A periodic bivariate integer-valued autoregressive model

Magda Monteiro, Manuel G. Scotto and Isabel Pereira

Abstract In this paper, a bivariate integer-valued autoregressive model with periodic structure is introduced and studied in some detail. The model can be viewed as a generalization of the one considered in Pedeli and Karlis (2011) [Pedeli, X., Karlis, D., 2011. A bivariate INAR(1) process with application. *Statist. Model.* 11, 325–349]. Emphasis is placed on models with periodic bivariate Poisson innovations. Basic probabilistic and statistical properties of the model are discussed. Moreover, parameter estimation and forecasting are also addressed.

1 Introduction

Periodically correlated processes play an important role in the analysis of a variety of data sets drawn from different areas such as economy ([10], [7], and [6]), hydrology ([23], [22], [21], and [24]), and signal processing ([8]) just to mention a few. Further examples can be viewed in [15] and [12] and the references therein. It is worth to mention that a large part of the literature on this topic is devoted to the continuous-valued Periodic AutoRegressive Moving Average (PARMA) models which are extensions of the commonly used ARMA models, having parameters which vary periodically in time. In contrast, however, the analysis of (univari-

Magda Monteiro

Escola Superior de Tecnologia e Gestão de Águeda and CIDMA, Universidade de Aveiro, Portugal, e-mail: msvm@ua.pt

Manuel G. Scotto

Departamento de Matemática and CIDMA, Universidade de Aveiro, Portugal, e-mail: mscotto@ua.pt

Isabel Pereira

Departamento de Matemática and CIDMA, Universidade de Aveiro, Portugal, e-mail: isabel.pereira@ua.pt

ate) periodically correlated time series of counts has not received much attention in the literature. The work in [15] introduced the periodic integer-valued autoregressive model of order one driven by a periodic sequence of independent Poisson-distributed random variables. The authors analyzed basic probabilistic and statistical properties of this models, namely the existence and uniqueness of a periodically stationary and causal process, its second-order structure, and issues related with parameter estimation. An application of the model proposed by [15] for the analysis of the number of hospital admissions per week caused by influenza can be found in [16]. In [11] was introduced a general class of periodic non-negative integer-valued moving average processes driven by a sequence of periodic integer-valued random variables with regularly varying tails. The authors analyzed some extremal properties related with this class of processes.

A related important problem which has not been addressed yet is the development of the bivariate integer-valued autoregressive model with periodic structure. This work aims at giving a contribution towards this direction. Many phenomena have in their essence a periodic structure and there are several potential applications for this class of models. For instance, these models can be applied in the environmental area, to model the monthly number of fires in neighbor counties; in epidemiological area, in the analysis of monthly (or daily) number of infections of different diseases related to each other, or in economy through the analysis of the monthly number of short term unemployed and long term unemployed or the monthly number of arrival flights and departure flights from an airport.

The literature on bivariate (and also multivariate) time series models for counts based on thinning operators is still in its infancy. An important contribution was made by Franke and Subba Rao in [5] who introduced the multivariate integer-valued autoregressive (MINAR) model of order one based on the binomial thinning operator, while a multivariate generalized INAR of order p was proposed by Latour in [14] in which matrices operate on vectors using the generalized thinning operator. More recently, Pedeli and Karlis introduced, in [17], the bivariate INAR (BINAR) model of order one with bivariate Poisson and bivariate negative binomial innovations. Pedeli and Karlis's model is defined as

$$\mathbf{X}_t = \mathbf{A} \circ \mathbf{X}_{t-1} + \mathbf{Z}_t \equiv \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \circ \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix}, t \in \mathbb{Z}, \quad (1)$$

where the (binomial) thinning operator “ $\circ$ ” is defined as $\alpha_j \circ X_j \stackrel{d}{=} \sum_{i=1}^{X_j} U_i(\alpha_j)$, being $U_i(\alpha_j)$, for $i = 1, \dots, X_j$, i.i.d. Bernoulli random variables with success probability $\alpha_j \in [0, 1]$, for $j = 1, 2$. Furthermore, the authors assumed that all thinning operations are performed independently of each other and of $(\mathbf{Z}_t) \equiv (\mathbf{Z}_t : t \in \mathbb{N})$ and that the thinning operations at each time t and $\mathbf{Z}_t$ are independent of $(\mathbf{X}_s)$ for $s < t$. Moreover, $(Z_{1,t}, Z_{2,t})$ are assumed to be independent $\mathbb{N}^2$ -valued random pairs. The authors illustrated the performance of the BINAR(1) model through an empirical application to the joint modeling of the number of daytime and nighttime

road accidents in the Netherlands for the year 2001. It is important to refer that in Pedeli and Karlis' model the autoregression matrix $\mathbf{A}$ is diagonal which means that there is no cross-autocorrelation in the counts; see also [18] for further details. A bivariate INAR model that accounts for cross-autocorrelation in the counts has been recently proposed by Boudreault and Charpentier in [2]. In order to also account for negative correlation between the time series, Karlis and Pedeli introduced in [13] a family of bivariate INAR(1) processes where negative cross-correlation is introduced through the innovations in terms of appropriate bivariate copulas. Extensions for bivariate INAR(1) models with positively correlated geometric marginals can be found in [20]. Bivariate INMA models based on the binomial thinning operator and non cross-autocorrelation in the count were proposed by Quoreshi, in [19] and by Brännäs and Quoreshi [3] who report an application to the number of transactions in intra-day data of stock.

In this work, the model proposed by Pedeli and Karlis, in [17], is generalized by assuming periodic time-varying parameters and periodic bivariate sequences of innovations, i.e., expression (1) takes the form

$$\mathbf{X}_t = \mathbf{A}_t \circ \mathbf{X}_{t-1} + \mathbf{Z}_t \equiv \begin{bmatrix} \phi_{1,t} & 0 \\ 0 & \phi_{2,t} \end{bmatrix} \circ \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix}, t \in \mathbb{Z} \quad (2)$$

with $\phi_{j,t} = \alpha_{j,i}$, for $t = i + kT$ ($i = 1, \dots, T$), $j = 1, 2$, and $k \in \mathbb{N}_0$. In this framework, the thinning operator is defined as

$$\phi_{j,t} \circ X_{j,t-1} \stackrel{d}{=} \sum_{i=1}^{X_{j,t-1}} U_{i,t}(\phi_{j,t}),$$

where $(U_{m,t}(\phi_{j,t}))$ is a periodic sequence of independent Bernoulli random variables with success probability $P(U_{m,t}(\phi_{j,t}) = 1) = \phi_{j,t}$. Note that by the properties of the binomial thinning operator

$$X_{j,t} = \phi_{j,t} \circ X_{j,t-1} + Z_{j,t}, j = 1, 2. \quad (3)$$

It is assumed that $(\mathbf{Z}_t)$ forms a periodic sequence of independent random vectors with mean $\boldsymbol{\delta}_t := [\delta_{1,t} \ \delta_{2,t}]'$ being $\delta_{j,t} = \lambda_{j,i}$ and covariance matrix $\boldsymbol{\Sigma}_t$ where $\sigma_{j,t}^2 = v_{j,i}\lambda_{j,i}$, with $v_{j,i} > 0$, $\sigma_{12,t} := \phi_i$, for $j = 1, 2$, and $t = i + kT$ ($i = 1, \dots, T$, $k \in \mathbb{N}_0$). Furthermore, for each t , $Z_{j,t}$ is assumed to be independent of $X_{j,t-1}$ and $\phi_{j,t} \circ X_{j,t-1}$. To avoid ambiguity T is taken as the smallest positive integer satisfying (2).

Throughout the rest of the work the model in (2) will be referred to as periodic bivariate integer-valued autoregressive model of order one (PBINAR(1), in short) with period $T \in \mathbb{N}$. Basic probabilistic and statistical properties of the PBINAR(1) model will be studied in some detail in the subsequent sections. Moreover, parameter estimation and forecasting will be also discussed.

The rest of the paper is organized as follows: in Section 2, we demonstrate the existence and uniqueness of a periodically stationary and causal PBINAR(1) process satisfying (2). Furthermore, expressions for the mean, variance, and autocovariance function are also derived. Parameter estimation is covered in Section 3. Forecasting is addressed in Section 4. In Section 5.2 we present a simulation study with a comparison between the different predictors referred in the previous section. Finally, concluding remarks are given in Section 6.

2 Basic properties of the PBINAR(1) model

The analysis of the existence and uniqueness of a periodically stationary and causal PBINAR(1) process follows easily by the arguments given by Pedeli and Karlis in [17], section 2, since $(\mathbf{X}_t)$ with $t = i + kT$ ($i = 1, \dots, T$) is a strictly stationary process. By iterating the equation in (3) and after rearranging some terms, it follows by Proposition 2.1 in [15] that for $i = 1, \dots, T$ the stationary distribution of $(X_{j,i+kT})$ is given by that of

$$V_{j,i} = \sum_{m=1}^{\infty} \sum_{a=0}^{T-1} (\beta_{i,i}^{(j)} \beta_{T,a}^{(j)} (\beta_{T,T}^{(j)})^{m-1}) \circ Z_{j,T(m+1)-a} + \sum_{m=0}^{i-1} \beta_{i,m}^{(j)} \circ Z_{j,i-m}, \quad j = 1, 2, \quad (4)$$

with

$$\beta_{t,m}^{(j)} := \begin{cases} \prod_{l=0}^{m-1} \phi_{j,t-l} & l > 0 \\ 1 & l = 0 \end{cases},$$

for $m \leq t$. Note that $(\beta_{t,m}^{(j)})$ is T -periodic and that $\beta_{t+kT,m}^{(j)} = \beta_{t,m}^{(j)}$ for $t > m$, satisfying that for $i = 1, 2, \dots, T$, and $k \in \mathbb{N}_0$, $\beta_{t,i+kT}^{(j)} = \beta_{t,i}^{(j)} (\beta_{T,T}^{(j)})^k$, $\beta_{i+T,i+a}^{(j)} = \beta_{i,i}^{(j)} \beta_{T,a}^{(j)}$ and $\beta_{T,T}^{(j)} = \prod_{m=1}^T \alpha_{j,m}$. The series on the right-hand side of (4) converges almost surely and also in L_2 .

From the representation in (4) the mean and autocovariance function of $(\mathbf{X}_t)$ can be obtained.

Lemma 1. *The mean value, variance and autocovariance structure of $(X_{j,t})$, for $j = 1, 2$ and $t = i + kT$ with $i = 1, 2, \dots, T$, ($T \in \mathbb{N}$), and $k \in \mathbb{N}_0$ are given by*

1. *Mean value:*

$$\mu_{j,i} \equiv E[X_{j,i}] = \frac{\sum_{m=0}^{i-1} \beta_{i,m}^{(j)} \lambda_{j,i-m} + \beta_{i,i}^{(j)} \sum_{m=0}^{T-i-1} \beta_{T,m}^{(j)} \lambda_{j,T-m}}{1 - \beta_{T,T}^{(j)}};$$

2. Variance:

$$\begin{aligned}
V[X_{j,t}] = & K_j(\beta) \sum_{m=0}^{i-1} \left\{ \beta_{T,T}^{(j)} \beta_{i,m}^{(j)} \lambda_{j,i-m} + \beta_{i,m}^{(j)} (1 - \beta_{i,m}^{(j)}) \lambda_{j,i-m} + \left(\beta_{i,m}^{(j)} \right)^2 v_{j,i-m} \lambda_{j,i-m} \right\} + \\
& + K_j(\beta) \sum_{m=0}^{T-i-1} \left\{ \beta_{T,T}^{(j)} \beta_{i,i}^{(j)} \beta_{T,m}^{(j)} \lambda_{j,T-m} + \right. \\
& \left. + \beta_{i,i}^{(j)} \beta_{T,m}^{(j)} \left(1 - \beta_{i,i}^{(j)} \beta_{T,m}^{(j)} \right) \lambda_{j,T-m} + \left(\beta_{i,i}^{(j)} \beta_{T,m}^{(j)} \right)^2 v_{j,T-m} \lambda_{j,T-m} \right\}
\end{aligned}$$

with $K_j(\beta) := 1/[1 - (\beta_{T,T}^{(j)})^2]$. The convention $\sum_{m=0}^{-1} = 0$ is adopted.

3. Autocovariance structure:

$$\begin{aligned}
\omega_i &:= \text{Cov}(X_{1,i}, X_{2,i}) = \text{Cov}(X_{1,i+kT}, X_{2,i+kT}) \\
&= \frac{1}{1 - \beta_{T,T}^{(1)} \beta_{T,T}^{(2)}} \beta_{i,i}^{(1)} \beta_{i,i}^{(2)} \sum_{m=0}^{T-i-1} \beta_{T,m}^{(1)} \beta_{T,m}^{(2)} \varphi_{T-m} + \\
&+ \frac{1}{1 - \beta_{T,T}^{(1)} \beta_{T,T}^{(2)}} \sum_{m=0}^{i-1} \beta_{i,m}^{(1)} \beta_{i,m}^{(2)} \varphi_{i-m}
\end{aligned}$$

and

$$\begin{aligned}
\text{Cov}(X_{1,t+h}, X_{2,t}) &= \frac{\beta_{t+h,h}^{(1)}}{1 - \beta_{T,T}^{(1)} \beta_{T,T}^{(2)}} \beta_{i,i}^{(1)} \beta_{i,i}^{(2)} \sum_{m=0}^{T-i-1} \beta_{T,m}^{(1)} \beta_{T,m}^{(2)} \varphi_{T-m} + \\
&+ \frac{\beta_{t+h,h}^{(1)}}{1 - \beta_{T,T}^{(1)} \beta_{T,T}^{(2)}} \sum_{m=0}^{i-1} \beta_{i,m}^{(1)} \beta_{i,m}^{(2)} \varphi_{i-m} \\
\text{Cov}(X_{1,t}, X_{2,t+h}) &= \frac{\beta_{t+h,h}^{(2)}}{1 - \beta_{T,T}^{(1)} \beta_{T,T}^{(2)}} \beta_{i,i}^{(1)} \beta_{i,i}^{(2)} \sum_{m=0}^{T-i-1} \beta_{T,m}^{(1)} \beta_{T,m}^{(2)} \varphi_{T-m} + \\
&+ \frac{\beta_{t+h,h}^{(2)}}{1 - \beta_{T,T}^{(1)} \beta_{T,T}^{(2)}} \sum_{m=0}^{i-1} \beta_{i,m}^{(1)} \beta_{i,m}^{(2)} \varphi_{i-m}
\end{aligned}$$

with $\varphi_i := \text{Cov}(Z_{1,i+T}, Z_{2,i+T})$.

Proof. The results follows by straightforward, although tedious, calculations. We skip the details.

Remark: The mean $\mu_{j,i}$ and ω_i can be calculated through the expressions

$$\mu_{j,i} = \beta_{i,i}^{(j)} \left(\mu_{j,T} + \frac{1}{\beta_{i,i}^{(j)}} \sum_{k=0}^{i-1} \beta_{i,k}^{(j)} \lambda_{j,i-k} \right), i = 1, \dots, T, j = 1, 2.$$

and

$$\omega_i = \sum_{k=0}^{i-1} \beta_{i,k}^{(1)} \beta_{i,k}^{(2)} \varphi_{i-k} + \beta_{i,i}^{(1)} \beta_{i,i}^{(2)} \omega_T, i = 1, \dots, T,$$

respectively.

Note that the probability generating function (pgf) of $\mathbf{X}_t$, for $t = i + kT$, takes the form

$$\begin{aligned} G_{\mathbf{X}_{i+kT}}(s_1, s_2) &\equiv G_{X_{1,i+kT}, X_{2,i+kT}}(s_1, s_2) = \\ &= G_{X_{1,0}}(1 - \beta_{i,i}^{(1)} (\beta_{T,T}^{(1)})^k + \beta_{i,i}^{(1)} (\beta_{T,T}^{(1)})^k s_1) G_{X_{2,0}}(1 - \beta_{i,i}^{(2)} (\beta_{T,T}^{(2)})^k + \beta_{i,i}^{(2)} (\beta_{T,T}^{(2)})^k s_2) \times \\ &\times \prod_{m=1}^k \prod_{a=0}^{T-1} G_{T-a; Z_1, Z_2} \left(1 - \beta_{i,i}^{(1)} \beta_{T,a}^{(1)} (\beta_{T,T}^{(1)})^{m-1} (1 - s_1), 1 - \beta_{i,i}^{(2)} \beta_{T,a}^{(2)} (\beta_{T,T}^{(2)})^{m-1} (1 - s_2) \right) \\ &\times \prod_{m=0}^{i-1} G_{i-m; Z_1, Z_2} \left(1 - \beta_{i,m}^{(1)} + \beta_{i,m}^{(1)} s_1, 1 - \beta_{i,m}^{(2)} + \beta_{i,m}^{(2)} s_2 \right), \end{aligned} \quad (5)$$

where $G_{i; Z_1, Z_2}$ represents the pgf of $\mathbf{Z}_{i+kT}$. The expression in (5) reduces to

$$\begin{aligned} G_{\mathbf{X}_{i+kT}}(s_1, s_2) &= \\ &\prod_{m=1}^{+\infty} \prod_{a=0}^{T-1} G_{T-a; Z_1, Z_2} \left(1 - \beta_{i,i}^{(1)} \beta_{T,a}^{(1)} (\beta_{T,T}^{(1)})^{m-1} (1 - s_1), 1 - \beta_{i,i}^{(2)} \beta_{T,a}^{(2)} (\beta_{T,T}^{(2)})^{m-1} (1 - s_2) \right) \\ &\times \prod_{m=0}^{i-1} G_{i-m; Z_1, Z_2} \left(1 - \beta_{i,m}^{(1)} + \beta_{i,m}^{(1)} s_1, 1 - \beta_{i,m}^{(2)} + \beta_{i,m}^{(2)} s_2 \right), \end{aligned} \quad (6)$$

as k tends to infinity.

Remark: For the particular case in which $\mathbf{Z}_{i+kT}$ ($i = 1, \dots, T$) follows the bivariate Poisson distribution (Johnson et al., 1997, p. 125)

$$\begin{aligned} P(Z_{1,i+kT} = z_1, Z_{2,i+kT} = z_2) &= \\ &= e^{-(\lambda_{1,i} + \lambda_{2,i} - \varphi_i)} \sum_{m=0}^{\min(z_1, z_2)} \frac{(\lambda_{1,i} - \varphi_i)^{z_1-m}}{(z_1-m)!} \frac{(\lambda_{2,i} - \varphi_i)^{z_2-m}}{(z_2-m)!} \frac{\varphi_i^m}{m!}, \end{aligned} \quad (7)$$

where $\lambda_{1,i}, \lambda_{2,i} > 0$ and $\varphi_i \in [0, \min(\lambda_{1,i}, \lambda_{2,i})]$, then

$$G_{\mathbf{X}_{i+kT}}(s_1, s_2) = \exp \{ \mu_{1,i}(s_1 - 1) + \mu_{2,i}(s_2 - 1) + \omega_i(s_1 - 1)(s_2 - 1) \}.$$

Parameters φ_i represent the covariance between the two time series within the i th period, for $i = 1, \dots, T$. The previous remark lead us to the following result.

Theorem 1. *The marginal distribution of $(\mathbf{X}_{i+kT})$ for $i = 1, \dots, T$ and $k \in \mathbf{N}_0$ is bivariate Poisson with parameters $(\mu_{1,i}, \mu_{2,i}, \omega_i)$ if and only if $(\mathbf{Z}_{i+kT})$ is bivariate Poisson with parameters $(\lambda_{1,i}, \lambda_{2,i}, \phi_i)$.*

3 Parameter estimation

Consider a finite time series $(\mathbf{X}_1, \dots, \mathbf{X}_{NT})$ from the PBINAR(1) model in (2). Let $\boldsymbol{\theta} := (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\phi})$ with $\boldsymbol{\alpha}_j = (\alpha_{j,1}, \dots, \alpha_{j,T})$, $\boldsymbol{\lambda}_j = (\lambda_{j,T}, \dots, \lambda_{j,1})$, for $j = 1, 2$ and $\boldsymbol{\phi} = (\phi_1, \dots, \phi_T)$ be the vector of unknown parameters. Without loss of generality it is assumed that $X_0 = x_0$. Note that the transition probabilities in this case take the form

$$\begin{aligned} p_i(\mathbf{y}|\mathbf{x}) &:= P(\mathbf{X}_{i+kT} = \mathbf{y} | \mathbf{X}_{i-1+kT} = \mathbf{x}) \\ &= \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} P(\alpha_{1,i} \circ X_{1,i-1+kT} = m_1, \alpha_{2,i} \circ X_{2,i-1+kT} = m_2 | \mathbf{X}_{i-1+kT} = \mathbf{x}) \times \\ &\quad \times P(Z_{1,i+kT} = y_1 - m_1, Z_{2,i+kT} = y_2 - m_2) \end{aligned}$$

with $\mathbf{y} := [y_1 \ y_2]'$, $\mathbf{x} := [x_1 \ x_2]'$, $M_1 := \min(x_1, y_1)$, $M_2 := \min(x_2, y_2)$, and $L := \min(y_1 - m_1, y_2 - m_2)$. The CML-estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by maximizing the conditional log-likelihood function

$$l(\boldsymbol{\theta}) := \ln(L(\boldsymbol{\theta})) = \sum_{n=0}^{N-1} \sum_{i=1}^T \ln(p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})).$$

Numerical maximization is straightforward with standard statistical packages.

Note that from Theorem 2.2. in [1], since $(\mathbf{X}_t)$ is a Markov chain, under standard assumptions, we can obtain asymptotically normality of the CML-estimators.

Theorem 2. *The CML-estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is asymptotically normal, i.e.,*

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} N(\mathbf{0}, I^{-1}(\boldsymbol{\theta})),$$

where $I(\boldsymbol{\theta})$ is the Fisher information matrix.

The choice of the joint distribution for the innovation bivariate process determines the properties of the underlying bivariate process. In the univariate case, the most frequently distributions assumed to the innovation process are the Poisson distribution and the negative binomial distribution. The first is appropriate for modeling equidispersed data and has the advantage of the stationary distribution has a closed form, also a Poisson distribution, while the second is adequate to model overdispersed count data. In the standard bivariate INAR model presented in [17], the bivariate distributions assumed for the innovation process were the bivariate Poisson

and the bivariate negative binomial. In the periodic case we will also give emphasis to these two distributions.

3.1 Innovations with periodic bivariate Poisson distribution

In the case of the PBINAR with periodic bivariate Poisson distribution for the innovation process, the transition probabilities is given by

$$\begin{aligned} p_i(\mathbf{y}|\mathbf{x}) &:= P(\mathbf{X}_{i+kT} = \mathbf{y} | \mathbf{X}_{i-1+kT} = \mathbf{x}) \\ &= e^{-(\lambda_{1,i} + \lambda_{2,i} - \phi_i)} \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} \sum_{l=0}^L \frac{\phi_i^l}{l!} \prod_{j=1}^2 C_{m_j}^{x_j} \alpha_{j,i}^{m_j} (1 - \alpha_{j,i})^{x_j - m_j} \frac{(\lambda_{j,i} - \phi_i)^{y_j - m_j - l}}{(y_j - m_j - l)!} \end{aligned}$$

In this case, from the partial derivatives of first order the following system is obtained

$$\left\{ \begin{array}{l} \sum_{n=0}^{N-1} \frac{x_{1,i-1+nT}}{1 - \alpha_{1,i}} \left(\frac{p_i(\mathbf{x}_{i+Tn} - (1,0) | \mathbf{x}_{i-1+Tn} - (1,0))}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} - 1 \right) = 0 \\ \sum_{n=0}^{N-1} \frac{x_{2,i-1+nT}}{1 - \alpha_{2,i}} \left(\frac{p_i(\mathbf{x}_{i+Tn} - (0,1) | \mathbf{x}_{i-1+Tn} - (0,1))}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} - 1 \right) = 0 \\ \sum_{n=0}^{N-1} \frac{p_i(\mathbf{x}_{i+Tn} - (1,0) | \mathbf{x}_{i-1+Tn})}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} = N \\ \sum_{n=0}^{N-1} \frac{p_i(\mathbf{x}_{i+Tn} - (0,1) | \mathbf{x}_{i-1+Tn})}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} = N \\ \sum_{n=0}^{N-1} \frac{p_i(\mathbf{x}_{i+Tn} - (1,1) | \mathbf{x}_{i-1+Tn})}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} = N \end{array} \right.$$

for $i = 1, \dots, T$. Analytical estimates for the above system cannot be found. Thus, to solve this system numerical procedures have to be employed. In order to find standard errors for the parameter estimates associated to the Theorem 2, the diagonal entries of the Hessian matrix are related to the expressions below and all other entries are calculated in a very straightforward manner:

$$\begin{aligned} \frac{\partial^2 l(\theta)}{\partial \alpha_{1,i}^2} &= \frac{1}{(1 - \alpha_{1,i})^2} \times \\ &\times \sum_{n=0}^{N-1} \left\{ -x_{1,i-1+nT} + x_{1,i-1+nT} (x_{1,i-1+nT} - 1) \frac{p_i(\mathbf{x}_{i+Tn} - (2,0) | \mathbf{x}_{i-1+Tn} - (2,0))}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} + \right. \\ &\left. + 2x_{1,i-1+nT} \frac{p_i(\mathbf{x}_{i+Tn} - (1,0) | \mathbf{x}_{i-1+Tn} - (1,0))}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} - \right. \end{aligned}$$

$$- \left(x_{1,i-1+nT} \frac{p_i(\mathbf{x}_{i+Tn} - (1,0) | \mathbf{x}_{i-1+Tn} - (1,0))}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} \right)^2 \Bigg\}; \quad (8)$$

$$\frac{\partial^2 l(\theta)}{\partial \lambda_{1,i}^2} = \sum_{n=0}^{N-1} \left\{ \frac{p_i(\mathbf{x}_{i+Tn} - (2,0) | \mathbf{x}_{i-1+Tn})}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} - \left(\frac{p_i(\mathbf{x}_{i+Tn} - (1,0) | \mathbf{x}_{i-1+Tn})}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} \right)^2 \right\}; \quad (9)$$

$$\frac{\partial^2 l(\theta)}{\partial \varphi_i^2} = \sum_{n=0}^{N-1} \left(\frac{p_i(\mathbf{x}_{i+Tn} - (1,1) | \mathbf{x}_{i-1+Tn})}{p_i(\mathbf{x}_{i+Tn} | \mathbf{x}_{i-1+Tn})} \right) \varphi_i' - \frac{\partial^2 l(\theta)}{\partial \varphi_i \partial \lambda_{1,i}} - \frac{\partial^2 l(\theta)}{\partial \varphi_i \partial \lambda_{2,i}},$$

for $i = 1, \dots, T$. The second derivatives in order to $\alpha_{2,i}^2$ and $\lambda_{2,i}^2$ are as in (8) and (9) with (2,0) and (1,0) replaced by (0,2) and (0,1). In the case of $\alpha_{2,i}^2$, $x_{1,i-1+nT}$ has to be replaced by $x_{2,i-1+nT}$ in (8).

3.2 Innovations with periodic bivariate negative binomial distribution

The case of the PBINAR with periodic bivariate negative binomial innovations is more flexible than the periodic Poisson BINAR(1). The transition probabilities is given by

$$p_i(\mathbf{y} | \mathbf{x}) = \left(\frac{\beta_i^{-1}}{\lambda_{1,i} + \lambda_{2,i} + \beta_i^{-1}} \right)^{\beta_i^{-1}} \sum_{m_1=0}^{M_1} \sum_{m_2=0}^{M_2} \frac{1}{\Gamma(\beta_i^{-1} + x_1 - m_1 + x_2 - m_2) \Gamma(\beta_i^{-1})} \times \\ \times \prod_{j=1}^2 \frac{C_{m_j}^{x_j}}{\Gamma(x_j - m_j + 1)} \alpha_{j,i}^{m_j} (1 - \alpha_{j,i})^{x_j - m_j} \left(\frac{\lambda_{j,i}}{\lambda_{1,i} + \lambda_{2,i} + \beta_i^{-1}} \right)^{y_j - m_j},$$

where $\lambda_{1,i}, \lambda_{2,i}, \beta_i > 0$, for all $i \in \{1, 2, \dots, T\}$, are the parameters associated with the periodic negative binomial bivariate distribution of the innovations. $\lambda_{1,i}$ and $\lambda_{2,i}$ are the mean of each component in season i and β_i is, for each season, the parameter associated with the overdispersion that this distribution allows. In fact, the variance, $\sigma_{j,i}^2$, is equal to $\lambda_{j,i}(1 + \beta_i \lambda_{j,i})$. The covariance between the two components, in each season i , is $\varphi_i = \lambda_{1,i} \lambda_{2,i} \beta_i$, $i = 1, \dots, T$, which only allows for positive correlation.

In this case the expressions of the partial derivatives of the log likelihood do not have a simple form as in the previous subsection.

4 Forecasting

In this section we consider the forecasting of future values $\mathbf{X}_{i+NT+h}$ of the periodic Poisson BINAR(1) process given past observations up through time $i + NT$, for

$i = 1, \dots, T$. Throughout the rest of the section it shall be assumed that $h = j + kT$, for $j \in \{1, \dots, T\}$. First note that by iterating equation (3) it follows that $X_{m,i+NT+h}$ can be expressed as

$$X_{m,i+NT+h} \stackrel{d}{=} \beta_{j+i,j}^{(m)} (\beta_{T,T}^{(m)})^k \circ X_{m,i+NT} + V_{m,j+i+kT},$$

where

$$V_{m,j+i+kT} = \sum_{r=0}^{j-1} \beta_{j+i,r}^{(m)} \circ Z_{m,j+i-r+NT} + \sum_{w=0}^{k-1} \sum_{r=0}^{T-1} \beta_{j+i+T(N+k),r+j+Tw}^{(m)} \circ Z_{i+T(N+k-w)-r}.$$

As in the univariate case,

$$\beta_{j+i,j}^{(m)} (\beta_{T,T}^{(m)})^k \circ X_{m,i+NT} | X_{m,i+NT} \sim Bi(X_{m,i+NT}, \beta_{j+i,j}^{(m)} (\beta_{T,T}^{(m)})^k), m = 1, 2.$$

Moreover, $V_{1,j+i+kT}$ and $V_{2,j+i+kT}$ are independent of $X_{1,i+NT}$ and $X_{2,i+NT}$, respectively with joint pgf

$$\begin{aligned} G_{V_{1,j+i+kT}, V_{2,j+i+kT}}(s_1, s_2) &= \prod_{r=0}^{j-1} G_{j+i-r; Z_1, Z_2} \left(1 - \beta_{j+i,r}^{(1)} (1 - s_1), 1 - \beta_{j+i,r}^{(2)} (1 - s_2) \right) \times \\ &\times \prod_{w=0}^{k-1} \prod_{r=0}^{T-1} G_{i-r+T; Z_1, Z_2} \left(1 - \beta_{j+i+T(N+k),r+j+Tw}^{(1)} (1 - s_1), 1 - \beta_{j+i+T(N+k),r+j+Tw}^{(2)} (1 - s_2) \right). \end{aligned}$$

Furthermore, it is assumed that $\mathbf{Z}_{i+kT}$ follows the bivariate Poisson distribution in (7) with parameters $(\delta_{1,t}, \delta_{2,t}, \psi_t)$ such that for $t = i + kT$, $\delta_{1,t} = \lambda_{1,i}$, $\delta_{2,t} = \lambda_{2,i}$ and $\psi_t = \varphi_i$. In this case, the joint pgf above takes the form

$$\begin{aligned} G_{V_{1,j+i+kT}, V_{2,j+i+kT}}(s_1, s_2) &= \exp \left\{ \left(\mu_{1,<i+j>} - (\beta_{T,T}^{(1)})^k \beta_{j+i,j}^{(1)} \mu_{1,i} \right) (s_1 - 1) \right\} \times \\ &\times \exp \left\{ \left(\mu_{2,<i+j>} - (\beta_{T,T}^{(2)})^k \beta_{j+i,j}^{(2)} \mu_{2,i} \right) (s_2 - 1) \right\} \times \\ &\times \exp \left\{ \left(\omega_{<i+j>} - (\beta_{T,T}^{(1)} \beta_{T,T}^{(2)})^k \beta_{j+i,j}^{(1)} \beta_{j+i,j}^{(2)} \omega_i \right) (s_1 - 1)(s_2 - 1) \right\} \end{aligned} \quad (10)$$

with

$$<i+j> := \begin{cases} i+j, & i+j \leq T \\ i+j-T, & i+j > T \end{cases}.$$

Note that the expression in (10) is the joint pgf of the bivariate Poisson distribution in (7) with parameters $(v_{1,i}, v_{2,i}, v_{3,i})$, being

$$\begin{aligned} v_{1,i} &:= \mu_{1,<i+j>} - (\beta_{T,T}^{(1)})^k \beta_{j+i,j}^{(1)} \mu_{1,i}; \\ v_{2,i} &:= \mu_{2,<i+j>} - (\beta_{T,T}^{(2)})^k \beta_{j+i,j}^{(2)} \mu_{2,i}; \\ v_{3,i} &:= \omega_{<i+j>} - (\beta_{T,T}^{(1)} \beta_{T,T}^{(2)})^k \beta_{j+i,j}^{(1)} \beta_{j+i,j}^{(2)} \omega_i. \end{aligned}$$

Thus, the distribution of $(X_{1,i+NT+h}, X_{2,i+NT+h})$ given $(X_{1,i+NT}, X_{2,i+NT})$ is the convolution of two binomial distributions, with parameters $(X_{1,i+NT}, \beta_{j+i,j}^{(1)}(\beta_{T,T}^{(1)})^k)$ and $(X_{2,i+NT}, \beta_{j+i,j}^{(2)}(\beta_{T,T}^{(2)})^k)$ respectively, with the bivariate distribution which has the joint pgf given in (10).

The discussion above leads to the following result.

Theorem 3. *For the bivariate Poisson periodic model $\mathbf{X}_{i+NT+h}|\mathbf{X}_{i+NT}$, $h = j + kT$, for $j \in \{1, 2, \dots, T\}$ and $k \in \mathbf{N}_0$, the following properties hold:*

a) *The pgf of $\mathbf{X}_{i+NT+h}|\mathbf{X}_{i+NT}$ is given by*

$$\begin{aligned} G_{\mathbf{X}_{i+NT+h}|\mathbf{X}_{i+NT}=(x_{1,i+NT}, x_{2,i+NT})}(s_1, s_2) &= \\ &= (1 - \beta_{j+i,j}^{(1)}(\beta_{T,T}^{(1)})^k(1-s_1))^{x_{1,i+NT}} (1 - \beta_{j+i,j}^{(2)}(\beta_{T,T}^{(2)})^k(1-s_2))^{x_{2,i+NT}} \\ &\times \exp\left\{\left(\mu_{1,<i+j>} - (\beta_{T,T}^{(1)})^k \beta_{j+i,j}^{(1)} \mu_{1,i}\right)(s_1 - 1)\right\} \times \\ &\times \exp\left\{\left(\mu_{2,<i+j>} - (\beta_{T,T}^{(2)})^k \beta_{j+i,j}^{(2)} \mu_{2,i}\right)(s_2 - 1)\right\} \times \\ &\times \exp\left\{\left(\omega_{<i+j>} - (\beta_{T,T}^{(1)}\beta_{T,T}^{(2)})^k \beta_{j+i,j}^{(1)}\beta_{j+i,j}^{(2)}\omega_i\right)(s_1 - 1)(s_2 - 1)\right\}; \end{aligned}$$

b) $E[X_{m,i+NT+j+kT}|\mathbf{X}_{m,i+NT}] = \left(\beta_{j+i,j}^{(m)}(\beta_{T,T}^{(m)})^k\right) X_{m,i+NT} + \mu_{m,<i+j>} - (\beta_{T,T}^{(m)})^k \beta_{j+i,j}^{(m)} \mu_{m,i}$,
 $m = 1, 2$;

c) $V[X_{m,i+NT+j+kT}|\mathbf{X}_{m,i+NT}] = \left(\beta_{j+i,j}^{(m)}(\beta_{T,T}^{(m)})^k\right) \left(1 - \beta_{j+i,j}^{(m)}(\beta_{T,T}^{(m)})^k\right) X_{m,i+NT} + \mu_{m,<i+j>} - (\beta_{T,T}^{(m)})^k \beta_{j+i,j}^{(m)} \mu_{m,i}$, $m = 1, 2$;

d) $Cov(X_{1,i+NT+j+kT}, X_{2,i+NT+j+kT}|\mathbf{X}_{1,i+NT}, \mathbf{X}_{2,i+NT}) = \omega_{<i+j>} - (\beta_{T,T}^{(1)}\beta_{T,T}^{(2)})^k \beta_{j+i,j}^{(1)}\beta_{j+i,j}^{(2)}\omega_i$.

e) *As $k \rightarrow +\infty$, $\mathbf{X}_{i+NT+h}|\mathbf{X}_{i+NT}$ has a bivariate Poisson distribution with parameters $(\mu_{1,<i+j>}, \mu_{2,<i+j>}, \omega_{<i+j>})$.*

To make a h -step ahead prediction we use the mode of the distribution of $\mathbf{X}_{i+NT+h}|\mathbf{X}_{i+NT}$ or the mean, median and mode of the marginal distributions of each component of $\mathbf{X}_{i+NT+h}|\mathbf{X}_{i+NT}$. The median and mode are considered estimates coherent with the model whereas the mean is considered an incoherent estimate since may not produce an integer value.

5 Simulation Study for a particular periodic Poisson BINAR(1) model

In this section a simulation study is conducted to illustrate the theoretical findings given in the previous section for the periodic Poisson BINAR(1) model and to assess the small, moderate and large sample behavior of the CML estimators. A comparison between the different predictors is also made in this section.

Throughout the analysis it shall be assumed that $T = 4$. The simulation study contemplates the following combination of α 's, λ 's and ϕ 's: $\alpha_1 = (0.5, 0.9, 0.3, 0.8)$, $\alpha_2 = (0.85, 0.4, 0.7, 0.2)$, $\lambda_1 = (4, 2, 8, 5)$, $\lambda_2 = (1.5, 5, 3, 10)$ and $\phi = (1, 1.5, 2.3, 3.8)$. We simulated times series of length $n = NT = 500, 1200, 2000$ with 200 independent replicates. Figure 1 shows the plots of a particular realization of the PBINAR model with dimension 1200 and Figure 2 presents the sample means, sample variances and sample covariance for each season.

Fig. 1 Plots of a particular realization of the PBINAR model

Fig. 2 Plots of sample means, sample variances and sample covariance of a particular realization of the PBINAR(1) model

5.1 Estimation

For the simulated model, and all replicates, the CML estimates of the parameters were calculated and the results are summarized in Table 1. Note that the estimates were calculated through numerical procedures that need initial values to start the optimization procedure. In this case we used the CML estimates for the parameters α 's and λ 's from each marginal PINAR model obtained by the use of the bisection method which does not require initial values. The initial values for the covariance parameters of the innovation bivariate process were found through the use of sample covariance of each season combined with the first equation of the third point of Lemma 1. The results obtained with these initial values were similar to those obtained by the use of the true values of the parameters as initial values in the optimization procedure.

Figures 3-5 display boxplots of the biases of the estimates for α_j and λ_j for $j = 1, 2$, and ϕ .

Table 1 Maximum likelihood estimates for θ . Standard errors in parentheses.

n	500	1200	2000
$\hat{\alpha}_{1,1}$	0.500 (0.056)	0.499 (0.027)	0.502 (0.019)
$\hat{\alpha}_{1,2}$	0.893 (0.070)	0.901 (0.008)	0.901 (0.005)
$\hat{\alpha}_{1,3}$	0.289 (0.069)	0.296 (0.044)	0.300 (0.034)
$\hat{\alpha}_{1,4}$	0.790 (0.083)	0.799 (0.019)	0.800 (0.013)
$\hat{\alpha}_{2,1}$	0.845 (0.065)	0.848 (0.009)	0.850 (0.007)
$\hat{\alpha}_{2,2}$	0.398 (0.066)	0.401 (0.035)	0.401 (0.027)
$\hat{\alpha}_{2,3}$	0.689 (0.063)	0.700 (0.021)	0.698 (0.015)
$\hat{\alpha}_{2,4}$	0.198 (0.085)	0.199 (0.053)	0.193 (0.041)
$\hat{\lambda}_{1,1}$	3.947 (0.635)	4.027 (0.366)	3.980 (0.305)
$\hat{\lambda}_{1,2}$	2.010 (0.293)	1.997 (0.111)	2.008 (0.090)
$\hat{\lambda}_{1,3}$	8.048 (0.998)	8.051 (0.531)	7.97 (0.440)
$\hat{\lambda}_{1,4}$	5.014 (0.747)	5.025 (0.243)	4.989 (0.174)
$\hat{\lambda}_{2,1}$	1.508 (0.199)	1.525 (0.134)	1.495 (0.099)
$\hat{\lambda}_{2,2}$	4.956 (0.758)	4.980 (0.429)	5.011 (0.310)
$\hat{\lambda}_{2,3}$	3.073 (0.416)	2.998 (0.215)	3.003 (0.149)
$\hat{\lambda}_{2,4}$	9.933 (1.116)	10.031 (0.537)	10.053 (0.416)
$\hat{\phi}_1$	0.889 (0.299)	0.976 (0.234)	0.952 (0.193)
$\hat{\phi}_2$	1.403 (0.422)	1.478 (0.236)	1.499 (0.182)
$\hat{\phi}_3$	2.216 (0.473)	2.241 (0.303)	2.234 (0.258)
$\hat{\phi}_4$	3.813 (0.957)	3.830 (0.385)	3.838 (0.300)

From Table 1, it can be observed that the standard errors of the estimators rapidly decrease to zero as n increases with special emphasis to the parameters related to the binomial thinning and the parameters associated with the average of the innovations. Furthermore, Figures 3-5 reveal that the estimates of α_1 and α_2 , componentwise, tend to be biased to the left and negatively skewed which implies that the CML estimation has a tendency of underestimating the α 's mainly in the case of small sample sizes. It also can be seen that CML estimation has a better performance componentwise, regardless the component of the bivariate model, for both α and λ , when thinning parameter is superior to 0.5. For ϕ 's this tendency is less obvious and is related with the magnitude of both thinning parameters. As expected, however, both the bias and skewness approach zero as the sample size increases. This is in agreement with the asymptotic properties of the CML estimators, namely unbiasedness and consistency.

5.2 Prediction

To compare and analyze the different predictors previously mentioned in section 4 the realizations of the PBINAR model were used to make h -step ahead predictions, from one to twenty. Consider $\hat{X}_{i,t+h}$, $\hat{m}_{i,t+h}$ and $\hat{m}o_{i,t+h}$ respectively the estimators of the mean, median and mode of the marginal conditional distribution $X_{i,n+h}|X_{i,n}$.

(a)

(b)

Fig. 3 Boxplots for (a) the biases of the CML estimators $\hat{\alpha}_1$ and (b) the biases of the CML estimators $\hat{\alpha}_2$. In (a), from left to right the first three boxplots display the biases of $\hat{\alpha}_{1,1}$ for $n = 500, 1200, 2000$. The subsequent three boxplots show the same information for $\hat{\alpha}_{1,2}$, the next three for $\hat{\alpha}_{1,3}$, and the last three boxplots for $\hat{\alpha}_{1,4}$. The boxplots in (b) show the same information for the four components of $\hat{\alpha}_2$.

In addition to these estimators was also used the mode of the joint conditional distribution $\mathbf{X}_{n+h}|\mathbf{X}_n, \hat{\mathbf{m}}_{\mathbf{O}_{t+h}}$.

In the different predictors the CML estimates were plugged-in in the prediction probability functions. To assess the performance of each estimator with the increase of dimension different measures were used. For the conditional mean is was considered the square root of the mean squared error (RMSE) while the mean absolute error (MAE) was used to evaluate the performance of the conditional median. For the conditional marginal mode and the mode of the joint distribution the loss func-

(a)

(b)

Fig. 4 Boxplots for (a) the biases of the CML estimators $\hat{\lambda}_1$ and (b) the biases of the CML estimators $\hat{\lambda}_2$. In (a), from left to right the first three boxplots display the biases of $\hat{\lambda}_{1,1}$ for $n = 500, 1200, 2000$. The subsequent three boxplots show the same information for $\hat{\lambda}_{1,2}$, the next three for $\hat{\lambda}_{1,3}$, and the last three boxplots for $\hat{\lambda}_{1,4}$. The boxplots in (b) show the same information for the four components of $\hat{\lambda}_2$.

tion everything or nothing (LFEN) was used to evaluate their performance. This last function is defined by

$$LFEN = \frac{1}{2mh} \sum_{k=1}^m \sum_{h=1}^{20} \sum_{i=1}^2 I\left(x_{i,t+h}^{(k)}\right),$$

where m represents the number of replicates and

Fig. 5 Boxplots for the biases of the CML estimators $\hat{\phi}$. From left to right the first three boxplots display the biases of $\hat{\phi}_1$ for $n = 500, 1200, 2000$. The subsequent three boxplots show the same information for $\hat{\phi}_2$, the next three for $\hat{\phi}_3$, and the last three boxplots for $\hat{\phi}_4$.

$$I\left(x_{i,t+h}^{(k)}\right)=\begin{cases} 1 & \text{if } |\hat{m}o_{i,t+h}^{(k)}-x_{i,t+h}^{(k)}|>1 \\ 0 & \text{if } |\hat{m}o_{i,t+h}^{(k)}-x_{i,t+h}^{(k)}|\leq 1 \end{cases}.$$

In order to compare the performance of the different predictors it was used the measure mean percentage absolute error (MPAE) given by

$$MPAE = \frac{1}{2Hm} \sum_{h=1}^H \sum_{k=1}^m \sum_{i=1}^2 |X_{i,t+h}^{*(k)} - X_{i,t+h}^{(k)}| / X_{i,t}^{(k)},$$

where m represents the number of replicates and H the number of predictions and $X_{i,t+h}^*$ represents one of the predictors used according to the methodology.

Table 2 presents a summary of the measures used to compare the predictors in the considered scenario. With the increase of n the measures RMSE, MAE and LFEN tend to decrease and the comparison of MPAE between predictors allows us to conclude that the mode of the joint conditional distribution of $\mathbf{X}_{n+h}|\mathbf{X}_n$ is the one that has a better performance for all dimensions that were used. Figure 6 presents the h -step-ahead predictive marginal distributions for each component, $P(x_{k,T+h}|x_{k,T})$, $k = 1, 2$ and $h = 1, 2, 3, 4$ for the same particular realization of PBINAR used in Figure 1. For each component, 1 and 2, the mode for the first season is respectively 9 and 11, for the second season is 10 and 9, for the third season is 11 and 9 and for the fourth season is 13 and 11. Figure 7 shows the h -step-ahead joint predictive distribution $P((x_{1,T+h}, x_{2,T+h})|(x_{1,T}, x_{2,T}))$, $h = 1, 2, 3, 4$. For the first season the mode is (10, 11), for the second season this pair is (11, 9), for the third season is (10, 9) and in the last season is (11, 9).

Table 2 RMSE, MAE, LFEN e MPAA of 20 predictions h -step ahead

	$\hat{X}_{t+h}$		$\hat{m}_{t+h}$		$\hat{m}o_{t+h}$		$\hat{m}o_{t+h}$	
	RMSE	MPAA	MAE	MPAA	LFEN	MPAA	LFEN	MPAA
$n = 500$	3.94	0.317	3.13	0.313	0.696	0.305	0.700	0.300
$n = 1200$	3.93	0.325	3.12	0.321	0.704	0.312	0.694	0.294
$n = 2000$	3.88	0.310	3.07	0.308	0.695	0.300	0.691	0.295

Fig. 6 Plots for the h -step-ahead predictive marginal distributions for each component, $P(x_{k,T+h}|x_{k,T})$, $k = 1, 2$ and $h = 1, 2, 3, 4$; (a) component 1 (b) component 2

6 Conclusions

In this article, a family of bivariate integer-valued autoregressive model of order one with periodic structure was proposed. This family is a generalization of the

Fig. 7 Plots for the h -step-ahead joint predictive distribution $P(\mathbf{x}_{T+h}|\mathbf{x}_T)$, $h = 1, 2, 3, 4$

BINAR model of Pedeli and Karlis (2011). Likelihood-based estimators for model parameters were derived and their asymptotic properties obtained and prediction was also addressed.

As referred throughout, an important limitation of Pedeli and Karlis' model is that the autoregression matrix is diagonal which means that it causes no cross-autocorrelation in the counts. This is also true for the PBINAR model. Therefore, extensions for PBINAR models accounting for cross-autocorrelation is also an impending problem. Moreover, similar to what happens with conventional PARMA models, PBINAR models can have an inordinately large number of parameters. Therefore, the development of procedures for dimensionality reduction remains an important topic for future work.

Acknowledgements This work was supported by Portuguese funds through the CIDMA - Center for Research and Development in Mathematics and Applications, and the Portuguese Foundation

for Science and Technology (FCT - Fundação para a Ciência e a Tecnologia), within project PEST-OE/MAT/UI4106/2014.

References

1. Billingsley, P. (1961) Statistical Inference for Markov Processes. University of Chicago Press, Chicago.
2. Boudreault, M., Charpentier, A., (2011) Multivariate integer-valued autoregressive models applied to earthquake counts. Submitted.
3. Brännäs, K., Quoreshi, A.M.M.S. (2010) Integer-valued moving average modelling of the number of transactions in stocks. *Appl. Finan. Econ.* 20, 1429–1440
4. Carvalho, A., Monteiro, A., Flannigan, M., Solman S., Miranda, A.I., and Borrego, C. (2011) Forest fires in a changing climate and their impacts on air quality. *Atmos. Environ.* 45, 5545–5553.
5. Franke, J., Subba Rao, T. (1993) Multivariate first-order integer-valued autoregressions. Technical report. Forschung Universität Kaiserslautern.
6. Franses, P.H. (1994) A multivariate approach to modeling univariate seasonal time series. *J. Econometrics* 63, 133–151.
7. Franses, P.H., Paap, R. (2004) Periodic Time Series. Oxford University Press. Oxford.
8. Gardner, W.A., Napolitano, A., Paura, L. (2006). Cyclostationary: half a century of research. *Signal Process.* 86, 639–697.
9. Gomes, J.F.P., (2006) Forest fires in Portugal: how they happen and why they happen. *Int. J. Environ. Studies* 63, 109–119.
10. Haldrup, N., Hyllerberg, S., Pons, G., Sansó, A., (2007) Common periodic correlation features and the interaction of stocks and flows in daily airport data. *J. Bus. Econom. Statist.* 25, 21–32.
11. Hall, A., Scotto, M.G., and Cruz, J., 2010. Extremes of integer-valued moving average sequences. *Test* 19, 359–374.
12. Hurd, H.L., Miamer, A. (2007) Periodically Correlated Random Sequences: Spectral Theory and Practice. John Wiley & Sons, New Jersey.
13. Karlis, D., Pedeli, X. (2013) Flexible bivariate INAR(1) processes using copulas. *Commun. Statist. Theor. Meth.* 42, 723–740.
14. Latour, A. (1997) The multivariate GINAR(p) process. *Adv. Appl. Probab.* 29, 228–248.
15. Monteiro, M., Scotto, M.G. and Pereira, I. (2010) Integer-valued autoregressive processes with periodic structure. *J. Statist. Plann. Inference* 140, 1529–1541.
16. Moriña, D., Puig, P., Ríos, J., Vilella, A., Trilla, A. (2011). A statistical model for hospital admissions caused by seasonal diseases. *Statist. Med.* 30, 3125–3136.
17. Pedeli, X., Karlis, D. (2011) A bivariate INAR(1) process with application. *Statist. Model.* 11, 325–349.
18. Pedeli, X., Karlis, D. (2013) On composite likelihood estimation of a multivariate INAR(1) model. *J. Time Ser. Anal.* 34, 206–220.
19. Quoreshi, A.M.M.S. (2006) Bivariate time series modelling of financial count data. *Commun. Statist. Theor. Meth.* 35, 1343–1358.
20. Ristic, M.M., Nastic, A.S., Jayakumar, K., Bakouch, H.S. (2012) A bivariate INAR(1) time series model with geometric marginals. *Appl. Math. Lett.* 25, 481–485.
21. Salas, J.D. (1993) Analysis and modeling of hydrologic time series. In: Maidment, D.R. (Ed.), *Handbook of Hydrology*, Chapter 19. McGraw-Hill, New York.
22. Tesfaye, Y.G., Meerschaert, M.M., Anderson, P.L. (2006) Identification of PARMA models and their application to the modeling of riverflows. *Water Resour. Res.* 42, W01419 (11pp.).
23. Ursu, E. and Turkman, K.F. (2012) Periodic autoregressive model identification using genetic algorithms. *J. Time Ser. Anal.* 33, 398–405.
24. Vecchia, A.V. (1985) Periodic autoregressive-moving average (PARMA) modelling with applications to water resources. *Water Res. Bull.* 21, 721–730.