methodology to sample from (29). For the Gibbs sampling algorithm (Gelfand and Smith, 1990), we need to derive the full conditional posterior distribution of each unknown variable. Thus, using the expression (29), the full conditional of λ is given by

$$p(\lambda|\alpha, \mathbf{X}_{r,n}) = \frac{p(\lambda, \alpha|\mathbf{X}_{r,n})}{p(\alpha|\mathbf{X}_{r,n})} \propto \exp[-(d + (n-1)r)\lambda]\lambda^{c-1}\prod_{k=1}^{r}\prod_{t=2}^{n}\sum_{i=0}^{M_{k,t}} C(k, t, i)\lambda^{(X_{k,t})-i}, \tag{30}$$

where

$$C(k, t, i) = \frac{1}{((X_{k,t})-i)!}\binom{X_{k,t-1}}{i}\alpha^i(1-\alpha)^{(X_{k,t-1})-i} \quad \text{and} \quad \lambda > 0.$$

Proceeding in a similar way it can be shown that the full conditional distribution of α is

$$p(\alpha|\lambda, \mathbf{X}_{r,n}) = \frac{p(\lambda, \alpha|\mathbf{X}_{r,n})}{p(\lambda|\mathbf{X}_{r,n})} \propto \alpha^{a-1}(1-\alpha)^{b-1} \prod_{k=1}^{r} \prod_{t=2}^{n} \sum_{i=0}^{M_{k,t}} K(k,t,i)\alpha^{i}(1-\alpha)^{(X_{k,t-1})-i}, \tag{31}$$

where

$$K(k,t,i) = \frac{\lambda^{(X_{k,t})-i}}{((X_{k,t})-i)!} \binom{X_{k,t-1}}{i} \quad 0 < \alpha < 1.$$

It is interesting to note that when a gamma prior is used for λ, the full conditional posterior density function of λ is a linear combination of gamma densities and if a beta distribution for α is considered, the full conditional distribution of α, is a linear combination of beta densities.

4 Monte Carlo simulation study

The purpose of the simulation study presented in this section is twofold: to study and compare the small sample properties of the different estimators and to assess the effect of the replicates in the estimates.

We consider $r = 1, 10, 20$ replicates of time series of $n = 25, 50, 100$ observations, generated by INAR(1) models for the following set of parameters values $\alpha = 0.1, 0.3, 0.7, 0.9$ and $\lambda = 1$ and 3. For every possible combination of the parameters α and λ, 500 sets of r replicates of length n are simulated and the sample mean, variance and mean squared error of the estimates are calculated. The main reason for choosing a Monte Carlo study based on 500 repetitions is the extremely large amount of time need for the computation of Bayes estimates.

The asymptotic variance of YW, CLS and CML estimators, as given by equations (8), (9), (14), (15), (16), (22), (23), (24) and (25), is also provided for comparison purposes.

The Yule-Walker estimates (YW) for α and λ are obtained from equations (6) and (7) and the Conditional Least Squares estimates (CLS) are calculated from the normal equations given in (12) and (13). The Iterative Weighted Conditional Least Squares estimates (IWCLS) are computed as described in section 3.2.2 and using the MATLAB function *lsqnonneg* to minimize (17). The Whittle estimates (WHT) of α and λ, are obtained using the constrained minimization algorithm implemented in the MATLAB function *fmincon*. This algorithm finds a constrained minimum of a function of several variables (here the function is given in (27)) by a Sequential Quadratic Programming method. The CLS estimates are chosen as initial values for the algorithm. The constraints considered are $0 < \alpha < 1$ and $\lambda > 0$. The Conditional Maximum Likelihood estimates (CML) of the parameters α and λ are computed from equation (21), as explained in Section 3.3, and using a bisection method to find the zero solution of (20). To calculate the Bayesian estimates (Bayes), we run the Gibbs sampler algorithm with initial value $\alpha = \alpha_{CLS}$. In order to sample from full conditionals which are not log-concave densities, we have to use the Adaptive Rejection Metropolis Sampling -ARMS- (Gilks and Best, 1987), inside the Gibbs sampler. To reduce autocorrelation between MCMC samples, we considered only samples from every 20 iterations. Among these, we ignored the first 1100 samples as burn-in time, and use 2000 samples after the burn-in for posterior inference. In order to use vague prior distributions we considered all the hyperparameters $a, b, c, d = 10^{-4}$.

Note that YW, CLS and IWCLS estimates for α do take values outside the admissible range $[0, 1]$ when α lies near zero or one. A constrained minimization algorithm was used with the Whittle criterion to avoid this. On the other hand, CML and Bayes estimates always lie in the admissible range.

Numerical results are presented only for the models with $\alpha = 0.1, 0.3, 0.9$ and $\lambda = 1$, in Tables 1 to 6, since these illustrate well the following overall conclusions.

The estimates for α and λ present sample mean biases and variances which decrease both with the sample size n and the number of replicates, r, in agreement with the asymptotic properties of the estimators: unbiasedness and consistency.

Also, it can be noted that the absolute sample biases are larger for larger values of α and λ, and for a fixed α, the sample variance of $\hat{\lambda}$ increase with λ. Furthermore, in general, $\hat{\alpha}$ presents negative sample mean biases for all the estimation methods regardless of the size and number of replicates, indicating that α is underestimated, whilst the estimates for λ shows positive sample biases, indicating overestimation for λ.

The YW estimates, among all the methods, present the larger sample biases. For $\alpha < 0.5$, the CLS, IWCLS, CML and WHT estimates of α and λ present the lower sample mean biases, while for $\alpha > 0.5$, the lower sample mean bias is presented by CML and Bayes. The sample variance of the Bayes estimates is the lowest, among all the methods.

The root mean squared errors of the estimates are close to the corresponding standard deviations, indicating small biases.

Generally the asymptotic and the sample standard deviations of the estimators are comparable and are, in fact, quite similar for larger values of n and/or r. However, it is noticeable that CML asymptotics are rather conservative, except for $\hat{\alpha}$ when α is large.

Boxplots of the sample bias are presented in Figures 1 to 3. The boxplots indicate that the marginal distributions of the estimators are, generally, symmetric in agreement to the theoretical results. However, for small sample sizes there is evidence of departure from symmetry in the marginal distributions, specially for values of the parameters near the non-stationary region.

The above conclusions are the same for other values of the parameter λ.

5 Example

Sunspots are magnetic regions on the Sun that appear as dark group of spots on its surface with many shapes and forms. The spots change from day to day, even from hour to hour, and vary in size, from small dot (pores) to large spots groups covering a vast area of the solar surface, which after a time get smaller and disappear. The time from birth to death of a sunspot group varies from a few days to six months, with the median less than two weeks.

Sporadic naked-eye observations exist in Chinese dynastic histories since 28 BC. Telescopic observations of sunspots have been made in Europe since 1610 AD. Modern systematic measurements of sunspots began in 1835. In order to quantify the results of the observations, Rudolf Wolf introduced, in 1848, the Relative Sunspot Numbers (now referred to as the International Sunspot Numbers) as a measure of sunspots activity. Recently, Hoyt and Schatten (1998) have introduced the Group Sunspot Number, that uses the number of sunspot groups observed, rather than groups and individual sunspots.

Here, we consider number of sunspot groups available on-line at the National Geophysical Data Cen-

			$\hat{\alpha}$					
$(\alpha, \lambda) = (0.1, 1.0)$	r	n	YW	CLS	IWCLS	CML	WHT	Bayes
sample bias	1	25	-0.0538	0.0052	0.0601	0.0873	0.0300	0.0430
		50	-0.0271	0.0012	0.0326	0.0427	0.0196	0.0105
		100	-0.0121	-0.0043	0.0108	0.0219	0.0026	-0.0122
	10	25	-0.0109	-0.0033	-0.0023	0.0015	-0.0027	-0.0289
		50	-0.0105	-0.0026	-0.0077	-0.0059	-0.0026	-0.0283
		100	-0.0039	0.0020	-0.0029	-0.0031	0.0017	-0.0165
	20	25	-0.0097	-0.0025	-0.0053	-0.0040	-0.0037	-0.0268
		50	-0.0041	-0.0006	-0.0022	-0.0028	-0.0013	-0.0167
		100	-0.0025	-0.0014	-0.0015	-0.0015	-0.0016	-0.0068
sample standard deviation (theoretical standard deviation)	1	25	0.1839	0.1342	0.1162	0.1330	0.1470	0.0800
			(0.1888)	(0.2070)		(0.1987)		
		50	0.1458	0.1049	0.0996	0.1010	0.1127	0.0693
			(0.1335)	(0.1464)		(0.1425)		
		100	0.1029	0.0849	0.0799	0.0762	0.0883	0.0566
			(0.0944)	(0.1035)		(0.1012)		
	10	25	0.0662	0.0600	0.0615	0.0539	0.0624	0.0458
			(0.0597)	(0.0655)		(0.0642)		
		50	0.0461	0.0480	0.0453	0.0412	0.0490	0.0412
			0.0422	0.0463		0.0454		
		100	0.0322	0.0346	0.0325	0.0332	0.0346	0.0374
			0.0298	0.0327		0.0321		
	20	25	0.0454	0.0447	0.0457	0.0412	0.0458	0.0412
			(0.0422)	(0.0463)		(0.0454)		
		50	0.0314	0.0332	0.0321	0.0316	0.0332	0.0374
			(0.0298)	(0.0327)		(0.0321)		
		100	0.0229	0.0245	0.0232	0.0224	0.0245	0.0245
			(0.0211)	(0.0231)		(0.0227)		
root mean square error	1	25	0.1916	0.1342	0.1308	0.1590	0.1500	0.0909
		50	0.1483	0.1049	0.1048	0.1097	0.1145	0.0698
		100	0.1037	0.0854	0.0807	0.0791	0.0883	0.0581
	10	25	0.0671	0.0600	0.0616	0.0535	0.0624	0.0540
		50	0.0473	0.0480	0.0459	0.0417	0.0490	0.0496
		100	0.0325	0.0346	0.0326	0.0326	0.0346	0.0406
	20	25	0.0465	0.0447	0.0460	0.0413	0.0458	0.0490
		50	0.0317	0.0332	0.0322	0.0323	0.0332	0.0408
		100	0.0231	0.0245	0.0232	0.0226	0.0245	0.0262

Table 1: Sample means, sample standard deviations, theoretical standard deviations (in brackets) and sample root mean square error for $\alpha = 0.1, \lambda = 1$.

$(\alpha,\lambda)=(0.1,1.0)$	r	n	$\hat{\lambda}$ YW	CLS	IWCLS	CML	WHT	Bayes
sample bias	1	25	0.0529	-0.0164	0.0525	-0.1176	-0.0420	-0.0739
		50	0.0209	-0.0043	0.0191	-0.0551	-0.0377	-0.0259
		100	0.0091	-0.0056	0.0083	-0.0258	-0.0158	0.0086
	10	25	0.0090	-0.0011	0.0051	-0.0018	-0.0032	0.0308
		50	0.0134	0.0039	0.0115	0.0068	0.0027	0.0310
		100	0.0067	-0.0024	0.0056	0.0065	0.0000	0.0215
	20	25	0.0104	0.0030	0.0062	0.0047	0.0057	0.0300
		50	0.0050	0.0002	0.0031	0.0058	-0.0022	0.0212
		100	0.0027	0.0027	0.0018	0.0038	0.0018	0.0097
sample standard deviation (theoretical standard deviation)	1	25	0.2835	0.2532	0.2955	0.2307	0.3612	0.2027
			(0.3113)	(0.2981)		(0.2743)		
		50	0.2084	0.1811	0.2127	0.1729	0.2502	0.1565
			(0.2201)	(0.2108)		(0.2021)		
		100	0.1479	0.1245	0.1489	0.1327	0.1688	0.1192
			(0.1556)	(0.1491)		(0.1454)		
	10	25	0.0932	0.0964	0.0962	0.0860	0.1249	0.0825
			(0.0984)	(0.0943)		(0.0931)		
		50	0.0643	0.0693	0.0656	0.0632	0.0872	0.0648
			(0.0696)	(0.0667)		(0.0661)		
		100	0.0455	0.0490	0.0456	0.0458	0.0608	0.0510
			(0.0492)	(0.0471)		(0.0468)		
	20	25	0.0669	0.0608	0.0685	0.0632	0.0860	0.0648
			(0.0696)	(0.0667)		(0.0660)		
		50	0.0458	0.0480	0.0464	0.0141	0.0632	0.0510
			(0.0492)	(0.0471)		(0.0468)		
		100	0.0338	0.0346	0.0339	0.0332	0.0860	0.0346
			(0.0348)	(0.0333)		(0.0330)		
root mean square error	1	25	0.2884	0.2538	0.3001	0.2586	0.3636	0.2156
		50	0.2095	0.1811	0.2136	0.1812	0.2532	0.1584
		100	0.1482	0.1245	0.1492	0.1352	0.1697	0.1193
	10	25	0.0937	0.0964	0.0963	0.0862	0.1249	0.0881
		50	0.0656	0.0693	0.0666	0.0635	0.0872	0.0715
		100	0.0460	0.0490	0.0460	0.0467	0.0608	0.0553
	20	25	0.0677	0.0608	0.0688	0.0637	0.0860	0.0716
		50	0.0461	0.0480	0.0465	0.0460	0.0632	0.0548
		100	0.0339	0.0346	0.0340	0.0329	0.0447	0.0363

Table 2: Sample means, sample standard deviations, theoretical standard deviations (in brackets) and sample root mean square error for $\alpha = 0.1, \lambda = 1$.

$(\alpha, \lambda) = (0.3, 1.0)$	r	n	$\hat{\alpha}$ YW	CLS	IWCLS	CML	WHT	Bayes
sample bias	1	25	-0.0826	-0.0672	-0.0458	-0.0003	-0.0305	-0.0596
		50	-0.0406	-0.0331	-0.0331	-0.0135	-0.0140	-0.0615
		100	-0.0261	-0.0304	-0.0233	-0.0110	-0.0235	-0.0446
	10	25	-0.0172	-0.0132	-0.0055	-0.0086	-0.0206	-0.0211
		50	-0.0109	-0.0049	-0.0049	-0.0039	-0.0086	-0.0085
		100	-0.0034	-0.0018	-0.0004	-0.0021	-0.0035	-0.0046
	20	25	-0.0162	-0.0040	-0.0041	-0.0039	-0.0124	-0.0090
		50	-0.0076	-0.0001	-0.0017	-0.0016	-0.0043	-0.0039
		100	-0.0044	0.0016	-0.0014	-0.0008	-0.0007	-0.0019
sample standard deviation (theoretical standard deviation)	1	25	0.1860	0.1811	0.1642	0.1510	0.1957	0.1265
			(0.1596)	(0.2056)		(0.1839)		
		50	0.1324	0.1345	0.1294	0.1257	0.1418	0.1192
			(0.1129)	(0.1454)		(0.1313)		
		100	0.0983	0.1077	0.0991	0.0911	0.1114	0.1005
			(0.0798)	(0.1028)		(0.0926)		
	10	25	0.0617	0.0640	0.0640	0.0600	0.0663	0.0663
			(0.0505)	(0.0650)		(0.0584)		
		50	0.0458	0.0447	0.0459	0.0400	0.0469	0.0412
			0.0357	0.0460		0.0412		
		100	0.0315	0.0316	0.0317	0.0300	0.0316	0.0300
			0.0252	0.0325		0.0291		
	20	25	0.0455	0.0469	0.0471	0.0424	0.0500	0.0436
			(0.0357)	(0.0460)		(0.0412)		
		50	0.0326	0.0316	0.0332	0.0300	0.0316	0.0300
			(0.0252)	(0.0325)		(0.0291)		
		100	0.0229	0.0245	0.0231	0.0200	0.0245	0.0200
			(0.0178)	(0.0230)		(0.0205)		
root mean square error	1	25	0.2035	0.1342	0.1705	0.1509	0.1500	0.1398
		50	0.1385	0.1049	0.1385	0.1261	0.1145	0.1341
		100	0.1018	0.0854	0.1018	0.0919	0.0883	0.1099
	10	25	0.0641	0.0600	0.0643	0.0608	0.0624	0.1093
		50	0.0471	0.0480	0.0472	0.0403	0.0490	0.0422
		100	0.0317	0.0346	0.0317	0.0292	0.0346	0.0299
	20	25	0.0483	0.0447	0.0473	0.0424	0.0458	0.0447
		50	0.0335	0.0332	0.0333	0.0293	0.0332	0.0298
		100	0.0233	0.0245	0.0231	0.0202	0.0245	0.0206

Table 3: Sample means, sample standard deviations, theoretical standard deviations (in brackets) and sample root mean square error for $\alpha = 0.3, \lambda = 1$.

$(\alpha, \lambda) = (0.3, 1.0)$	r	n	$\hat{\lambda}$ YW	CLS	IWCLS	CML	WHT	Bayes
sample bias	1	25	0.1229	0.0767	0.1077	-0.0310	0.0036	0.0385
		50	0.0501	0.0483	0.0441	-0.0041	0.0024	0.0623
		100	0.0315	0.0349	0.0273	0.0064	0.0130	0.0538
	10	25	0.0194	0.0119	0.0034	0.0091	0.0199	0.0270
		50	0.0104	0.0057	0.0018	0.0053	0.0038	0.0115
		100	0.0068	0.0027	0.0026	0.0055	0.0049	0.0093
	20	25	0.0244	0.0053	0.0071	0.0055	0.0147	0.0124
		50	0.0089	-0.0014	0.0001	0.0047	0.0078	0.0081
		100	0.0041	-0.0020	-0.0002	0.0022	0.0030	0.0039
sample standard deviation (theoretical standard deviation)	1	25	0.3396	0.3489	0.3554	0.2704	0.3581	0.2608
			(0.3719)	(0.3381)		(0.3031)		
		50	0.2312	0.2520	0.2348	0.2064	0.2632	0.2138
			(0.2630)	(0.2390)		(0.2186)		
		100	0.1696	0.1811	0.1726	0.1546	0.2027	0.1703
			(0.1859)	(0.1690)		(0.1560)		
	10	25	0.1037	0.1049	0.1075	0.0985	0.1200	0.1058
			(0.1176)	(0.1069)		(0.0989)		
		50	0.0758	0.0742	0.0769	0.0686	0.0837	0.0707
			(0.0832)	(0.0756)		(0.0698)		
		100	0.0536	0.0548	0.0538	0.0469	0.0616	0.0480
			(0.0588)	(0.0535)		(0.0494)		
	20	25	0.0760	0.0800	0.0782	0.0700	0.0894	0.0721
			(0.0832)	(0.0756)		(0.0698)		
		50	0.0541	0.0539	0.0549	0.0469	0.0608	0.0469
			(0.0588)	(0.0535)		(0.0494)		
		100	0.0401	0.0387	0.0404	0.0332	0.0424	0.0332
			(0.0416)	(0.0378)		(0.0348)		
root mean square error	1	25	0.3612	0.2538	0.3714	0.2718	0.3636	0.2634
		50	0.2366	0.1811	0.2389	0.2061	0.2532	0.2224
		100	0.1725	0.1245	0.1748	0.1546	0.1697	0.1784
	10	25	0.1055	0.0964	0.1076	0.0988	0.1249	0.0697
		50	0.0765	0.0693	0.0770	0.0689	0.0872	0.0714
		100	0.0541	0.0490	0.0539	0.0473	0.0608	0.0486
	20	25	0.0798	0.0608	0.0786	0.0699	0.0860	0.0729
		50	0.0549	0.0480	0.0549	0.0469	0.0632	0.0480
		100	0.0403	0.0346	0.0404	0.0332	0.0447	0.0340

Table 4: Sample means, sample standard deviations, theoretical standard deviations (in brackets) and sample root mean square error for $\alpha = 0.3, \lambda = 1$.

$(\alpha, \lambda) = (0.9, 1.0)$	r	n	$\hat{\alpha}$ YW	CLS	IWCLS	CML	WHT	Bayes
sample bias	1	25	-0.2156	-0.1755	-0.1686	-0.0083	-0.0980	-0.0140
		50	-0.1040	-0.0868	-0.0818	-0.0074	-0.0547	-0.0099
		100	-0.0476	-0.0389	-0.0383	-0.0031	-0.0216	-0.0041
	10	25	-0.0510	-0.0101	-0.0143	-0.0011	-0.0030	-0.0015
		50	-0.0253	-0.0061	-0.0064	-0.0006	-0.0086	-0.0009
		100	-0.0125	-0.0034	-0.0036	-0.0006	-0.0100	-0.0007
	20	25	-0.0426	-0.0058	-0.0062	-0.0006	-0.0043	-0.0008
		50	-0.0216	-0.0038	-0.0034	-0.0006	-0.0112	-0.0007
		100	-0.0111	-0.0021	-0.0019	-0.0002	-0.0095	-0.0003
sample standard deviation (theoretical standard deviation)	1	25	0.1517	0.1766	0.1644	0.0374	0.2114	0.0400
			(0.0276)	(0.0892)		(0.0354)		
		50	0.0946	0.0943	0.0945	0.0265	0.1245	0.0283
			(0.0195)	(0.0631)		(0.0248)		
		100	0.0625	0.0566	0.0628	0.0173	0.0707	0.0173
			(0.0138)	(0.0446)		(0.0168)		
	10	25	0.0334	0.0316	0.0330	0.0100	0.0970	0.0100
			(0.0087)	(0.0282)		(0.0104)		
		50	0.0216	0.0200	0.0209	0.0100	0.0469	0.0100
			(0.0062)	(0.0199)		(0.0074)		
		100	0.0146	0.0141	0.0141	0.0100	0.0200	0.0100
			(0.0044)	(0.0141)		(0.0052)		
	20	25	0.0223	0.0224	0.0205	0.0100	0.0775	0.0100
			(0.0062)	(0.0199)		(0.0074)		
		50	0.0150	0.0141	0.0146	0.0100	0.0374	0.0100
			(0.0044)	(0.0141)		(0.0052)		
		100	0.0105	0.0100	0.0100	0.0100	0.0141	0.0100
			(0.0031)	(0.0100)		(0.0037)		
root mean square error	1	25	0.2636	0.2490	0.2356	0.0380	0.2330	0.0422
		50	0.1406	0.1285	0.1250	0.0280	0.1360	0.0296
		100	0.0786	0.0686	0.0736	0.0178	0.0735	0.0183
	10	25	0.0610	0.0332	0.0360	0.0107	0.0970	0.0108
		50	0.0332	0.0200	0.0219	0.0072	0.0469	0.0073
		100	0.0193	0.0141	0.0145	0.0051	0.0224	0.0051
	20	25	0.0481	0.0224	0.0214	0.0072	0.0775	0.0072
		50	0.0263	0.0141	0.0150	0.0051	0.0387	0.0052
		100	0.0153	0.0100	0.0102	0.0037	0.0173	0.0038

Table 5: Sample means, sample standard deviations, theoretical standard deviations (in brackets) and sample root mean square error for $\alpha = 0.9, \lambda = 1$.

			$\hat{\lambda}$					
$(\alpha, \lambda) = (0.9, 1.0)$	r	n	YW	CLS	IWCLS	CML	WHT	Bayes
sample bias	1	25	2.1159	1.7143	1.6724	0.0035	0.2893	0.0509
		50	1.0130	0.8573	0.7883	0.0108	0.2283	0.0314
		100	0.4728	0.3791	0.3766	0.0094	0.1092	0.0185
	10	25	0.5092	0.0983	0.1396	0.0101	0.3522	0.0147
		50	0.2561	0.0567	0.0653	0.0054	0.1828	0.0070
		100	0.1278	0.0299	0.0391	0.0099	0.0893	0.0107
	20	25	0.4200	0.0596	0.0566	0.0055	0.3513	0.0076
		50	0.2124	0.0388	0.0296	0.0104	0.1756	0.0112
		100	0.1114	0.0170	0.0194	0.0052	0.0890	0.0053
sample standard deviation (theoretical standard deviation)	1	25	1.6835	1.8866	1.7528	0.3282	0.6145	0.3527
			(0.9338)	(0.8944)		(0.3342)		
		50	0.9570	0.9531	0.9392	0.2280	0.4260	0.2337
			(0.6603)	(0.6325)		(0.2379)		
		100	0.6196	0.5913	0.6211	0.1673	0.2512	0.1688
			(0.4669)	(0.4472)		(0.1680)		
	10	25	0.3349	0.3124	0.3258	0.1049	0.2040	0.1058
			(0.2953)	(0.2828)		(0.1063)		
		50	0.2260	0.2054	0.2198	0.0735	0.1225	0.0742
			(0.2088)	(0.2000)		(0.0749)		
		100	0.1497	0.1400	0.1436	0.0510	0.0781	0.0510
			(0.1476)	(0.1414)		(0.0532)		
	20	25	0.2216	0.2128	0.2028	0.0728	0.1404	0.0735
			(0.2088)	(0.2000)		(0.0749)		
		50	0.1475	0.1493	0.1426	0.0510	0.0831	0.0510
			(0.1476)	(0.1414)		(0.0532)		
		100	0.1095	0.1025	0.1039	0.0374	0.0510	0.0387
			(0.1044)	(0.1000)		(0.0374)		
root mean square error	1	25	2.7039	2.5491	2.4226	0.3278	0.6792	0.3560
		50	1.3935	1.2820	1.2262	0.2280	0.4833	0.2356
		100	0.7794	0.7024	0.7264	0.1674	0.2739	0.1698
	10	25	0.6095	0.3276	0.3545	0.1052	0.4069	0.1069
		50	0.3416	0.2133	0.2293	0.0738	0.2200	0.0745
		100	0.1969	0.1432	0.1488	0.0519	0.1183	0.0523
	20	25	0.4748	0.2211	0.2106	0.0729	0.3783	0.0735
		50	0.2586	0.1543	0.1456	0.1007	0.1942	0.0525
		100	0.1562	0.1039	0.1057	0.0380	0.1025	0.0385

Table 6: Sample means, sample standard deviations, theoretical standard deviations (in brackets) and sample root mean square error for $\alpha = 0.9, \lambda = 1$.

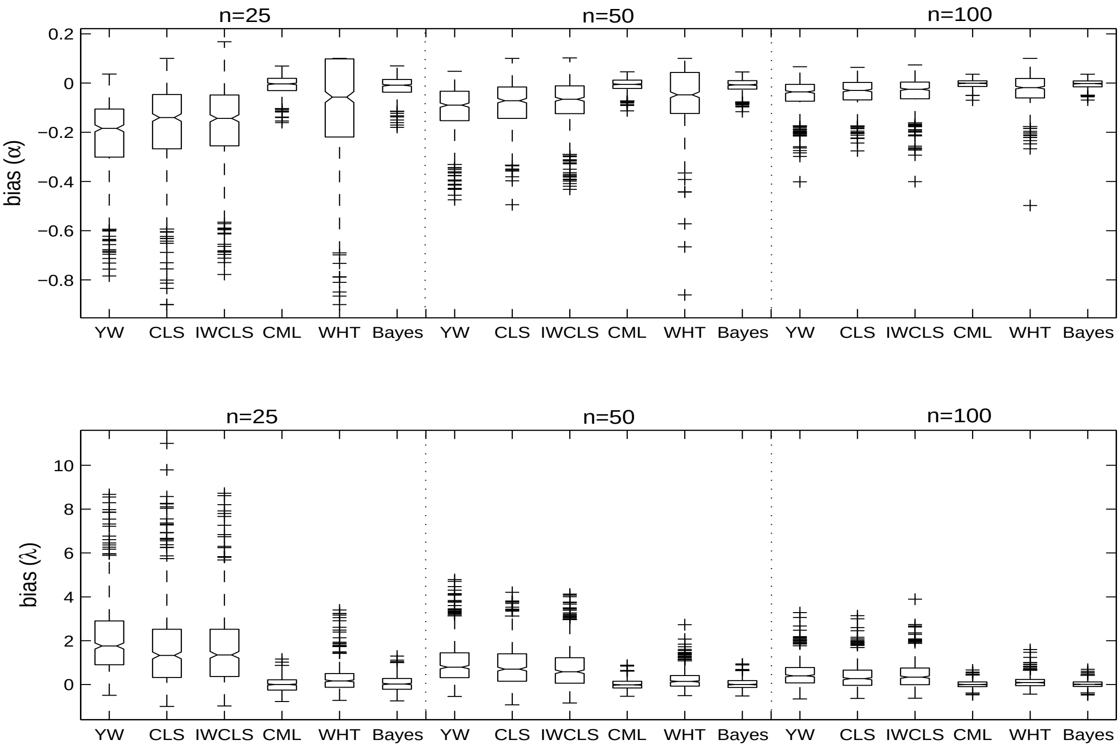

Figure 1: Boxplots of the biases for $\alpha = 0.9, \lambda = 1, r = 1, n = 25, 50, 100$.

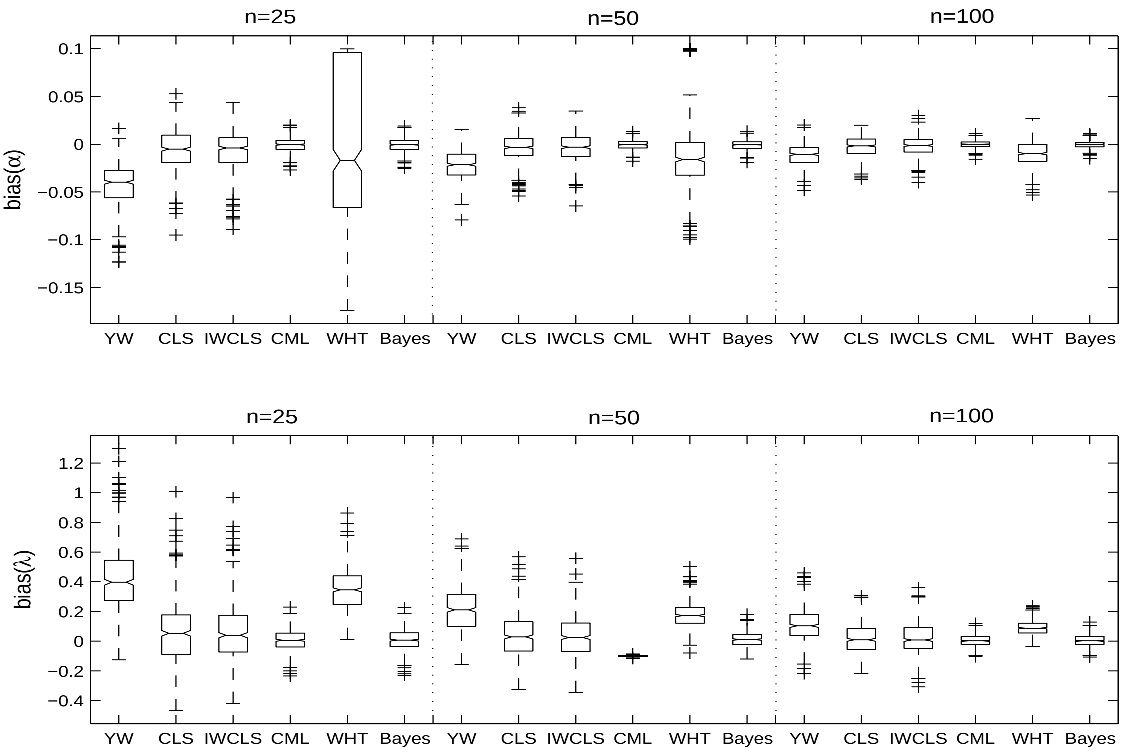

Figure 2: Boxplots of the biases for $\alpha = 0.9, \lambda = 1, r = 20, n = 25, 50, 100$.

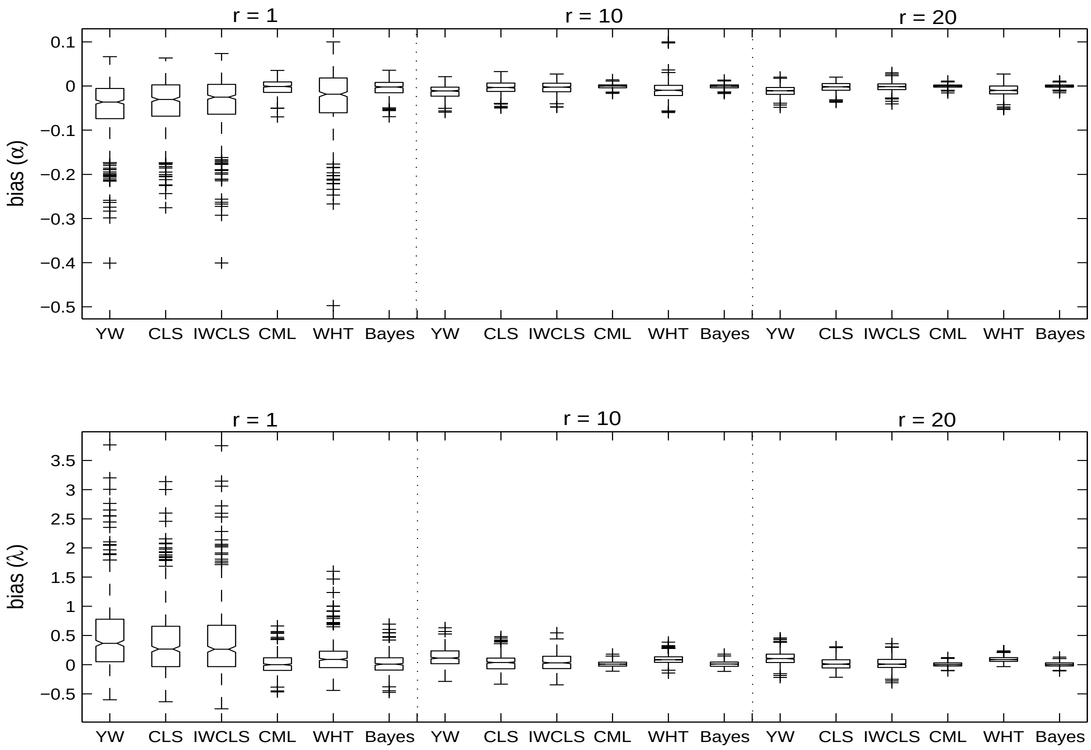

Figure 3: Boxplots of the biases for $\alpha = 0.9, \lambda = 1, r = 1, 10, 20, n = 100$.

ter (`http://www.ngdc.noaa/gov/`), in the section about Solar Sunspot Region. The data consists of the total number of sunspot groups per week, during two years (1990-1991), in a total of $n = 104$ observations, registered in two solar observatories: National Geophysical Data Center at Boulder (Colorado, USA) and Palehua Solar Observatory (Hawaii, USA). Figure 4 shows the two series with the corresponding sample autocorrelation functions and sample partial correlation functions.

Note that the number of sunspot groups in a week can be considered as the number of sunspot groups existing in the previous week that have not disappeared, with probability α, plus the new spot groups that appear in the current week.

For the Palehua series, the analysis of the correlogram and partial correlogram indicates a first-order model. The choice of $p = 1$ is corroborated by the AICC criterion for order selection in INAR models (I. Silva, 2005), which attains a minimum value of 403.32 for $p = 1$, when p is allowed to vary up to 10. On the other hand, for the Boulder series, the correlogram and partial correlogram indicate orders 1 and 3 as candidates for the order of the model. In this case, the AICC criterion gives a minimum value 383.2491 for $p = 1$ *versus* a value 404.8081 when $p = 3$. In addition, the variance of the residuals (when the parameters of the model are estimated by constrained Whittle criterion) is 17.6546 for the INAR(1) model and 17.9486 for the INAR(3) model. Therefore, we find that a first order model is suitable for both series. Further, considering that both observatories are observing the Sun we assume that the same INAR(1) model is appropriate for both series. Thus, although these series may present some degree of dependence, we consider that the series are a realization of a Poisson RINAR(1) process with $r = 2$ replicates. The parameters, (α, λ), are estimated by the methods proposed in the previous sections and presented in Table 7.

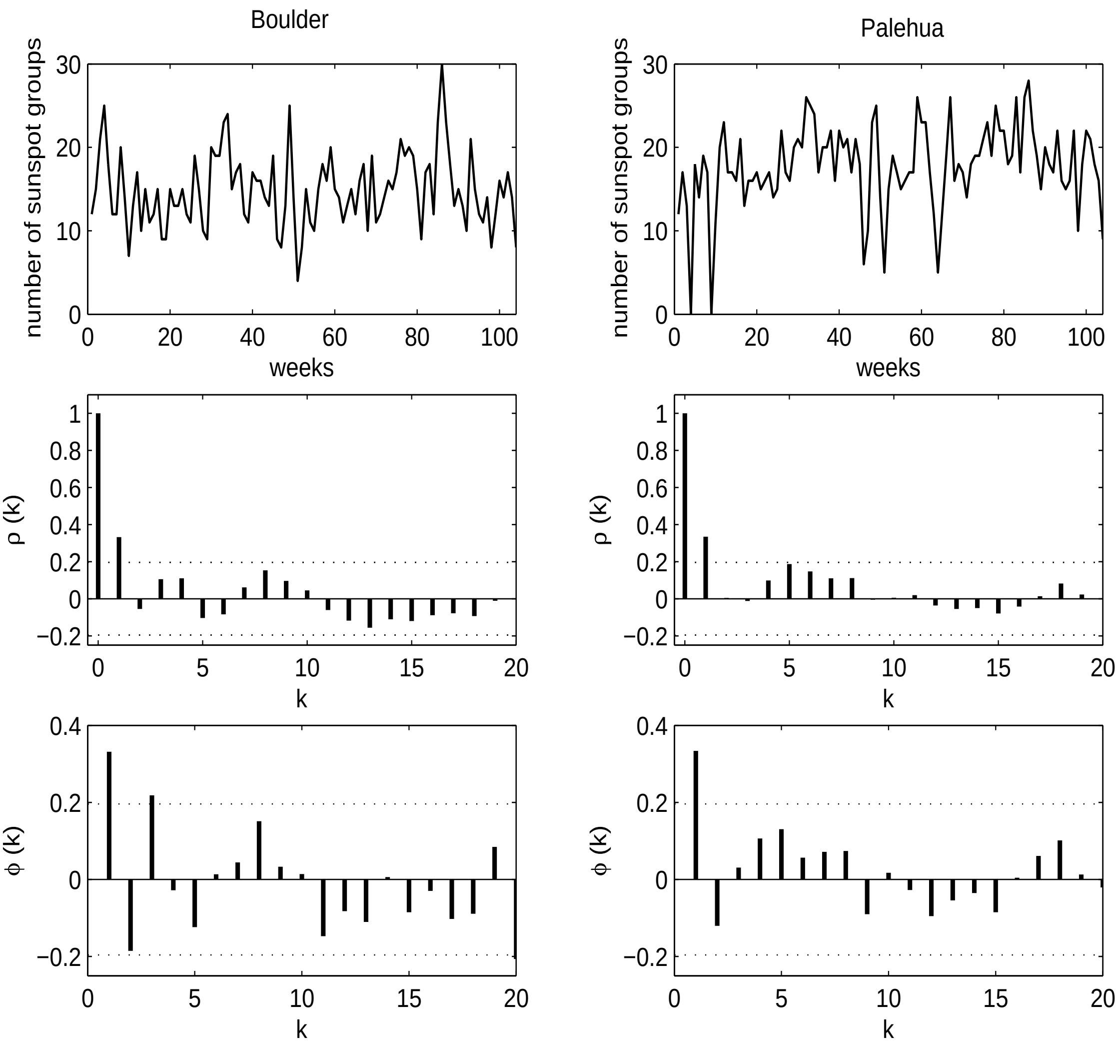

Figure 4: Number of sunspot groups per week, during two years (1990-1991) registered in two solar observatories and sample autocorrelation and sample partial autocorrelation functions.

In order to verify the goodness-of-fit of the RINAR(1) to the observations, we calculate the sample correlogram and sample partial correlogram for the residuals defined as

$$res_{M,t} = X_{k,t} - \hat{\alpha}_M X_{k,t-1} - \hat{\lambda}_M,$$

where $k = 1, 2;\ t = 2, \ldots, 104$ and M represents the estimation method. The usual randomness tests for the standardized residuals

$$Res_{M,t} = \frac{res_{M,t} - \overline{res}_M}{\hat{\sigma}_{res_M}},$$

where $\overline{res}_M$ is the sample mean of the residuals and $\hat{\sigma}_{res_M}$ is the sample standard deviation of the residuals, do not reject the hypothesis of uncorrelatedness. Thus, the RINAR(1) process is a reasonable model for the description of the data.

Method	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\mu}_X$
YW	0.390 (0.047)	9.936 (1.105)	16.284
CLS	0.399 (0.064)	9.795 (1.059)	16.303
IWCLS	0.399	9.795	16.303
CML	0.297 (0.122)	11.466 (2.025)	16.311
WHT	0.366	10.313	16.272
Bayes	0.289	11.627	16.344

Table 7: Parameter estimates of the RINAR(1) model for the total number of sunspot groups, per week ($r = 2, n = 104$), with the corresponding standard errors (in brackets).

Acknowledgements

The authors would like to thank the referees for their comments which helped to improve the paper. The first author would like to thank the PRODEP III for the financial support during the course of this work.

References

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

Azzalini, A. (1981). Replicated observations of low order autoregressive time series. *Journal of Time Series Analysis*, Vol. 2, pp. 63-70.

Azzalini, A. (1984). Estimation and hypothesis testing for collections of autoregressive time series. *Biometrika*, Vol. 71, pp. 85-90.

Brännäs, K. (1995). Explanatory variables in the AR(1) count data model. Umeå Economic Studies, 381, Umeå University.

Brillinger, D.R. (2001). *Time series: Data Analysis and Theory*, SIAM, Philadelphia.

Chandler, R.E. (1997). A spectral method for estimating parameters in rainfall models. *Bernoulli*, Vol. 3, pp. 301-322.

Congdon, P. (2003). *Applied Bayesian Modelling*, John Wiley and Sons.

Dégerine, S. (1987). Maximum likelihood estimation of autocovariance matrices from replicated short time series. *Journal of Time Series Analysis*, Vol. 8, pp. 135-146.

Diggle, P.J. and al-Wasel, I. (1997). Spectral Analysis of Replicated Biomedical Time Series. *Applied Statistics*, Vol. 46, pp. 31-71.

Dion, J-P., Gauthier, G. and Latour, A. (1995). Branching processes with immigration and integer-valued time series. *Serdica Mathematical Journal*, Vol. 21, pp. 123-136.

Du, Jin-Guan and Li, Yuan (1991). The integer-valued autoregressive (INAR(p)) model. *Journal of Time Series Analysis*, Vol. 12, pp. 129-142.

Dzhaparidze, K.O. and Yaglom, A.M. (1983). Spectrum parameter estimation in time series analysis. In *Developments in Statistics*, Vol. 4 (ed. P. R. Krishnaiah). Academic Press, New York. pp. 1-96.

Fox, R. and Taqqu, M. S. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, Vol. 14, pp. 517-532.

Franke, J. and Seligmann, T. (1993). Conditional maximum likelihood estimates for INAR(1) processes and their application to modelling epileptic seizure counts. In *Developments in Time Series Analysis* (ed. T. Subba Rao). Chapmann and Hall, London. pp. 310-330

Franke, J. and Subba Rao, T. (1995). Multivariate first order integer valued autoregressions. *Technical Report*, Math Dep., UMIST.

Freeland, R.K. and McCabe, B.P.M. (2004). Analysis of low count time series data by Poisson Autoregression. *Journal of the Time Series Analysis*, Vol. 25, pp. 701-722.

Gauthier, G. and Latour, A. (1994). Convergence forte des estimateurs des paramètrés d'un processus GENAR(p). *Annales des Sciences Mathématiques du Québec*, Vol. 18, pp. 49-71.

Gelfand, A.E. and Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association,* Vol. 85, pp. 398-409.

Gilks, W.R. and Best, N.G. (1987). Adaptive rejection Metropolis sampling within Gibbs sampling. *Applied Statistics*, Vol. 44, pp. 455-472.

Godambe, V.P. (1960). An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics*, Vol. 31, pp. 1208-1212.

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

Hoyt, D.V. and Schatten, K.H. (1998). Group sunspot numbers: a new solar activity reconstruction. *Solar Physics*, Vol. 181, pp. 491-512.

Johnson, N.L. and Kotz, S. (1969). *Distributions in Statistics: Discrete distributions*, John Wiley & Sons, New York.

Klimko, L. and Nelson, P. (1978). On conditional least square estimation for stochastic processes. *The Annals of Statistics*, Vol. 6, pp. 629-642.

Latour, A. (1997). The Multivariate GINAR(p) process. *Advances in Applied Probability*, Vol. 29, pp. 228-248.

Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process. *Journal of Time Series Analysis*, Vol. 19, pp. 439-455.

Park, Y. and Oh, C.W. (1997). Some asymptotic properties in INAR(1) processes with Poisson marginals. *Statistical Papers*, Vol. 38, pp. 287-302.

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

Rosenblatt, M. (1971). *Markov process - Structure and asymptotic behavior*. Springer-Verlag, New York.

Rosenblatt, M. (1983). Cumulants and cumulant spectra. In *Handbook of Statistics,* Vol. 3 (ed. D.R. Brillinger and P.R. Krishnaiah), Elsevier Science Publisher B.V., North Holland. pp. 369-382.

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

Silva, I. (2005). *Contributions to the analysis of discrete-valued time series.* PhD Thesis submitted to Universidade do Porto, Portugal.

Silva, N. (2005). *Análise Bayesiana de séries temporais de valores inteiros com observações repetidas.* PhD Thesis submitted to Universidade de Aveiro, Portugal.

Silva, M.E. and Oliveira, V.L. (2004). Difference equations for the higher order moments and cumulants of the INAR(1) model. *Journal of the Time Series Analysis*, Vol. 25, pp. 317-333.

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

Steutel, F. W. and Van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *The Annals of Probability*, Vol. 7, pp. 893-899.

Subba Rao, T. 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,* Vol. II (ed. P. Robinson and M. Rosenblatt). Springer-Verlag, Lecture Notes in Statistics No. 115, pp. 392-405.

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*, Vol. 4, pp. 363-384.

Wedderburn, R.W.M. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika*, Vol. 61, pp. 439-447.

Wei, C.Z. and Winnicki, J. (1989). Some asymptotic results for the branching process with immigration. *Stochastic Processes and their Applications*, Vol. 31, pp. 261-282.

Whittle, P. (1953). Estimation and information in stationary time series. *Arkiv fõr Matematik*, Vol. 2, pp. 423-434.

Whittle, P. (1954). Some recent contributions to the theory of stationary processes. In *A study in the analysis of stationary time series*, Appendix 2 (ed. H. Wold). 2nd ed. Almquist and Wiksell, Stockholm. pp. 196-228.

Winnicki, J. (1988). Estimation theory for the branching process with immigration. *Contemporary Mathematics*, Vol. 80, pp. 301-322.

Wooldridge, J.M. (1994). Estimation and inference for dependent processes. In *Handbook of Econometrics,* Vol 4 (ed. R.F. Engle and D.L. McFadden). Elsevier Science B.V., North-Holland, Amsterdam. pp. 2639-2738