Noname manuscript No.
(will be inserted by the editor)

On the theory of periodic multivariate INAR processes

Cláudia Santos · Isabel Pereira · Manuel G. Scotto

Received: date / Accepted: date

Abstract In this paper, a multivariate integer-valued autoregressive model of order one with periodic time-varying parameters, and driven by a periodic innovations sequence of independent random vectors is introduced and studied in detail. Emphasis is placed on models with periodic multivariate negative binomial innovations. Basic probabilistic and statistical properties of the novel model are discussed. Aiming to reduce computational burden arising from the use of the conditional maximum likelihood method, a composite likelihood-based approach is adopted. The performance of such method is compared with that of some traditional competitors, namely moment estimators and conditional maximum likelihood estimators. Forecasting is also addressed. Furthermore, an application to a real data set concerning the monthly number of fires in three counties in Portugal is presented.

Keywords Periodic autoregression · binomial thinning operator · parameter estimation

Cláudia Santos
CIDMA, University of Aveiro & Coimbra College of Agriculture, Polytechnic Institute of Coimbra, Portugal
E-mail: csps@ua.pt

Isabel Pereira
CIDMA and Department of Mathematics
University of Aveiro, Aveiro, Portugal
E-mail: isabel.pereira@ua.pt

Manuel G. Scotto
CEMAT and Department of Mathematics
IST University of Lisbon, Lisbon, Portugal
E-mail: manuel.scotto@tecnico.ulisboa.pt

1 Introduction

Recently, there has been a considerable interest in univariate thinning-based operators integer-valued time series models and nowadays, a voluminous literature has resulted from such interest in specialized books (Weiß 2018a; Davis et al. 2016; Turkman et al. 2014) and review papers (Scotto et al. 2015; Weiß 2008). However, the literature on multivariate integer-valued time series is much less developed. Extensions of univariate integer-valued autoregressive (`INAR`) processes to the multivariate case have been introduced by several authors. Franke and Subba Rao (1993) made an important contribution by introducing the multivariate `INAR` (`MINAR`) model of order one based upon a matrix of univariate independent binomial thinning operators. Karlis and Pedeli (2013) and Pedeli and Karlis (2011, 2013a,b) restrict their attention to the diagonal case, which means that the thinning operators causes no cross-correlation in the counts. Popović (2015) also considers the diagonal case, although the author assumes that the model is based on random coefficient thinning operators. Another important extension of Franke and Subba Rao's `MINAR` model was proposed by Latour (1997), in which the binomial thinning operators are replaced by generalized thinning operators (that is, thinning operators based on non-necessarily Bernoulli-distributed counting random variables). The `MINAR` model introduced by Boudreault and Charpentier (2011) and Pedeli and Karlis (2013c) also resembles Franke and Subba Rao's model and therefore accounts for cross-correlation in the counts. Moreover, models based on random coefficient thinning operators allowing for cross-correlation in the counts have been proposed by Popović (2016), Popović et al. (2016), Nastić et al. (2016) and Ristić et al. (2012). However, a major drawback of the aforementioned models is that they only allow for positive correlations between the time series. In order to also account for negative correlation Karlis and Pedeli (2013) introduced a bivariate `INAR`(1) model, in which negative cross-correlation is induced through the innovations, by defining its distribution in terms of appropriate bivariate copulas. Also to this end, Bulla et al. (2017) introduced the bivariate integer-valued autoregressive model (`B-SINAR`) based on the signed thinning operator (Kim and Park 2008). The advantage of the `B-SINAR` model is to fit integer-valued time series with positive and negative observations. A different approach was adopted by Scotto et al. (2014b) who introduced the so-called bivariate binomial thinning operator. Upon this new thinning operator the authors proposed a bivariate extension of the binomial autoregressive model of order one of McKenzie (1985), based on the bivariate binomial distribution of type II, which used to be referred to as $\texttt{BVB}_{\text{II}}\texttt{-AR(1)}$ model. It is important to stress the fact that the bivariate binomial thinning operator induces both positive and negative cross-correlation. An empirical application of such models, in statistical process control, can be found in He et al. (2016). Moreover, Möller et al. (2016) proposed an extension of the $\texttt{BVB}_{\text{II}}\texttt{-AR(1)}$ model for the analysis of the temporal characterization of integer-valued time series exhibiting piecewise-type patterns. More recently, Ristić and Popović (2019) have introduced a new bivariate binomial time series model with identical binomial marginal distribuitons.

Multivariate models for time series of counts based on moving average models have been also proposed in the literature. A remarkable contribution is due to Quoreshi (2006, 2008) who introduced the class of bivariate moving average time-series (`BINMA`) models and also the class of vector integer-valued moving average (`VINMA`) models. The `BINMA` class can be seen as an extension of the conventional integer-valued moving average introduced by Al-Osh and Alzaid (1988) and McKenzie (1988). The `VINMA` model is more general than the `BINMA` model and allows for both negative and positive correlation in the counting series. Extensions of Quoreshi's models have been proposed by Ristić et al. (2019), Sunecher et al. (2018) and Jowaheer et al. (2018) who consider `BINMA` models driven by COM-Poisson innovations and negative binomial innovations, respectively, under very general non-stationary moment assumptions. For further approaches and references on multivariate integer-valued time series models see Mamode Khan et al. (2019), Sunecher et al. (2019), Weiß (2018a), Karlis (2016), Scotto et al. (2015) and the references therein.

It is worth to mention here that all references given in the previous paragraphs deal with the case of non-periodically integer-valued time series. Such models, however, are useless to cope with periodically correlated processes. Although a large variety of integer-valued time series encountered in practice are periodically stationary (e.g. time series of tourism demand, fire activity and social science), the analysis of periodically correlated series of counts has not received much attention in the literature. Aiming this issue, Monteiro et al. (2010) introduced a class of univariate `INAR` models based on periodically varying thinning parameters. More recently, Monteiro et al. (2015) generalized the class of univariate `INAR` models of Monteiro et al. (2010) to the bivariate case. Two different distributional forms of the innovations were considered by the authors, namely: bivariate Poisson and bivariate negative binomial. In practice, however, the negative binomial distribution can better account for overdispersion (variance exceeds mean), a common feature in real data applications.

In this paper, we extend the results of Monteiro et al. (2015) to the multivariate case. Keeping an eye on practical applications, we will restrict our attention to the diagonal matrix case (Pedeli and Karlis 2011). Furthermore, the distribution of the innovation processes will be assumed to be periodic multivariate negative binomial. Motivation for considering the diagonal matrix case comes from the fact that in the application presented in section 5 to a multivariate data set of time series concerning the monthly number of fires in three counties in mainland Portugal, the assumption of independent counts is tenable.

The remainder of the paper is organized as follows. In section 2 the periodic multivariate integer-valued autoregressive model of order one is introduced. Parameter estimation is addressed in section 3. Furthermore, the performance of the estimation procedures is illustrated through a simulation study. Forecasting is covered in section 4. A real environmental data application based on fire activity is presented in section 5. Finally, some concluding remarks are summarized in section 6.

2 Periodic MINAR model of order one

Let $(\mathbf{X}_t)$ be a periodic m-variate integer-valued autoregressive process of first-order defined by the recursion

$$\mathbf{X}_t = \mathbf{M}_t \circ \mathbf{X}_{t-1} + \mathbf{Z}_t, \ \ t \in \mathbb{Z}, \tag{1}$$

where $\mathbf{X}_t, \mathbf{X}_{t-1}$ and $\mathbf{Z}_t$ are random ms-vectors with $\mathbf{X}_t = [\mathbf{X}_{1,t} \ \mathbf{X}_{2,t} \ \cdots \ \mathbf{X}_{m,t}]'$ for $t = v + ns$, $v = 1, \ldots, s$ and $n \in \mathbb{N}_0$, and $\mathbf{X}_{j,t} = [X_{j,1+ns} \ \cdots \ X_{j,s+ns}]'$, $j = 1, \ldots, m$. The ms-dimensional vector $\mathbf{Z}_t = [\mathbf{Z}_{1,t} \ \mathbf{Z}_{2,t} \ \cdots \ \mathbf{Z}_{m,t}]'$ constitutes a periodic sequence of independent random vectors with

$$\mathbf{Z}_{j,t} = [Z_{j,1+ns} \ Z_{j,2+ns} \ \cdots \ Z_{j,s+ns}]'. \tag{2}$$

The matrix $\mathbf{M}_t$ in (1) is a $(ms \times ms)$-diagonal matrix defined as

$$\mathbf{M}_t = \begin{bmatrix} \boldsymbol{\phi}_{1,t} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\phi}_{2,t} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\phi}_{m,t} \end{bmatrix},$$

which includes the periodic autoregressive coefficients for season v $(v = 1, \ldots, s)$. Moreover, the $\mathbf{0}$'s are $(s \times s)$-null matrices and the $\boldsymbol{\phi}_{j,t}$'s $(j = 1, \ldots, m)$ are $(s \times s)$-diagonal matrices. Note that for a fixed $t = v + ns$ $(v = 1, \ldots, s)$ and $n \in \mathbb{N}_0$, diagonal elements of $\boldsymbol{\phi}_{j,t} = \alpha_{j,v} \in (0, 1)$, i.e.,

$$\boldsymbol{\phi}_{j,t} = diag(\alpha_{j,1}, \alpha_{j,2}, \ldots, \alpha_{j,s}). \tag{3}$$

The model in (1) will be referred to as `P`eriodic `M`ultivariate `IN`teger-valued `A`uto`R`egressive model of order one (`PMINAR`(1) in short) with period $s \in \mathbb{N}$. The `PMINAR`(1) model admits the following matricial representation

$$\begin{bmatrix} \mathbf{X}_{1,t} \\ \mathbf{X}_{2,t} \\ \vdots \\ \mathbf{X}_{m,t} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_{1,t} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\phi}_{2,t} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\phi}_{m,t} \end{bmatrix} \circ \begin{bmatrix} \mathbf{X}_{1,t-1} \\ \mathbf{X}_{2,t-1} \\ \vdots \\ \mathbf{X}_{m,t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{Z}_{1,t} \\ \mathbf{Z}_{2,t} \\ \vdots \\ \mathbf{Z}_{m,t} \end{bmatrix}. \tag{4}$$

For each t, $\mathbf{Z}_{j,t}$ is assumed to be independent of $\mathbf{X}_{j,t-1}$ and $\boldsymbol{\phi}_{j,t} \circ \mathbf{X}_{j,t-1}$. Note that the j-th component in (4) is

$$\mathbf{X}_{j,t} = \boldsymbol{\phi}_{j,t} \circ \mathbf{X}_{j,t-1} + \mathbf{Z}_{j,t} \tag{5}$$

and each element of the vector $\boldsymbol{\phi}_{j,t} \circ \mathbf{X}_{j,t-1}$, for $t = v + ns$ fixed, is given by

$$\alpha_{j,v} \circ X_{j,v+ns-1} \stackrel{d}{=} \sum_{r=1}^{X_{j,v+ns-1}} U_{r,v}\left(\alpha_{j,v}\right),$$

where $\{U_{r,v}(\alpha_{j,v})\}_{r \in \mathbb{N}}$ is a periodic sequence of i.i.d. Bernoulli-distributed random variables with probability of success $P\left(U_{r,v}\left(\alpha_{j,v}\right) = 1\right) = \alpha_{j,v}$. Since the autocorrelation matrix $\mathbf{M}_t$ is diagonal, the only source of dependence between the series $(\mathbf{X}_{1,t}, \ldots, \mathbf{X}_{m,t})$ in (4) is provided through $\mathbf{Z}_t$. Therefore, the innovations will play a central role in the specification of the PMINAR(1) process.

Due to the fact that $t = v + ns$, it follows that $\mathbf{X}_{j,t-s} \stackrel{d}{=} \mathbf{X}_{j,v+(n-1)s}$, for $v = 1, \ldots, s$. Now, take element

$$X_{j,1+ns} = \alpha_{j,1} \circ X_{j,1+ns-1} + Z_{j,1+ns}$$

and replace it in $X_{j,2+ns}$, then

$$\begin{aligned} X_{j,2+ns} &= \alpha_{j,2} \circ X_{j,2+ns-1} + Z_{j,2+ns} = \alpha_{j,2} \circ X_{j,1+ns} + Z_{j,2+ns} = \\ &= \alpha_{j,2} \circ (\alpha_{j,1} \circ X_{j,ns} + Z_{j,1+ns}) + Z_{j,2+ns} = \\ &= \alpha_{j,2}\alpha_{j,1} \circ X_{j,ns} + \alpha_{j,2} \circ Z_{j,1+ns} + Z_{j,2+ns} \end{aligned}$$

and recursively

$$\begin{aligned} X_{j,3+ns} &= \alpha_{j,3} \circ X_{j,3+ns-1} + Z_{j,3+ns} = \alpha_{j,3} \circ X_{j,2+ns} + Z_{j,3+ns} = \\ &= \alpha_{j,3}\alpha_{j,2}\alpha_{j,1} \circ X_{j,ns} + \alpha_{j,3}\alpha_{j,2} \circ Z_{j,1+ns} + \alpha_{j,3} \circ Z_{j,2+ns} + Z_{j,3+ns} \end{aligned}$$

and so on allowing the j-th component $\mathbf{X}_{j,t}$ in (5) to be expressed in matricial notation as

$$\mathbf{X}_{j,t} = \mathbf{A}_j \circ \mathbf{X}_{j,t-s} + \mathbf{B}_j \circ \mathbf{Z}_{j,t}, \tag{6}$$

where the $(s \times s)$-matrices $\mathbf{A}_j$ and $\mathbf{B}_j$ $(j = 1, \ldots, m)$ are given by

$$\mathbf{A}_j = \begin{bmatrix} 0 \cdots & 0 & \alpha_{j,1} \\ 0 \cdots & 0 & \alpha_{j,2}\alpha_{j,1} \\ 0 \cdots & 0 & \alpha_{j,3}\alpha_{j,2}\alpha_{j,1} \\ \vdots \; \vdots & \ddots & \vdots \\ 0 \cdots & 0 & \prod_{k=0}^{s-1} \alpha_{j,s-k} \end{bmatrix} \tag{7}$$

and

$$\mathbf{B}_j = \begin{bmatrix} 1 & 0 & 0 & \ldots 0 \\ \alpha_{j,2} & 1 & 0 & \ldots 0 \\ \alpha_{j,3}\alpha_{j,2} & \alpha_{j,3} & 1 & \ldots 0 \\ \vdots & \vdots & \vdots & \ddots \vdots \\ \prod_{k=0}^{s-2} \alpha_{j,s-k} & \prod_{k=0}^{s-3} \alpha_{j,s-k} & \prod_{k=0}^{s-4} \alpha_{j,s-k} & \ldots 1 \end{bmatrix}, \tag{8}$$

respectively, with coefficients $\alpha_{j,v} \in (0, 1)$, $j = 1, \ldots, m$ and $v = 1, \ldots, s$. All columns of matrices $\mathbf{A}_j$, except the last one, are null. The matrices $\mathbf{B}_j$ are lower triangular matrices.

Taking all m components, the PMINAR(1) model in (1) can be rewritten in the form

$$\mathbf{X}_t = \widetilde{\mathbf{A}} \circ \mathbf{X}_{t-s} + \widetilde{\mathbf{B}} \circ \mathbf{Z}_t, \tag{9}$$

with matricial representation

$$\begin{bmatrix} \mathbf{X}_{1,t} \\ \mathbf{X}_{2,t} \\ \vdots \\ \mathbf{X}_{m,t} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_m \end{bmatrix} \circ \begin{bmatrix} \mathbf{X}_{1,t-s} \\ \mathbf{X}_{2,t-s} \\ \vdots \\ \mathbf{X}_{m,t-s} \end{bmatrix} + \begin{bmatrix} \mathbf{B}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{B}_m \end{bmatrix} \circ \begin{bmatrix} \mathbf{Z}_{1,t} \\ \mathbf{Z}_{2,t} \\ \vdots \\ \mathbf{Z}_{m,t} \end{bmatrix}.$$

The $(ms \times ms)$-matrices $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ in (9) are block-diagonal matrices, that is

$$\widetilde{\mathbf{A}} = diag(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m) \tag{10}$$

and

$$\widetilde{\mathbf{B}} = diag(\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_m) \tag{11}$$

with matrices $\mathbf{A}_j$ and $\mathbf{B}_j$ $(j = 1, \ldots, m)$ as in (7) and in (8), respectively. Generally, matrix $\widetilde{\mathbf{A}}$ has entries a^j_{ik} satisfying $0 \leq a^j_{ik} < 1$ and matrix $\widetilde{\mathbf{B}}$ has entries b^j_{ik} satisfying $0 \leq b^j_{ik} \leq 1$ with $i, k = 1, \ldots, ms$ and $j = 1, \ldots, m$. Furthermore, it will be assumed that the innovations $\mathbf{Z}_t$ have finite first- and second-order moments being

$$E[\mathbf{Z}_t] \equiv E \begin{bmatrix} \mathbf{Z}_{1,t} \\ \mathbf{Z}_{2,t} \\ \vdots \\ \mathbf{Z}_{m,t} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\delta}_{1,t} \\ \boldsymbol{\delta}_{2,t} \\ \vdots \\ \boldsymbol{\delta}_{m,t} \end{bmatrix} =: \boldsymbol{\delta}_t. \tag{12}$$

The ms-mean vector $\boldsymbol{\delta}_t$ with $t = v + ns; v = 1, \ldots, s$ and $n \in \mathbb{N}_0$ has m $(s \times 1)$-vectors, i.e.,

$$E[\mathbf{Z}_{j,t}] = \boldsymbol{\delta}_{j,t} = \begin{bmatrix} \lambda_{j,1} \\ \lambda_{j,2} \\ \vdots \\ \lambda_{j,s} \end{bmatrix}, \tag{13}$$

for $j = 1, \ldots, m$. For a fixed v, each element in (13) is

$$E[Z_{j,v+ns}] = \lambda_{j,v}. \tag{14}$$

Turning to the variance-covariance matrix of $\mathbf{Z}_t$, it follows that

$$\sum\nolimits_{\mathbf{Z}_t} := \begin{bmatrix} Var[\mathbf{Z}_{1,t}] & Cov(\mathbf{Z}_{1,t}, \mathbf{Z}_{2,t}) & \cdots & Cov(\mathbf{Z}_{1,t}, \mathbf{Z}_{m,t}) \\ Cov(\mathbf{Z}_{2,t}, \mathbf{Z}_{1,t}) & Var[\mathbf{Z}_{2,t}] & \cdots & Cov(\mathbf{Z}_{2,t}, \mathbf{Z}_{m,t}) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\mathbf{Z}_{m,t}, \mathbf{Z}_{1,t}) & Cov(\mathbf{Z}_{m,t}, \mathbf{Z}_{2,t}) & \cdots & Var[\mathbf{Z}_{m,t}] \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{\Psi}_{11,t} & \boldsymbol{\Psi}_{12,t} & \cdots & \boldsymbol{\Psi}_{1m,t} \\ & \boldsymbol{\Psi}_{22,t} & \cdots & \boldsymbol{\Psi}_{2m,t} \\ & & \ddots & \vdots \\ & & & \boldsymbol{\Psi}_{mm,t} \end{bmatrix} =: \boldsymbol{\Psi}_t, \tag{15}$$

where $\boldsymbol{\Psi}_{jk,t}$ $(j,k = 1,\ldots,m; t = v + ns; v = 1,\ldots,s; n \in \mathbb{N}_0)$ are $(s \times s)$-diagonal matrices of the form

$$\boldsymbol{\Psi}_{jk,t} = \begin{bmatrix} \sigma_{jk,1} & 0 & \ldots & 0 \\ 0 & \sigma_{jk,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_{jk,s} \end{bmatrix}. \tag{16}$$

Note that for a fixed v, each element of the diagonal in matrix (16) takes the form

$$\sigma_{jk,v} = Cov(Z_{j,v+ns}, Z_{k,v+ns}). \tag{17}$$

For notational simplicity, we use $\sigma^2_{j,t}$ instead of $\sigma_{jj,t}$ for $j = k$ $(j = 1,\ldots,m)$ and for $t = v + ns; v = 1,\ldots,s$,

$$\boldsymbol{\Psi}_{jj,t} \equiv Var[\mathbf{Z}_{j,t}] = \begin{bmatrix} \sigma^2_{j,1} & 0 & \ldots & 0 \\ 0 & \sigma^2_{j,2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma^2_{j,s} \end{bmatrix}. \tag{18}$$

For a fixed v, each element of the diagonal in matrix (18) is given by

$$\sigma^2_{j,v} = Var[Z_{j,v+ns}].$$

Furthermore, the $(ms \times ms)$-matrix $\boldsymbol{\Psi}_t$ in (15) has m on-diagonal matrices equal to $\boldsymbol{\Psi}_{jj,t}$ in (18) and $(m-1)m$ off-diagonal matrices equal to $\boldsymbol{\Psi}_{jk,t}$ in (16) with $j \neq k$; $j,k = 1,\ldots,m$.

2.1 Strictly periodically stationary distribution

The existence of a periodically stationary solution to (9) depends on the largest eigenvalue of the non-negative matrix $\widetilde{\mathbf{A}}$ in (10), whose coefficients $\alpha_{j,v} \in (0,1)$ for all components. Take the $(ms \times ms)$ block-diagonal matrix $\lambda\mathbf{I} - \widetilde{\mathbf{A}}$, where

$\mathbf{I}$ denotes the identity matrix as usual, then $\lambda\mathbf{I} - \widetilde{\mathbf{A}} = diag(\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_m)$ with $(s \times s)$-matrix $\mathbf{C}_j$ $(j = 1, \ldots, m)$ defined by

$$\mathbf{C}_j = \begin{bmatrix} \lambda & 0 & \cdots & 0 & & -\alpha_{j,1} \\ 0 & \lambda & \cdots & 0 & & -\alpha_{j,2}\alpha_{j,1} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda & & -\prod_{k=0}^{s-2} \alpha_{j,s-1-k} \\ 0 & 0 & \cdots & 0 & & \lambda - \prod_{k=0}^{s-1} \alpha_{j,s-k} \end{bmatrix}.$$

The determinant of the matrix $\lambda\mathbf{I} - \widetilde{\mathbf{A}}$, denoted by $det(\lambda\mathbf{I} - \widetilde{\mathbf{A}})$, can easily be determined since the matrices $\mathbf{C}_j$ $(j = 1, \ldots, m)$ are upper triangular matrices (Harville 2008). The characteristic polynomial of $\widetilde{\mathbf{A}}$ is

$$det(\lambda\mathbf{I} - \widetilde{\mathbf{A}}) = \left(\lambda^{s-1}\right)^m \prod_{j=1}^{m} \left(\lambda - \prod_{k=0}^{s-1} \alpha_{j,s-k}\right).$$

For convenience in notation let $\prod_{k=0}^{s-1} \alpha_{j,s-k} =: T_j$. The polynomial takes the form

$$det(\lambda\mathbf{I} - \widetilde{\mathbf{A}}) = \lambda^{ms}\lambda^{-m} \prod_{j=1}^{m} (\lambda - T_j) = \lambda^{ms} + \sum_{j=1}^{m} (-1)^j \beta_j \lambda^{ms-j}$$

with coefficients β_j $(j = 1, \ldots, m)$ given by

$$\beta_1 = \sum_{j=1}^{m} T_j;\ \beta_2 = \sum_{j=1}^{m-1} \sum_{i=j+1}^{m} T_j T_i; \ldots;\ \beta_{m-1} = \sum_{j=1}^{m} \prod_{\substack{i=1 \\ i \neq j}}^{m} T_i;\ \beta_m = \prod_{j=1}^{m} T_j.$$

Let ρ be the maximal eigenvalue of $\widetilde{\mathbf{A}}$, then by Proposition B in Dion et al. (1995), $\sum_{j=1}^{m} \beta_j < 1$ if and only if $\rho < 1$.

Lemma 1 *For a fixed v $(v = 1, \ldots, s)$, $\alpha_{j,v} \in (0,1)$ where $j = 1, \ldots, m$ and for $t = v + ns$, $0 < P(\mathbf{Z}_t = \underline{0}) < 1$. Furthermore, any solution to $(\mathbf{X}_t)$, $t = v + ns$ and $n \in \mathbb{N}_0$ in (9) is an irreducible and aperiodic Markov chain.*

Proof Let $\underline{r} = [\underline{r}_1 \ \cdots \ \underline{r}_m]'$ where $\underline{r}_j = [r_{j1} \ \cdots \ r_{js}]$ and $\underline{d} = [\underline{d}_1 \ \cdots \ \underline{d}_m]'$ with $\underline{d}_j = [d_{j1} \ \cdots \ d_{js}]$ for each $j = 1, \ldots, m$. Note that

$$P_{\underline{r},\underline{d}} := P(\mathbf{X}_t = \underline{r} | \mathbf{X}_{t-s} = \underline{d}) =$$

$$= P\left(\left[\begin{array}{c} \mathbf{A}_1 \circ \mathbf{X}_{1,t-s} + \mathbf{B}_1 \circ \mathbf{Z}_{1,t} \\ \vdots \\ \mathbf{A}_m \circ \mathbf{X}_{m,t-s} + \mathbf{B}_m \circ \mathbf{Z}_{m,t} \end{array}\right] = \underline{r} \,\middle|\, \mathbf{X}_{t-s} = \underline{d}\right) =$$

$$= \sum_{j=1}^{m} \left(\sum_{i_{j1}=0}^{d_{j1}} \sum_{i_{j2}=0}^{d_{j2}} \cdots \sum_{i_{js}=0}^{d_{js}} \left[\prod_{v=1}^{s} P(Z_{j,v+ns} = i_{jv}) P \left(\prod_{k=0}^{v-1} \alpha_{j,v-k} \circ X_{m,s+(n-1)s} + \right.\right.\right.$$

$$\left.\left.\left. + \sum_{l=1}^{v-1} \left(\prod_{k=0}^{l-1} \alpha_{j,v-k} \right) \circ Z_{j,v-l+ns} = r_{jv} - i_{jv} | Z_{j,1+ns} = i_{j1}, \ldots, Z_{j,s+ns} = i_{js} \right) \right] \right)$$

$$\geq \sum_{j=1}^{m} \left(\prod_{v=1}^{s} P(Z_{j,v+ns} = r_{jv}) P \left(\prod_{k=0}^{v-1} \alpha_{j,v-k} \circ d_{js} + \right.\right.$$

$$\left.\left. + \sum_{l=1}^{v-1} \left(\prod_{k=0}^{l-1} \alpha_{j,v-k} \right) \circ Z_{j,v-l+ns} = 0 | Z_{j,1+ns} = r_{j1}, \ldots, Z_{j,s+ns} = r_{js} \right) \right)$$

$$\geq \sum_{j=1}^{m} \left(\prod_{v=1}^{s} P(Z_{j,v+ns} = r_{jv}) \left(1 - \prod_{k=0}^{v-1} \alpha_{j,v-k} \right)^{d_{js}} \times \right.$$

$$\left. \times \prod_{k=1}^{v-1} P \left(\left(\prod_{k=0}^{v-1} \alpha_{j,v-k} \right) \circ Z_{j,v-l+ns} = 0 | Z_{j,1+ns} = r_{j1}, \ldots, Z_{j,s+ns} = r_{js} \right) \right)$$

$$\geq \sum_{j=1}^{m} \left(\prod_{v=1}^{s} P(Z_{j,v+ns} = r_{jv}) \left(1 - \prod_{k=0}^{v-1} \alpha_{j,v-k} \right)^{d_{js}} \prod_{k=1}^{v-1} \left(1 - \prod_{k=0}^{v-1} \alpha_{j,v-k} \right)^{r_{jv-k}} \right)$$

$$> 0.$$

Therefore,

$$P_{\underline{0},\underline{d}} = \sum_{j=1}^{m} \left(\prod_{v=1}^{s} P(Z_{j,v+ns} = 0) \left(1 - \prod_{k=0}^{v-1} \alpha_{j,v-k} \right)^{d_{js}} \right) > 0$$

and similarly $P_{\underline{r},\underline{0}} = P(\mathbf{X}_t = \underline{r} | \mathbf{X}_{t-s} = \underline{0}) > 0$, leading to conclude that $(\mathbf{X}_t)$ is irreducible. Moreover,

$$P_{\underline{0},\underline{0}} = \sum_{j=1}^{m} \prod_{v=1}^{s} P(Z_{j,v+ns} = 0) > 0,$$

which, in turn, implies that for a fixed v $(v = 1, \ldots, s)$, the process $(\mathbf{X}_t)$ with $t = v + ns$ and $n \in \mathbb{N}_0$ is an aperiodic Markov chain. This concludes the proof.

The main result of this subsection is formalized through the theorem below.

Theorem 1 *For a fixed v $(v = 1, \ldots, s)$, let $(\mathbf{X}_t)$ with $t = v + ns$ and $n \in \mathbb{N}_0$ as in* (9) *be an irreducible, aperiodic Markov chain on $\mathbb{N}_0^m$. If $E||\mathbf{Z}_t||$ is finite and if the largest eigenvalue of $\widetilde{\mathbf{A}}$ is less than one, then there exists a strictly periodically stationary (or cyclostationary) m-variate* `INAR`(1) *process satisfying recursion* (9).

Proof From Lemma 1, $(\mathbf{X}_t)$ with $t = v + ns$ and fixed $v = 1, \dots, s$ is an irreducible and aperiodic Markov chain being the eigenvalues of matrix $\widetilde{\mathbf{A}}$ are less than one. Thus, by Franke and Subba Rao (1993) a strictly periodically stationary m-variate non-negative integer-valued process satisfying the equation (9) exists.

The `PMINAR`(1) model in (9) can be expressed as

$$\mathbf{X}_t = \widetilde{\mathbf{A}} \circ \mathbf{X}_{t-s} + \mathbf{R}_t, \tag{19}$$

where $\mathbf{R}_t = \widetilde{\mathbf{B}} \circ \mathbf{Z}_t$ with matrix $\widetilde{\mathbf{B}}$ in (11). Let

$$\mathbf{R}_t = [\mathbf{R}_{1,t} \ \mathbf{R}_{2,t} \ \cdots \ \mathbf{R}_{m,t}]' = [\mathbf{B}_1 \circ \mathbf{Z}_{1,t} \ \mathbf{B}_2 \circ \mathbf{Z}_{2,t} \ \cdots \ \mathbf{B}_m \circ \mathbf{Z}_{m,t}]'.$$

The innovation series $(\mathbf{R}_t)$ is a sequence of independent non-negative integer-valued random vectors with periodic structure.

2.2 Mean vector of cyclostationary PMINAR(1)

In this subsection, the periodic mean and autocovariance function of the `PMINAR`(1) model are derived. First note that the expectation of $\mathbf{R}_t$ is

$$E[\mathbf{R}_t] = E[\widetilde{\mathbf{B}} \circ \mathbf{Z}_t] = \widetilde{\mathbf{B}} E[\mathbf{Z}_t] = \widetilde{\mathbf{B}} \boldsymbol{\delta}_t$$

with matrices $\widetilde{\mathbf{B}}$ and $\boldsymbol{\delta}_t$ as in (11) and (12), respectively. Furthermore, for each component $j = 1, \dots, m$, the mean vector of $\mathbf{R}_{j,t}$ takes the form

$$E[\mathbf{R}_{j,t}] = \begin{bmatrix} \lambda_{j,1} \\ \lambda_{j,1}\alpha_{j,2} + \lambda_{j,2} \\ \lambda_{j,1}\alpha_{j,3}\alpha_{j,2} + \lambda_{j,2}\alpha_{j,3} + \lambda_{j,3} \\ \vdots \\ \lambda_{j,1} \prod_{k=0}^{s-2} \alpha_{j,s-k} + \lambda_{j,2} \prod_{k=0}^{s-3} \alpha_{j,s-k} + \cdots + \lambda_{j,s-1}\alpha_{j,s} + \lambda_{j,s} \end{bmatrix}.$$

Moreover,

$$\boldsymbol{\mu}_t = E[\mathbf{X}_t] = E[\widetilde{\mathbf{A}} \circ \mathbf{X}_{t-s} + \mathbf{R}_t] = (\mathbf{I} - \widetilde{\mathbf{A}})^{-1} \widetilde{\mathbf{B}} \boldsymbol{\delta}_t \tag{20}$$

with matrices $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$, and vector $\boldsymbol{\delta}_t$ as in (10), (11) and (12), respectively. Next we prove that $\mathbf{I} - \widetilde{\mathbf{A}}$ is regular and therefore $(\mathbf{I} - \widetilde{\mathbf{A}})^{-1}$ exists. Note that matrix $\mathbf{I} - \widetilde{\mathbf{A}}$ is a $(ms \times ms)$ block-diagonal matrix given by

$$\mathbf{I} - \widetilde{\mathbf{A}} = diag(\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_m)$$

with $(s \times s)$-matrix $\mathbf{C}_j$ $(j = 1, \dots, m)$ as

$$\mathbf{C}_j = \begin{bmatrix} 1 \; 0 \cdots 0 & -\alpha_{j,1} \\ 0 \; 1 \cdots 0 & -\alpha_{j,2}\alpha_{j,1} \\ \vdots \; \vdots \; \ddots \; \vdots & \vdots \\ 0 \; 0 \cdots 1 & -\prod_{k=0}^{s-2} \alpha_{j,s-1-k} \\ 0 \; 0 \cdots 0 & 1 - \prod_{k=0}^{s-1} \alpha_{j,s-k} \end{bmatrix}.$$

The determinant of the matrix $\mathbf{I} - \widetilde{\mathbf{A}}$ is

$$d := det(\mathbf{I} - \widetilde{\mathbf{A}}) = \prod_{j=1}^{m} \left(1 - \prod_{k=0}^{s-1} \alpha_{j,s-k} \right) \neq 0,$$

since $\alpha_{j,v} \in (0,1)$ for $j = 1, \dots, m$ and $v = 1, \dots, s$, leading to conclude that $\mathbf{I} - \widetilde{\mathbf{A}}$ is regular. The ms-dimensional mean vector $\boldsymbol{\mu}_t$ for $t = v + ns$; $v = 1, \dots, s$ and $n \in \mathbb{N}_0$ in (20) takes the form

$$\boldsymbol{\mu}_t = \left[\boldsymbol{\mu}_{1,t} \; \boldsymbol{\mu}_{2,t} \; \dots \; \boldsymbol{\mu}_{m,t} \right]',$$

where

$$\boldsymbol{\mu}_{j,t} = \begin{bmatrix} \lambda_{j,1} + \frac{d_{(-j)}}{d} \alpha_{j,1} \left(\lambda_{j,1} \prod_{k=0}^{s-2} \alpha_{j,s-k} + \dots + \lambda_{j,s-1}\alpha_{j,s} + \lambda_{j,s} \right) \\ \lambda_{j,1}\alpha_{j,2} + \lambda_{j,2} + \frac{d_{(-j)}}{d} \prod_{k=0}^{2-1} \alpha_{j,2-k} \left(\lambda_{j,1} \prod_{k=0}^{s-2} \alpha_{j,s-k} + \dots + \lambda_{j,s-1}\alpha_{j,s} + \lambda_{j,s} \right) \\ \vdots \\ \frac{d_{(-j)}}{d} \left(\lambda_{j,1} \prod_{k=0}^{s-2} \alpha_{j,s-k} + \lambda_{j,2} \prod_{k=0}^{s-3} \alpha_{j,s-k} + \dots + \lambda_{j,s-1}\alpha_{j,s} + \lambda_{j,s} \right) \end{bmatrix}$$

with $d_{(-j)} := \prod_{\substack{r=1 \\ r \neq j}}^{m} \left(1 - \prod_{k=0}^{s-1} \alpha_{r,s-k} \right)$. Now consider for each $j = 1, \dots, m$ and $l \geq i$,

$$\varphi_{l,i}^{(j)} = \begin{cases} \prod_{k=0}^{i-1} \alpha_{j,l-k}, & i \geq 1 \\ 1, & i = 0 \end{cases}. \tag{21}$$

It follows by tedious (although straightforward) calculations that for a fixed v and j, each entry in $\boldsymbol{\mu}_{j,t} = [E(X_{j,1+ns}) \; E(X_{j,2+ns}) \cdots E(X_{j,s+ns})]'$ is given by

$$E(X_{j,v+ns}) = \frac{\sum_{k=0}^{v-1} \varphi_{v,k}^{(j)} \lambda_{j,v-k} + \varphi_{v,v}^{(j)} \sum_{i=0}^{s-(v+1)} \varphi_{s,i}^{(j)} \lambda_{j,s-i}}{1 - \varphi_{s,s}^{(j)}}. \tag{22}$$

Hence, $\boldsymbol{\mu}_{j,t}$ can be expressed as

$$\boldsymbol{\mu}_{j,t} = \frac{1}{1-\varphi_{s,s}^{(j)}} \begin{bmatrix} \sum_{k=0}^{1-1} \varphi_{1,k}^{(j)} \lambda_{j,1-k} + \varphi_{1,1}^{(j)} \sum_{i=0}^{s-2} \varphi_{s,i}^{(j)} \lambda_{j,s-i} \\ \sum_{k=0}^{2-1} \varphi_{2,k}^{(j)} \lambda_{j,2-k} + \varphi_{2,2}^{(j)} \sum_{i=0}^{s-3} \varphi_{s,i}^{(j)} \lambda_{j,s-i} \\ \vdots \\ \sum_{k=0}^{s-1} \varphi_{s,k}^{(j)} \lambda_{j,s-k} + \varphi_{s,s}^{(j)} \sum_{i=0}^{s-(s+1)} \varphi_{s,i}^{(j)} \lambda_{j,s-i} \end{bmatrix},$$

for $j = 1,\ldots,m$; $t = v + ns$; $v = 1,\ldots,s$ and $n \in \mathbb{N}_0$. Here, we adopt the convention

$$\sum_{i=0}^{s-(s+1)} \varphi_{s,i}^{(j)} \lambda_{j,s-i} = 0.$$

2.3 Variance-covariance matrix

In order to derive the variance-covariance matrix of $(\mathbf{X}_t)$ we start by calculating the variance-covariance matrix $\sum_{\mathbf{R}_t}$ of $(\mathbf{R}_t)$. From lemma 1 in Franke and Rao (1993) [see also lemma 2.1 in Latour (1997)], it follows that

$$\begin{aligned} \sum\nolimits_{\mathbf{R}_t} &= Var[\mathbf{R}_t] = Var[\widetilde{\mathbf{B}} \circ \mathbf{Z}_t] = \\ &= Var[E(\widetilde{\mathbf{B}} \circ \mathbf{Z}_t|\mathbf{Z}_t)] + E[Var(\widetilde{\mathbf{B}} \circ \mathbf{Z}_t|\mathbf{Z}_t)] = \\ &= Var[\widetilde{\mathbf{B}}\mathbf{Z}_t] + diag(\mathbf{Q}E(\mathbf{Z}_t)) = \\ &= \widetilde{\mathbf{B}} \sum\nolimits_{\mathbf{Z}_t} \widetilde{\mathbf{B}}' + diag(\mathbf{Q}\boldsymbol{\delta}_t) \\ &= \widetilde{\mathbf{B}} \boldsymbol{\Psi}_t \widetilde{\mathbf{B}}' + diag(\mathbf{Q}\boldsymbol{\delta}_t) \end{aligned} \tag{23}$$

with matrices $\widetilde{\mathbf{B}}$, $\boldsymbol{\delta}_t$ and $\boldsymbol{\Psi}_t$ in (11), (12) and (15), respectively. Thus, matrix $\mathbf{Q}$ is also block-diagonal with m $(s \times s)$-matrices $\mathbf{Q}_j$, i.e., $\mathbf{Q} = diag(\mathbf{Q}_1,\ldots,\mathbf{Q}_m)$, where

$$\mathbf{Q}_j = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ \alpha_{j,2}(1-\alpha_{j,2}) & 0 & \ldots & 0 \\ \alpha_{j,3}\alpha_{j,2}(1-\alpha_{j,3}\alpha_{j,2}) & \alpha_{j,3}(1-\alpha_{j,3}) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \prod_{k=0}^{s-2} \alpha_{j,s-k} \left(1 - \prod_{k=0}^{s-2} \alpha_{j,s-k}\right) & \prod_{k=0}^{s-3} \alpha_{j,s-k} \left(1 - \prod_{k=0}^{s-3} \alpha_{j,s-k}\right) & \ldots & 0 \end{bmatrix},$$

leading to

$$diag(\mathbf{Q}\boldsymbol{\delta}_t) = \begin{bmatrix} \mathbf{Q}_1^* & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2^* & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{Q}_m^* \end{bmatrix}$$

with $\mathbf{Q}_j^* = diag(\mathbf{Q}_j \boldsymbol{\delta}_{j,t})$ for $j = 1, \ldots, m$. Hence

$$\mathbf{Q}_j \boldsymbol{\delta}_{j,t} = \begin{bmatrix} 0 \\ \alpha_{j,2}(1-\alpha_{j,2})\lambda_{j,1} \\ \alpha_{j,3}\alpha_{j,2}(1-\alpha_{j,3}\alpha_{j,2})\lambda_{j,1} + \alpha_{j,3}(1-\alpha_{j,3})\lambda_{j,2} \\ \vdots \\ \prod_{k=0}^{s-2} \alpha_{j,s-k} \left(1 - \prod_{k=0}^{s-2} \alpha_{j,s-k}\right) \lambda_{j,1} + \ldots + \alpha_{j,s}(1-\alpha_{j,s})\lambda_{j,s-1} \end{bmatrix}$$

and the variance-covariance matrix of $\mathbf{R}_t$ in (23) takes the form

$$\sum\nolimits_{\mathbf{R}_t} = \begin{bmatrix} \mathbf{B}_1\boldsymbol{\Psi}_{11,t}\mathbf{B}_1^{'} + \mathbf{Q}_1^* & \mathbf{B}_1\boldsymbol{\Psi}_{12,t}\mathbf{B}_2^{'} & \ldots & \mathbf{B}_1\boldsymbol{\Psi}_{1m,t}\mathbf{B}_m^{'} \\ \mathbf{B}_2\boldsymbol{\Psi}_{12,t}\mathbf{B}_1^{'} & \mathbf{B}_2\boldsymbol{\Psi}_{22,t}\mathbf{B}_2^{'} + \mathbf{Q}_2^* & \ldots & \mathbf{B}_2\boldsymbol{\Psi}_{2m,t}\mathbf{B}_m^{'} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_m\boldsymbol{\Psi}_{1m,t}\mathbf{B}_1^{'} & \mathbf{B}_m\boldsymbol{\Psi}_{2m,t}\mathbf{B}_2^{'} & \ldots & \mathbf{B}_m\boldsymbol{\Psi}_{mm,t}\mathbf{B}_m^{'} + \mathbf{Q}_m^* \end{bmatrix}.$$

Furthermore, for each component $j = 1, \ldots, m$,

$$Var[\mathbf{R}_{j,t}] = \mathbf{B}_j\boldsymbol{\Psi}_{jj,t}\mathbf{B}_j^{'} + \mathbf{Q}_j^*$$

and for $j \neq k, k = 1, \ldots, m$,

$$Cov(\mathbf{R}_{j,t}, \mathbf{R}_{k,t}) = \mathbf{B}_j\boldsymbol{\Psi}_{jk,t}\mathbf{B}_k^{'},$$

where combining (8) and (21) we have

$$\mathbf{B}_j = \begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ \varphi_{2,1}^{(j)} & 1 & 0 & \ldots & 0 \\ \varphi_{3,2}^{(j)} & \varphi_{3,1}^{(j)} & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varphi_{s,s-1}^{(j)} & \varphi_{s,s-2}^{(j)} & \varphi_{s,s-3}^{(j)} & \ldots & 1 \end{bmatrix}.$$

Recall from (19) that $\mathbf{R}_t = \widetilde{\mathbf{B}} \circ \mathbf{Z}_t$ and $\mathbf{Z}_t$ are independent of $\mathbf{X}_{t-s}$, the variance-covariance matrix, $\sum_{\mathbf{X}_t}$, of $\mathbf{X}_t$ is obtained from

$$\sum\nolimits_{\mathbf{X}_t} = \widetilde{\mathbf{A}} Var[\mathbf{X}_{t-s}]\widetilde{\mathbf{A}}^{'} + diag\left(\mathbf{D} \cdot E[\mathbf{X}_{t-s}]\right) + \sum\nolimits_{\mathbf{R}_t},$$

where $\mathbf{D}$ is a matrix with entries $\mathbf{D}_j$ for $j = 1, \ldots, m$,

$$\mathbf{D}_j = \begin{bmatrix} 0 & \cdots & 0 & \varphi_{1,1}^{(j)}\left(1-\varphi_{1,1}^{(j)}\right) \\ 0 & \cdots & 0 & \varphi_{2,2}^{(j)}\left(1-\varphi_{2,2}^{(j)}\right) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \varphi_{s,s}^{(j)}\left(1-\varphi_{s,s}^{(j)}\right) \end{bmatrix}.$$

Furthermore,

$$diag\left(\mathbf{D}\cdot E[\mathbf{X}_{t-s}]\right) \equiv diag(\mathbf{D}\boldsymbol{\mu}_t) = \begin{bmatrix} \mathbf{D}_1^* & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2^* & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{D}_m^* \end{bmatrix},$$

where

$$\mathbf{D}_j^* = \frac{\sum_{k=0}^{s-1} \varphi_{s,k}^{(j)}\lambda_{j,s-k}}{1-\varphi_{s,s}^{(j)}} \begin{bmatrix} \varphi_{1,1}^{(j)}\left(1-\varphi_{1,1}^{(j)}\right) & 0 & \dots & 0 \\ 0 & \varphi_{2,2}^{(j)}\left(1-\varphi_{2,2}^{(j)}\right) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \varphi_{s,s}^{(j)}\left(1-\varphi_{s,s}^{(j)}\right) \end{bmatrix}.$$

For simplicity in notation we define

$$\begin{aligned} \sum\nolimits_{\mathbf{X}_t} &= \begin{bmatrix} Var[\mathbf{X}_{1,t}] & Cov(\mathbf{X}_{1,t},\mathbf{X}_{2,t}) & \dots & Cov(\mathbf{X}_{1,t},\mathbf{X}_{m,t}) \\ Cov(\mathbf{X}_{2,t},\mathbf{X}_{1,t}) & Var[\mathbf{X}_{2,t}] & \dots & Cov(\mathbf{X}_{2,t},\mathbf{X}_{m,t}) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\mathbf{X}_{m,t},\mathbf{X}_{1,t}) & Cov(\mathbf{X}_{m,t},\mathbf{X}_{2,t}) & \dots & Var[\mathbf{X}_{m,t}] \end{bmatrix} \\ &=: \begin{bmatrix} \sum_{1,1} & \sum_{1,2} & \cdots & \sum_{1,m} \\ & \sum_{2,2} & \cdots & \sum_{2,m} \\ & & \ddots & \vdots \\ & & & \sum_{m,m} \end{bmatrix}. \end{aligned}$$

Note that

$$\sum\nolimits_{j,j} = \begin{bmatrix} Var[X_{j,1+ns}] & Cov(X_{j,1+ns},X_{j,2+ns}) & \dots & Cov(X_{j,1+ns},X_{j,s+ns}) \\ & Var[X_{j,2+ns}] & \dots & Cov(X_{j,2+ns},X_{j,s+ns}) \\ & & \ddots & \vdots \\ & & & Var[X_{j,s+ns}] \end{bmatrix}$$

with diagonal elements

$$Var[X_{j,v+ns}] = \frac{h(\varphi,\lambda)}{1-\left(\varphi_{s,s}^{(j)}\right)^2}, \tag{24}$$

where

$$\begin{aligned} h(\varphi,\lambda) := &\sum_{k=0}^{v-1} \varphi_{s,s}^{(j)}\varphi_{v,k}^{(j)}\lambda_{j,v-k} + \varphi_{v,k}^{(j)}\left(1-\varphi_{v,k}^{(j)}\right)\lambda_{j,v-k} + \left(\varphi_{v,k}^{(j)}\right)^2\sigma_{j,v-k}^2 + \\ &+ \sum_{m=0}^{s-(v+1)} \varphi_{s,s}^{(j)}\varphi_{v,v}^{(j)}\varphi_{s,m}^{(j)}\lambda_{j,s-m} + \varphi_{v,v}^{(j)}\varphi_{s,m}^{(j)}\left(1-\varphi_{v,v}^{(j)}\varphi_{s,m}^{(j)}\right)\lambda_{j,s-m} + \\ &+ \sum_{m=0}^{s-(v+1)} \left(\varphi_{v,v}^{(j)}\varphi_{s,m}^{(j)}\right)^2\sigma_{j,s-k}^2, \end{aligned}$$

for a fixed v $(v = 1, \ldots, s)$ and off-diagonal elements

$$Cov(X_{j,v+ns}, X_{j,v+ns+l}) = \varphi_{v+l,l}^{(j)} Var[X_{j,v+ns}]. \tag{25}$$

Moreoever,

$$\sum\nolimits_{j,k} = \begin{bmatrix} Cov(X_{j,1+ns}, X_{k,1+ns}) & Cov(X_{j,1+ns}, X_{k,2+ns}) & \ldots & Cov(X_{j,1+ns}, X_{k,s+ns}) \\ Cov(X_{j,2+ns}, X_{k,1+ns}) & Cov(X_{j,2+ns}, X_{k,2+ns}) & \ldots & Cov(X_{j,2+ns}, X_{k,s+ns}) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(X_{j,s+ns}, X_{k,1+ns}) & Cov(X_{j,s+ns}, X_{k,2+ns}) & \ldots & Cov(X_{j,s+ns}, X_{k,s+ns}) \end{bmatrix}$$

with diagonal elements

$$\begin{aligned} Cov(X_{j,v+ns}, X_{k,v+ns}) &= \frac{1}{1-\varphi_{s,s}^{(j)}\varphi_{s,s}^{(k)}} \sum_{i=0}^{v-1} \varphi_{v,i}^{(j)}\varphi_{v,i}^{(k)}\sigma_{jk,v-i} + \\ &+ \frac{\varphi_{v,v}^{(j)}\varphi_{v,v}^{(k)}}{1-\varphi_{s,s}^{(j)}\varphi_{s,s}^{(k)}} \sum_{i=0}^{s-(v+1)} \varphi_{s,i}^{(j)}\varphi_{s,i}^{(k)}\sigma_{jk,s-i}, \end{aligned}$$

for a fixed v $(v = 1, \ldots, s)$ and off-diagonal elements

$$\begin{aligned} Cov(X_{j,v+ns+h}, X_{k,v+ns}) &= \frac{\varphi_{v+h,h}^{(j)}}{1-\varphi_{s,s}^{(j)}\varphi_{s,s}^{(k)}} \sum_{i=0}^{v-1} \varphi_{v,i}^{(j)}\varphi_{v,i}^{(k)}\sigma_{jk,v-i} + \\ &+ \frac{\varphi_{v+h,h}^{(j)}\varphi_{v,v}^{(j)}\varphi_{v,v}^{(k)}}{1-\varphi_{s,s}^{(j)}\varphi_{s,s}^{(k)}} \sum_{i=0}^{s-(v+1)} \varphi_{s,i}^{(j)}\varphi_{v,i}^{(k)}\sigma_{jk,s-i} \end{aligned}$$

and

$$\begin{aligned} Cov(X_{j,v+ns}, X_{k,v+ns+h}) &= \frac{\varphi_{v+h,h}^{(k)}}{1-\varphi_{s,s}^{(j)}\varphi_{s,s}^{(k)}} \sum_{i=0}^{v-1} \varphi_{v,i}^{(j)}\varphi_{v,i}^{(k)}\sigma_{jk,v-i} + \\ &+ \frac{\varphi_{v+h,h}^{(k)}\varphi_{v,v}^{(j)}\varphi_{v,v}^{(k)}}{1-\varphi_{s,s}^{(j)}\varphi_{s,s}^{(k)}} \sum_{i=0}^{s-(v+1)} \varphi_{s,i}^{(j)}\varphi_{v,i}^{(k)}\sigma_{jk,s-i} \end{aligned}$$

with $\sigma_{jk,v}$ defined in (17). Now, for any positive lag h and each component $j = 1, \ldots, m$ it follows that

$$Cov(\mathbf{X}_{j,t}, \mathbf{X}_{j,t+h}) = \mathbf{A}_j^h Cov(\mathbf{X}_{j,t}, \mathbf{X}_{j,t}) = \mathbf{A}_j^h Var[\mathbf{X}_{j,t}], \tag{26}$$

and

$$\begin{aligned} Cov(\mathbf{X}_{j,t+h}, \mathbf{X}_{k,t}) &= \mathbf{A}_j^h Cov(\mathbf{X}_{j,t}, \mathbf{X}_{k,t}), \\ Cov(\mathbf{X}_{j,t}, \mathbf{X}_{k,t+h}) &= \mathbf{A}_k^h Cov(\mathbf{X}_{j,t}, \mathbf{X}_{k,t}). \end{aligned}$$

2.4 PMINAR(1) process with MVNB Innovations

In this subsection, we derive the first- and second-order moment structure of the `PMINAR`(1) process driven by periodic multivariate negative binomial (MVNB) innovations. First note that the joint probability mass function of the innovations is given by

$$P(Z_{1,v+ns} = z_1, \ldots, Z_{m,v+ns} = z_m) =$$
$$= \frac{\Gamma\left(\beta_v^{-1} + \sum_{j=1}^m z_j\right)}{\Gamma(\beta_v^{-1})} \left(\frac{\beta_v^{-1}}{\beta_v^{-1} + \sum_{j=1}^m \lambda_{j,v}}\right)^{\beta_v^{-1}} \left(\beta_v^{-1} + \sum_{j=1}^m \lambda_{j,v}\right)^{-\sum_{j=1}^m z_j} \cdot$$
$$\cdot \prod_{j=1}^m \frac{\lambda_{j,v}^{z_j}}{z_j!}, \ (z_1, \ldots, z_m) \in \mathbb{N}_0^m. \tag{27}$$

Notice the marginal distribution of $\mathbf{Z}_{j,t}$ is univariate negative binomial with parameters β_v^{-1} and $p_{j,v}$ $(j = 1, \ldots, m; v = 1, \ldots, s)$ with

$$p_{j,v} = \frac{\beta_v^{-1}}{\lambda_{j,v} + \beta_v^{-1}}.$$

As previously mentioned, the innovation process $(\mathbf{Z}_t)$, $t = v + ns;\ v = 1, \ldots, s$ and $n \in \mathbb{N}_0$ is generally defined as a periodic sequence of independent random vectors with mean as in (12) and variance-covariance matrix as in (15), respectively. Thus,

$$\lambda_{j,v} = E[Z_{j,v+ns}] = \beta_v^{-1}\ \frac{1 - p_{j,v}}{p_{j,v}}, \tag{28}$$

$$\sigma_{j,v}^2 = Var[Z_{j,v+ns}] = \beta_v^{-1}\ \frac{1 - p_{j,v}}{p_{j,v}^2} = \lambda_{j,v}(1 + \beta_v \lambda_{j,v}), \tag{29}$$

$$\sigma_{jk,v} = Cov(Z_{j,v+ns}, Z_{k,v+ns}) = \beta_v \lambda_{j,v} \lambda_{k,v}, \tag{30}$$

for a fixed v $(v = 1, \ldots, s)$, $j \neq k;\ j, k = 1, \ldots, m$. Note that $Var[Z_{j,v+ns}]$ is greater than $E[Z_{j,v+ns}]$, implying overdispersion. Thus, the first-order moment and the auto- and cross-covariance structure `PMINAR`(1) process are obtained from (22), (24) and (25) by plugging-in the values of $\lambda_{j,v}$, $\sigma_{j,v}^2$ and $\sigma_{jk,v}$ as in (28)-(30).

3 Parameter estimation

Consider a finite time series $(\mathbf{X}_{j,ts})$ with $1 \leq t \leq N,\ j = 1, \ldots, m$ (N-number of complete cycles) from the `PMINAR`(1) model in (19) with MVNB innovations. Without loss of generality it will be assumed that $\mathbf{X}_0 = \mathbf{x}_0$. The vector of parameters $\boldsymbol{\theta}$ is a $(2m+1)s$-dimensional vector

$$\boldsymbol{\theta} := (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_m, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_m, \boldsymbol{\beta}) \tag{31}$$

with s-vectors ($j = 1, \ldots, m$):

$$\boldsymbol{\alpha}_j = (\alpha_{j,1}, \ldots, \alpha_{j,s}) \; ; \; \boldsymbol{\lambda}_j = (\lambda_{j,1}, \ldots, \lambda_{j,s}) \; ; \; \boldsymbol{\beta} = (\beta_1, \ldots, \beta_s).$$

In order to estimate the unknown parameters in $\boldsymbol{\theta}$, three estimation methods are proposed, namely: Yule-Walker (YW), conditional maximum likelihood (CML) and composite likelihood (CL).

3.1 Yule-Walker estimation

The YW estimator of $\boldsymbol{\theta}$, $\widehat{\boldsymbol{\theta}}_{YW} := (\widehat{\boldsymbol{\alpha}}_1^{YW}, \ldots, \widehat{\boldsymbol{\alpha}}_m^{YW}, \widehat{\boldsymbol{\lambda}}_1^{YW}, \ldots, \widehat{\boldsymbol{\lambda}}_m^{YW}, \widehat{\boldsymbol{\beta}}^{YW})$ are calculated as follows: first, the YW estimators of parameters $\boldsymbol{\lambda}_j$ are calculated through the solution of the system of s linear equations yielding

$$\widehat{\lambda}_{j,v}^{YW} = \begin{cases} \overline{X}_{j,v} - \widehat{\alpha}_{j,v}^{YW} \overline{X}_{j,s} & , v = 1 \\ \overline{X}_{j,v} - \widehat{\alpha}_{j,v}^{YW} \overline{X}_{j,v-1} & , v = 2, 3, \ldots, s \end{cases},$$

where

$$\overline{X}_{j,v} = \frac{1}{N} \sum_{n=0}^{N-1} X_{j,v+ns}, \; j = 1, \ldots, m.$$

Further, from relation (26) and taking lag $h = 1$, it follows that $\boldsymbol{\gamma}_{j,t}(1)$ equals $\mathbf{A}_j \boldsymbol{\gamma}_{j,t}(0)$ and therefore, the YW estimators of parameters $\boldsymbol{\alpha}_j$ are

$$\widehat{\alpha}_{j,v}^{YW} = \begin{cases} \frac{S_{j,v}^2}{S_{j,s}^2} & , v = 1 \\ \\ \frac{\gamma_{j,v-1}(1)}{S_{j,v-1}^2} & , v = 2, 3, \ldots, s \end{cases}, \tag{32}$$

where $S_{j,v}^2$ and $\gamma_{j,v}(1)$ ($j = 1, \ldots, m$) are defined as

$$S_{j,v}^2 = \frac{1}{N-1} \sum_{n=0}^{N-1} (X_{j,v+ns} - \overline{X}_{j,v})^2,$$

and

$$\begin{aligned} &\gamma_{j,v}(1) = Cov(X_{j,v+ns}, X_{j,v+1+ns}) = \\ &= \begin{cases} \frac{1}{N-1} \sum\limits_{n=0}^{N-1} (X_{j,v+ns} - \overline{X}_{j,v})(X_{j,v+1+ns} - \overline{X}_{j,v+1}) & , v = 1, \ldots, s-1 \\ \\ \frac{1}{N-1} \sum\limits_{n=0}^{N-1} (X_{j,v+ns} - \overline{X}_{j,v})(X_{j,1+(n+1)s} - \overline{X}_{j,1}^*) & , v = s \end{cases}, \end{aligned}$$

where $\overline{X}_{j,1}^* = \frac{1}{N} \sum\limits_{n=0}^{N} X_{j,1+ns}$. Finally, the YW estimator of $\boldsymbol{\beta}$ is

$$\widehat{\beta}_v^{YW} = \frac{\left(1 - \widehat{\varphi}_{s,s}^{(j)} \widehat{\varphi}_{s,s}^{(k)}\right) \gamma_{jk,v}(0)}{\sum\limits_{i=0}^{v-1} \widehat{\varphi}_{v,i}^{(j)} \widehat{\varphi}_{v,i}^{(k)} \widehat{\lambda}_{j,v-i} \widehat{\lambda}_{k,v-i} + \widehat{\varphi}_{v,v}^{(j)} \widehat{\varphi}_{v,v}^{(k)} \sum\limits_{i=0}^{s-(v+1)} \widehat{\varphi}_{s,i}^{(j)} \widehat{\varphi}_{s,i}^{(k)} \widehat{\lambda}_{j,s-i} \widehat{\lambda}_{k,s-i}},$$

for $v = 1, \dots, s$ and $j, k = 1, \dots, m$ ($j \neq k$). The Yule-Walker estimators of φ are the product of estimators $\widehat{\alpha}_{j,v}^{YW}$ defined in (32).

3.2 Conditional maximum likelihood estimation

In order to obtain the CML estimator $\widehat{\boldsymbol{\theta}}_{CML}$ of $\boldsymbol{\theta}$ we proceed as follows. First note that the transition probabilities for the PMINAR(1) model can be expressed as the convolution of m binomials with parameters $(x_{j,v-1+ns}, \alpha_{j,v})$ where $j = 1, \dots, m$; $v = 1, \dots, s$ with probability mass function

$$f_j(r_j) = C_{r_j}^{x_{j,v-1+ns}} \alpha_{j,v}^{r_j} (1 - \alpha_{j,v})^{x_{j,v-1+ns} - r_j}, \tag{33}$$

and the periodic discrete m-variate distribution defined as in (27). Thus, the conditional density is the multiple sum

$$\begin{aligned}
&p_v(\mathbf{x}_{v+ns}|\mathbf{x}_{v-1+ns}) = P(\mathbf{X}_{v+ns} = \mathbf{x}_{v+ns}|\mathbf{X}_{v-1+ns} = \mathbf{x}_{v-1+ns}) = \\
&= \sum_{r_1=0}^{g_1} \sum_{r_2=0}^{g_2} \cdots \sum_{r_m=0}^{g_m} \left(\prod_{j=1}^{m} f_j(r_j) \right) \frac{\Gamma\left(\beta_v^{-1} + \sum_{j=1}^{m} (x_{j,v+ns} - r_j)\right)}{\Gamma(\beta_v^{-1})} \cdot \\
&\cdot \left(\frac{\beta_v^{-1}}{\beta_v^{-1} + \sum_{j=1}^{m} \lambda_{j,v}} \right)^{\beta_v^{-1}} \left(\beta_v^{-1} + \sum_{j=1}^{m} \lambda_{j,v} \right)^{-\sum_{j=1}^{m} (x_{j,v+ns} - r_j)} \cdot \\
&\cdot \prod_{j=1}^{m} \frac{\lambda_{j,v}^{(x_{j,v+ns} - r_j)}}{(x_{j,v+ns} - r_j)!}
\end{aligned} \tag{34}$$

with $g_j := min(x_{j,v-1+ns}, x_{j,v+ns})$ and $f_j(r_j)$ as in (33) for $j = 1, \dots, m$; $v = 1, \dots, s$ and $n \in \mathbb{N}_0$. Hence, the CML estimator is obtained by maximizing the conditional log-likelihood

$$ln(L(\boldsymbol{\theta}|\mathbf{x})) = \sum_{n=0}^{N-1} \sum_{v=1}^{s} ln\left(p_v(\mathbf{x}_{v+ns}|\mathbf{x}_{v-1+ns})\right)$$

with transition probabilities $p_v(\mathbf{x}_{v+ns}|\mathbf{x}_{v-1+ns})$ in (34). Explicit CML estimators are not available so numerical procedures have to be employed. The asymptotic properties of $\widehat{\boldsymbol{\theta}}_{CML}$ are given through the following result.

Theorem 2 *The conditional maximum likelihood estimator* $\widehat{\boldsymbol{\theta}}_{CML}$ *of* $\boldsymbol{\theta}$ *is asymptotically normal*

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

where $I(\boldsymbol{\theta})$ *represents the Fisher information matrix*

$$I = \begin{bmatrix} \mathbf{M}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{M}_s \end{bmatrix}$$

with matrices $\mathbf{M}_v$ $(v = 1, \ldots, s)$ *given by*

$$(-1) \times \begin{bmatrix} E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{1,v}^2}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{1,v}\partial \alpha_{m,v}}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{1,v}\partial \lambda_{1,v}}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{1,v}\partial \lambda_{m,v}}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{1,v}\partial \beta_v}\right] \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{m,v}\partial \alpha_{1,v}}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{m,v}^2}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{m,v}\partial \lambda_{1,v}}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{m,v}\partial \lambda_{m,v}}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \alpha_{m,v}\partial \beta_v}\right] \\ E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{1,v}\partial \alpha_{1,v}}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{1,v}\partial \alpha_{m,v}}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{1,v}^2}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{1,v}\partial \lambda_{m,v}}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{1,v}\partial \beta_v}\right] \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{m,v}\partial \alpha_{1,v}}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{m,v}\partial \alpha_{m,v}}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{m,v}\partial \lambda_{1,v}}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{m,v}^2}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \lambda_{m,v}\partial \beta_v}\right] \\ E\left[\frac{\partial^2 C(\theta)}{\partial \beta_v\partial \alpha_{1,v}}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \beta_v\partial \alpha_{m,v}}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \beta_v\partial \lambda_{1,v}}\right] & \cdots & E\left[\frac{\partial^2 C(\theta)}{\partial \beta_v\partial \lambda_{m,v}}\right] & E\left[\frac{\partial^2 C(\theta)}{\partial \beta_v^2}\right] \end{bmatrix}.$$

Proof This result is a particular case of theorem 2.2 in Billingsley (1961). For each v, with $v = 1, \ldots, s$, $p_v(\cdot|\cdot)$ is the transition probabilities in (34) of the `PMINAR`(1) model, therefore the regularity conditions in Billingsley's theorem are satisfied. We postpone those assumptions to the Appendix A.

3.3 Composite likelihood estimation

For periodic multivariate processes, the number of parameters can be quite large. The inflation of parameters is due to season v $(v = 1, \ldots, s)$ with s representing the period. Computational issues often arise when applying the conditional maximum likelihood approach, the complexity of the method augments with dimensional increase. To overcome the limitations in computing the exact likelihood, Lindsay (1988) proposed the composite likelihood as a pseudo-likelihood for inference. The pseudo-likelihood may take various forms such as combinations of likelihoods for small subsets of the data or combinations of conditional likelihoods. These procedures adopt some features of the full likelihood which are useful for inference while keeping the computation feasible, producing answers in reasonable time.

Composite likelihood inherits many of the good properties of inference based on the full likelihood function, but is more easily implemented with high-dimensional data sets. Pairwise likelihood is one special case of a composite likelihood, in which the pseudo-likelihood is defined as the product of the bivariate likelihood of all possible pairs of observations. A more general discussion of pairwise likelihood can be found in Davis and Yau (2011). Issues and strategies in the selection of composite likelihoods are given in Lindsay et al. (2011).

Composite likelihood methods based on optimizing sums of log-likelihoods of low-dimensional margins have become popular in recent years (Pedeli and Karlis 2013a); being useful for multivariate models in which the likelihood of multivariate data is very time-consuming. The methodology has drawn considerable attention in a broad range of applied disciplines in which complex data structures arise (e.g. Varin et al. 2011; Varin 2008). Asymptotic results and computational aspects of construction of, and inference from, composite likelihood are available in Varin et al. (2011). Analogues of the Akaike information criteria for model selection can be derived in the framework of composite likelihoods, having a similar form, see e.g. Ng and Joe (2014).

Note that the bivariate marginal log-likelihood function between two random elements, say X_a and X_b, can be defined as

$$l_{ab}(\boldsymbol{\theta};\mathbf{x}_a,\mathbf{x}_b)=\frac{1}{Ns}\sum_{n=0}^{N-1}\sum_{v=1}^{s} log f_{X_a,X_b}(x_{a,v+ns},x_{b,v+ns}|x_{a,v-1+ns},x_{b,v-1+ns};\boldsymbol{\theta}),$$

where

$$\begin{aligned}&f_{X_a,X_b}(x_{a,v+ns},x_{b,v+ns}|x_{a,v-1+ns},x_{b,v-1+ns};\boldsymbol{\theta})=\\&=\sum_{k_a=0}^{g_1}\sum_{k_b=0}^{g_2}\binom{x_{a,v-1+ns}}{x_{a,v+ns}-k_a}\alpha_{a,v}^{x_{a,v+ns}-k_a}(1-\alpha_{a,v})^{x_{a,v-1+ns}-x_{a,v+ns}+k_a}\cdot\\&\cdot\binom{x_{b,v-1+ns}}{x_{b,v+ns}-k_a}\alpha_{b,v}^{x_{b,v+ns}-k_b}(1-\alpha_{b,v})^{x_{b,v-1+ns}-x_{b,v+ns}+k_b}\cdot h_{R_a,R_b}(k_a,k_b),\end{aligned}$$

for $g_1 := min(x_{a,v+ns}, x_{a,v-1+ns})$ and $g_2 := min(x_{b,v+ns}, x_{b,v-1+ns})$. The bivariate function $h_{R_a,R_b}(k_a,k_b)$ represents the bivariate marginal probability density with bivariate negative binomial innovations being a particular case ($m = 2$) of the multivariate negative binomial distribution in (27).

The composite log-likelihood function, $cl(\boldsymbol{\theta};\mathbf{x}_a,\mathbf{x}_b)$, then arises as the sum of all bivariate log-likelihood functions, i.e.,

$$cl(\boldsymbol{\theta};\mathbf{x}_a,\mathbf{x}_b)=\sum_{a=1}^{m-1}\sum_{b=a+1}^{m} w_{ab}\ l_{ab}(\boldsymbol{\theta};\mathbf{x}_a,\mathbf{x}_b),$$

where w_{ab} is a constant weight for l_{ab}. Typically, the weights are chosen in order to eliminate distant pairs of observations, which should be less informative (Varin and Vidoni 2005). For sake of simplicity, it is common to set w_{ab} equal to 1, $1 \le a \le b \le m$. Further details on weighting of bivariate margins in pairwise likelihood can be found in Joe and Lee (2009).

3.4 Simulation study

The performance of the three estimation methods (YW, CML, CL) for the `PMINAR`(1) model driven by multivariate negative binomial innovations are compared in this subsection through a simulation study for $m = 3$ (trivariate) thus the vector of unknown parameters $\boldsymbol{\theta}$ in (31) is $(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\boldsymbol{\alpha}_3,\boldsymbol{\lambda}_1,\boldsymbol{\lambda}_2,\boldsymbol{\lambda}_3,\boldsymbol{\beta})$.

Further, it will be assumed throughout the study that $s = 4$. This simulation study contemplates the three scenarios displayed in Table 1. Notice that for set A, parameters $\boldsymbol{\alpha}$ are mainly above 0.5. For set B, parameters $\boldsymbol{\alpha}$ take values below and above 0.5 and for set C, parameters take at most the value 0.5. Correlations between series for set A are 0.52 (series 1 and 2), 0.71 (series 1 and 3) and 0.64 (series 2 and 3). For set B, 0.48, 0.35 and 0.28 and for set C, 0.74, 0.18 and 0.37, respectively.

(Table 1 about here)

Three alternative samples sizes are considered: $n = 400, 1000, 2000$. Thus $n = sN$, $N = 100, 250, 500$ complete cycles. For each experiment we conducted 200 independent replications.

The simulated data sets that produced YW estimates in an inadmissible range were disregarded and iterations were continued till reaching the specified number of 200 replications per experiment. The tendency of the YW method to produce inadmissible estimates was greater for smaller sample sizes. YW estimates were used as initial values in numerical routines for the optimization procedure of CML and CL methods.

Comparison of the YW, CML and CL estimators was carried out in terms of their corresponding mean square error (MSE) and the biases of the produced point estimates. For set A, Table 2 displays point estimates for the parameters of the periodic trivariate $\texttt{INAR}(1)$ model with trivariate negative binomial innovations. Each point estimate includes MSE in parenthesis. For sets B and C, the corresponding point estimates are summarized in Tables 3 and 4, respectively. Those tables are available in Appendix B.

(Table 2 about here)

For set A, Table 2 reports the estimates for autocorrelation parameters $\boldsymbol{\alpha}_j$ $(j = 1, 2, 3)$ where small MSEs characterize all estimates of $(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \boldsymbol{\alpha}_3)$. The performance of the estimators $\hat{\boldsymbol{\lambda}}_j$ $(j = 1, 2, 3)$ and estimator $\hat{\boldsymbol{\beta}}$ is slightly worse. The same behavior can be observed for sets B and C in Tables 3-4 (see Appendix B). The YW estimator does not perform well in general, revealing to be a not so good estimator for the dispersion parameter $\boldsymbol{\beta}$. The estimates obtained by adopting either the CML or the CL method are very close to the real parameter values, even in the case of a moderate sample size $(n = 400)$. For larger samples $(n = 1000$ and $n = 2000)$ both estimators seem to perform well and in a similar fashion.

Graphical inspection is provided through the boxplots of the biases of the produced estimates. Regarding set A, Figure 1 displays the boxplots of biases' estimates of parameters $\boldsymbol{\alpha}_j = (\alpha_{j,1}, \alpha_{j,2}, \alpha_{j,3}, \alpha_{j,4})$, for $j = 1, 2, 3$. Furthermore, Figures 2-3 refer to the boxplots of biases' estimates for the parameters of the trivariate negative binomial distributed innovation distribution

$\boldsymbol{\lambda}_j = (\lambda_{j,1}, \lambda_{j,2}, \lambda_{j,3}, \lambda_{j,4})$, for $j = 1, 2, 3$, and $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)$, respectively. Likeliwise, boxplots regarding set B are in Figures 4-6 and regarding set C in Figures 7-9. Due to space constraints, figures related to sets B and C are displayed in Appendix B. Also, the effect of sample size on the behavior of the estimators can be checked throughout Figures 1-9.

(Figures 1-3 about here)

As expected, increasing the sample size improves the performance of all estimators in terms of both location (median closer to zero) and dispersion (narrower interquartile ranges). Therefore, this indicates the superiority of CML and CL estimators over the YW estimators.

Closing this section, it is worth mentioning that numerical maximization of the conditional maximum likelihood is very time-consuming. The composite likelihood method was suggested in order to overcome the computational difficulties of the conditional maximum likelihood approach in multivariate models. The main advantage of the CL approach is the replacement of the full likelihood with a pseudo-likelihood which effectively captures the model properties while at the same time is computationally less demanding. Through this simulation study, the CL method revealed a good performance within a reasonable amount of time. One estimate with the CL method took about 25 sec, 66 sec and 129 sec for the three sample sizes $n = 400, 1000, 2000$ respectively; with the CML method the time increased to 54 sec, 162 sec and 324 sec. The CL method requires significantly less time for the optimization of the likelihood function without obvious losses in precision.

4 Forecasting

In this section we consider the problem of predicting the future values $\mathbf{X}_{t+h}$ $(t = v + ns;\ v = 1, \ldots, s)$ of the periodic `MINAR`(1) process given past observations through time $t = v + ns$ for $v = 1, \ldots, s$. Set $h = u + ls$ for $u = 1, \ldots, s$. Due to model's definition and by iterating equation (5), the j-th component $\mathbf{X}_{j,t}$ can be expressed as

$$\mathbf{X}_{j,t} \stackrel{d}{=} \left(\prod_{i=0}^{n-1} \boldsymbol{\phi}_{j,t-i}\right) \circ \mathbf{X}_{j,t-n} + \sum_{k=1}^{n-1} \left(\prod_{i=0}^{k-1} \boldsymbol{\phi}_{j,t-i}\right) \circ \mathbf{Z}_{j,t-k} + \mathbf{Z}_{j,t}$$

with $\mathbf{Z}_{j,t}$ in (2) and $\boldsymbol{\phi}_{j,t}$ defined in (3). Then each element of $\mathbf{X}_{j,t}$ for a fixed t can be written as

$$X_{j,t} \stackrel{d}{=} \zeta_{t,n}^{(j)} \circ X_{j,t-n} + \sum_{k=0}^{n-1} \zeta_{t,k}^{(j)} \circ Z_{j,t-k},$$

where, for $t \geq i$

$$\zeta_{t,i} := \begin{cases} \prod_{k=0}^{i-1} \phi_{t-k} & , i > 0 \\ 1 & , i = 0 \end{cases}$$

and also $\zeta_{t,i} := \zeta_{t,v}(\zeta_{s,s})^l$ for $i = v + ls;\ v = 1, \dots, s$, leading to

$$X_{j,v+ns+h} \stackrel{d}{=} \zeta^{(j)}_{v+ns+h,h} \circ X_{j,v+ns} + \sum_{k=0}^{h-1} \zeta^{(j)}_{v+ns+h,k} \circ Z_{j,v+ns+h-k}.$$

Since $h = u + ls$ for $u = 1, \dots, s$, it follows that

$$\begin{aligned} X_{j,v+ns+h} &\stackrel{d}{=} \zeta^{(j)}_{v+u+(n+l)s,u+ls} \circ X_{j,v+ns} + \sum_{k=1}^{u+ls-1} \zeta^{(j)}_{v+u+(n+l)s,k} \circ Z_{j,v+u+(n+l)s-k} \\ &\stackrel{d}{=} \zeta^{(j)}_{v+u,u} \left(\zeta^{(j)}_{s,s}\right)^l \circ X_{j,v+ns} + Y_{j,v+u+ls} \end{aligned}$$

with

$$\begin{aligned} Y_{j,v+u+ls} = &\sum_{k=0}^{v-1} \zeta^{(j)}_{v+u,k} \circ Z_{j,v+u+ns-k} + \\ &+ \sum_{w=0}^{l-1} \sum_{k=0}^{s-1} \zeta^{(j)}_{v+u+(n+l)s,k+u+ws} \circ Z_{j,v+(n+l-w)s-k}. \end{aligned}$$

As usual, to generate the h-step ahead prediction we can employ the mean, median or mode of the predictive distribution of $\mathbf{X}_{v+ns+h}|\mathbf{X}_{v+ns}$ as a point forecast. The median and mode are considered as coherent predictions but the mean is not. The h-step ahead point predictor that minimizes the mean square error (MSE) is given by

$$\begin{aligned} \widehat{X}_{j,v+ns+h} &= E[X_{j,v+ns+h}|X_{j,v+ns}] \\ &= E\left[\zeta^{(j)}_{v+u,u} \left(\zeta^{(j)}_{s,s}\right)^l \circ X_{j,v+ns}|X_{j,v+ns}\right] + E[Y_{j,v+u+ls}], \end{aligned}$$

where

$$\begin{aligned} &E[Y_{j,v+u+ls}] = \\ &= E\left[\sum_{k=0}^{v-1} \zeta^{(j)}_{v+u,k} \circ Z_{j,v+u+ns-k} + \sum_{w=0}^{l-1} \sum_{k=0}^{s-1} \zeta^{(j)}_{v+u+(n+l)s,k+u+ws} \circ Z_{j,v+(n+l-w)s-k}\right] \\ &= \sum_{k=0}^{v-1} \zeta^{(j)}_{v+u,k} \lambda_{j,v+u-k} + \sum_{w=0}^{l-1} \sum_{k=0}^{s-1} \zeta^{(j)}_{v+u+(n+l)s,k+u+ws} \lambda_{j,v+(n+l-w)s-k} \end{aligned} \tag{35}$$

with $E[Z_{j,v+ns}] = \lambda_{j,v}$ in (14). For the particular case, $h = 1$, the one-step ahead predictive function is

$$\begin{aligned} &p_v(\mathbf{x}_{v+ns+1}|\mathbf{x}_{v+ns}) = \\ &= \sum_{r_1=0}^{g_1} \sum_{r_2=0}^{g_2} \dots \sum_{r_m=0}^{g_m} \left(\prod_{j=1}^{m} f_j(r_j)\right) k(x_{1,v+ns+1} - r_1, \dots, x_{m,v+ns+1} - r_m) \end{aligned}$$

with $g_j = min(x_{j,v+ns}, x_{j,v+ns+1}),\ j = 1, \ldots, m$. In particular, for the MVNB distribution defined in (27) function $k(\cdot)$ above takes the form

$$k(x_{1,v+ns+1} - r_1, x_{2,v+ns+1} - r_2, \ldots, x_{m,v+ns+1} - r_m) =$$
$$= \frac{\Gamma\left(\beta_v^{-1} + \sum_{j=1}^{m}(x_{j,v+ns+1} - r_j)\right)}{\Gamma(\beta_v^{-1})} \left(\frac{\beta_v^{-1}}{\beta_v^{-1} + \sum_{j=1}^{m} \lambda_{j,v}}\right)^{\beta_v^{-1}} \times$$
$$\times \left(\beta_v^{-1} + \sum_{j=1}^{m} \lambda_{j,v}\right)^{-\sum_{j=1}^{m}(x_{j,v+ns+1} - r_j)} \prod_{j=1}^{m} \frac{\lambda_{j,v}^{(x_{j,v+ns+1} - r_j)}}{(x_{j,v+ns+1} - r_j)!}.$$

Furthermore, from equations (22) and (35), the one-step ahead predictor of $X_{j,v+ns+1}$ takes the form

$$\widehat{X}_{j,v+ns+1} = E[X_{j,v+ns+1} | X_{j,v+ns}]$$
$$= \frac{\sum_{k=0}^{v} \varphi_{v+1,k}^{(j)} \lambda_{j,v+1-k} + \varphi_{v+1,v+1}^{(j)} \sum_{i=0}^{s-(v+2)} \varphi_{s,i}^{(j)} \lambda_{j,s-i}}{1 - \varphi_{s,s}^{(j)}} +$$
$$+ \sum_{k=0}^{v-1} \zeta_{v+1,k}^{(j)} \lambda_{j,v+1-k} + \sum_{w=0}^{l-1} \sum_{k=0}^{s-1} \zeta_{v+1+(n+l)s,k+1+ws}^{(j)} \lambda_{j,v+(n+l-w)s-k}.$$

5 An application to time series of fire activity

This section illustrates the performance of the PMINAR(1) model to the analysis of a trivariate real environmental data set. The data refers to the number of fires collected in three counties in Portugal, namely Anadia, Oliveira do Bairro (O.Bairro) and Vagos, during 32 consecutive years, from 1986 to 2017 (www2.icnf.pt/portal/florestas/dfci/inc/estat-sgif). The data represents monthly observations of daily fires in those counties. This collection of fires is showed in Figure 10.

(Figure 10 about here)

Forest fires are a major problem in many European countries of the northern arch of the Mediterranean Sea, namely Portugal, Spain, France, Italy and Greece, as they are a threat not only to forests but also to people and their surroundings. In Europe, Portugal is the country with the highest number of forest fires per unit surface and per number of inhabitants (San Miguel-Ayanz and Camia 2009). Fire frequency is markedly different from north to south and from east to west (Nunes et al. 2016; Nunes 2012). The distribution of fires across the year follows a regular pattern, strongly influenced by seasonal variations of temperature and rainfall. Hence, it is expected to find the highest number of fires in the summer season, with a peak in July/August and the lowest number of fires in the rainy season. For further details see Tonini et al.

(2017) and Scotto et al. (2014a). The sample autocorrelation function (ACF) in Figure 11 reveals a periodic pattern of 12 months.

(Figure 11 about here)

The mean values and variances of the number of fires per month are shown in Figure 12 and cross-correlations in Figure 13. In the three counties, most months have variance greater than the mean, implying overdispersion.

(Figure 12 about here)

(Figure 13 about here)

As stressed above, the distribution for the innovations is assumed to be trivariate negative binomial. It is worth to mention that for this particular application, the Yule-Walker approach produces non-admissible point estimates for some months and, hence, is disregarded for further analysis. Table 5 summarizes the CL and CML estimates and the corresponding standard errors (SE) obtained by fitting the periodic trivariate `INAR`(1) model with period $s = 12$ and trivariate negative binomial innovations. The SE are calculated numerically from the Hessian matrix during the optimization procedure in `R`.

(Table 5 about here)

In many cases, the estimates from both methods (CL and CML) are very close. Some loss of efficiency is noticed when the CL method is employed although it is important to emphasize here that the CL is an approximate likelihood, leading to inevitable losses. The CL method could be regarded as a satisfactory approach for the estimation of the unknown parameters of the `PMINAR`(1) process, especially when other alternatives are not available. The CL estimates were used to initialize the CML method. Some estimates of the autocorrelation parameters in Table 5 are not statistically significant, suggesting that on those months the number of fires is being mainly modeled through the innovation process.

In order to check the adequacy of the periodic trivariate model for the considered data of the monthly number of fires in Anadia, O.Bairro and Vagos counties, two approaches will be adopted: residual-based and parametric bootstrap-based methods. The use of the parametric bootstrap method relies on a resampling exercice proposed by Tsay (1992). Such techniques to assess model adequacy have also been employed by Ristić and Popović (2019).

Concerning the residual-based method, the standardized Pearson residuals are calculated through the expression

$$R^*_{j,t} = \frac{X_{j,t} - E[X_{j,t}|X_{j,t-1}]}{\sqrt{Var[X_{j,t}|X_{j,t-1}]}},$$

where $j = 1, 2, 3, t = v + ns$ and $v = 2, \ldots, s$. The analysis of the standardized residuals for the trivariate model reveal sample means of -0.003 for Anadia, -0.001 for O.Bairro and -0.007 for Vagos county. The sample variances are quite close to one namely, 0.97, 1.01 and 1.03 for Anadia, O.Bairro and Vagos counties, respectively. Figure 14 displays the sample autocorrelation function of Pearson residuals in Anadia, O.Bairro and Vagos counties. Overall, the results are consistent with the assumptions of zero-mean, unit variance and uncorrelatedness of the standardized residuals.

(Figure 14 about here)

Pursing with the parametric bootstrap method, we generated 5000 artificial data sets using the fitted periodic trivariate `INAR`(1) model with period $s = 12$ and trivariate negative binomial innovations with CML estimates as in Table 5. Based on these bootstrap data sets, we compared some fitted characteristics. Namely, for each obtained simulated sample, the sample mean and the sample standard deviation were calculated. Bootstrap confidence intervals for such characteristics based on their corresponding 2.5% and 97.5% quantiles were obtained and graphically displayed in Figures 15 and 16, respectively.

(Figure 15 about here)

(Figure 16 about here)

Considering the bootstrap data sets, 5000 autocorrelation functions were computed. For each fixed lag, the sample ACF, 2.5% and 97.5% quantiles were determined to constitute the lower and upper bounds, displayed in Figure 17. Almost all of the sample autocorrelation values lie within the confidence bounds. Note that in this case we also need to compute the sample pairwise cross-correlation function between the three time series. Therefore, for the earlier bootstrap samples the cross-correlation is computed between pairs of time series. The bootstrap confidence intervals for lags from -10 to 10 were derived based on their 2.5% and 97.5% quantiles. Figure 18 reveals no evidence of model inadequacy for the cross-correlation structure between the three time series. Hence, we conclude that the fitted periodic trivariate `INAR`(1) model is adequate for fitting to the monthly number of fires in Anadia, O.Bairro and Vagos counties.

(Figure 17 about here)

(Figure 18 about here)

As a final remark we would like to point out the fact that analytic expressions concerning the resulting marginal distribution for the proposed periodic model are not known, so marginal goodness-of-fit statistics cannot be computed for this model, see e.g. Weiß (2018b, p. 627) for further details.

6 Conclusions

In this paper, a class of periodic multivariate integer-valued autoregressive models of order one with period s has been introduced and explained in detail. The PMINAR(1) model can be viewed as a generalization of the bivariate model in Monteiro et al. (2015) to the multivariate case. Emphasis was placed upon models with periodic negative binomial innovations in order to account for overdispersion. To overcome the computational difficulties arising from the use of the conditional maximum likelihood method a composite likelihood-based approach is discussed. An application to a real data set related with the analysis of fire activity was presented.

It is worth mentioning here that the autoregression matrix of the PMINAR(1) process considered is diagonal, meaning no cross-correlation in the counts, which implies that the marginals behave like the univariate model in (5). Hence, extensions for PMINAR models accounting for cross-correlation is a topic for future investigation. Furthermore, it is also important to stress that beyond this extension, there are a number of open problems for future research in this area. For example, in order to make the PMINAR models more flexible with respect to real data applications, such as the one presented in this paper, it may be of interest to include covariates in the model to account for dependence through the thinning parameters on several factors. Moreover, extensions of periodic multivariate models to fit multivariate count data time series with a finite range of counts will be also very welcome.

A Assumptions of Billingsley's theorem

For a fixed v ($v = 1, \dots, s$), let the vector of parameters from the innovation process be

$$\xi_v = (\lambda_{1,v}, \lambda_{2,v}, \dots, \lambda_{m,v}, \beta_v) = (\xi_{1,v}, \xi_{2,v}, \dots, \xi_{m,v}, \xi_{m+1,v}) \in B.$$

(**C1**) The set $\{a : P(\mathbf{Z}_{v+ns} = a) = f(a, \xi_v)\}$ does not depend on ξ_v;

(**C2**) $E[\mathbf{Z}^3_{v+ns}] < \infty$;

(**C3**) $f(a, \xi_v)$ is three times continuously differentiable on the set of parameters B;

(**C4**) For any $\xi_v \in B$, there exists a neighbourhood U of ξ_v such that

$$\sum_{a=0}^{\infty} \sup_{\xi_v \in U} f(a, \xi_v) < \infty \ ,$$
$$\sum_{a=0}^{\infty} \sup_{\xi_v \in U} \left| \frac{\partial}{\partial \xi_{u,v}} f(a, \xi_v) \right| < \infty, \ u = 1, \dots, m+1,$$
$$\sum_{a=0}^{\infty} \sup_{\xi_v \in U} \left| \frac{\partial^2}{\partial \xi_{u,v} \partial \xi_{w,v}} f(a, \xi_v) \right| < \infty, \ u, w = 1, \dots, m+1;$$

(**C5**) For any $\xi_v \in B$ there exists a neighbourhood U of ξ_v and increasing sequences $\psi_u(n)$, $\psi_{u,w}(n)$, $\psi_{u,w,y}(n), n \geq 0$ such that for all $\xi_v \in B$ and all $a \leq n$ with nonvanishing $f(a, \xi_v)$

$$\left| \frac{\partial}{\partial \xi_{u,v}} f(a, \xi_v) \right| \leq \psi_u(n) \ f(a, \xi_v),$$
$$\left| \frac{\partial^2}{\partial \xi_{u,v} \partial \xi_{w,v}} f(a, \xi_v) \right| \leq \psi_{u,w}(n) \ f(a, \xi_v),$$
$$\left| \frac{\partial^3}{\partial \xi_{u,v} \partial \xi_{w,v} \partial \xi_{y,v}} f(a, \xi_v) \right| \leq \psi_{u,w,y}(n) \ f(a, \xi_v), \ u, w, y = 1, \dots, m+1;$$

and also concerning the cyclostationary distribution of $\mathbf{X}_t$, with $t = v + ns$:

$$E[\psi^3_u(\mathbf{X}_v)] < \infty, \qquad E[\mathbf{X}_v \psi_{u,w}(\mathbf{X}_{v+1})] < \infty,$$

$$E[\psi_u(\mathbf{X}_v) \psi_{u,w}(\mathbf{X}_{v+1})] < \infty, \quad E[\psi_{u,w,y}(\mathbf{X}_v)] < \infty;$$

(**C6**) The Fisher information matrix, $I(\boldsymbol{\theta})$, is nonsingular.

B Tables and Figures for Sets B and C from simulation study

(Table 3 about here)

(Figures 4-6 about here)

(Table 4 about here)

(Figures 7-9 about here)

Acknowledgements The authors wish to thank the Associate Editor and the three anonymous referees for carefully reading the article and for all helpful comments. This research was supported by the Portuguese national funding agency for science, research and technology (FCT-Fundação para a Ciência e a Tecnologia), within the Center for Research and Development in Mathematics and Applications, project UID/MAT/04106/2019 (CIDMA) and the Center for Computational and Stochastic Mathematics, Instituto Superior Técnico, Univ. of Lisbon, project UID/Multi/04621/2019 (CEMAT/IST-ID).

References

1. Al-Osh M, Alzaid A (1988) Integer-valued moving average (INMA) process. Stat Pap 29:281–300
2. Billingsley P (1961) Statistical inference for Markov processes. Statistical Research Monographs. University of Chicago Press
3. Boudreault M, Charpentier A (2011) Multivariate integer-valued autoregressive models applied to earthquake counts. arXiv:1112.0929v1 [stat.AP]
4. Bulla J, Chesneau C, Kachour M (2017) A bivariate first-order signed integer-valued autoregressive process. Commun Stat Theory Methods 46:6590–6604
5. Davis RA, Holan SH, Lund R, Ravishanker N (eds) (2016) Handbook of discrete-valued time series. CRC Press, Boca Raton
6. Davis RA, Yau CY (2011) Comments on pairwise likelihood in time series models. Stat Sin 255–277
7. Dion J, Gauthier G, Latour A (1995) Branching processes with immigration and integer-valued time series. Serdica Math J 21:123–136
8. Franke J, Subba Rao T (1993) Multivariate first-order integer-valued autoregression. Technical report, Universität Kaiserslautern
9. Harville DA (2008) Matrix algebra from a statistician's perspective. Springer-Verlag, New York
10. He Z, Wang Z, Tsung F, Shang Y (2016) A control scheme for autocorrelated bivariate binomial data. Comput Ind Eng 98:350–359
11. Joe H, Lee Y (2009) On weighting of bivariate margins in pairwise likelihood. J Multivar Anal 100:670–685
12. Jowaheer V, Mamode Khan N, Sunecher Y (2018) A BINAR(1) time series model with cross-correlated COM-Poisson innovations. Commun Stat Theory Methods 47: 1133–1154
13. Karlis D (2016) Modelling multivariate times series for counts. In: Davis RA, Holan SH, Lund R, Ravishanker N (eds) Handbook of discrete-valued time series. CRC Press, Boca Raton, pp 407-424
14. Karlis D, Pedeli X (2013) Flexible bivariate INAR(1) processes using copulas. Commun Stat Theory Methods 42:723–740
15. Kim HY, Park YS (2008) A non-stationary integer-valued autoregressive model. Stat Pap 49:485–502
16. Latour A (1997) The multivariate GINAR(p) process. Adv Appl Prob 29:228–248
17. Lindsay BG (1988) Composite likelihood methods. Contemp Math 80:221–239
18. Lindsay BG, Yi GY, Sun J (2011) Issues and strategies in the selection of composite likelihoods. Stat Sin 21:71–105
19. Mamode Khan N, Sunecher Y, Jowaheer V, Ristić MM and Heenaye-Mamode Khan M (2019) Investigating GQL-based inferential approaches for non-stationary BINAR(1) model under different quantum of over-dispersion with application. Comput Stat 34:1275–1313
20. McKenzie E (1985) Some simple models for discrete variate time series. Water Resour Bull 21:645–650
21. McKenzie E (1988) Some ARMA models for dependent sequences of Poisson counts. Adv Appl Prob 20:822–835
22. Möller TA, Silva ME, Weiß CH, Scotto MG, Pereira I (2016) Self-exciting threshold binomial autoregressive processes. AStA–Adv Stat Anal 100:369–400
23. Monteiro M, Scotto MG, Pereira I (2010) Integer-valued autoregressive processes with periodic structure. J Stat Plann Inference 140:1529–1541

24. Monteiro M, Scotto MG, Pereira I (2015) A periodic bivariate integer-valued autoregressive model. In: Bourguignon JP, Jelstch R, Pinto A, Viana M (eds) Dynamics, games and science - international conference. Advanced school planet earth DGS II. Springer, Switzerland, pp 455–477
25. Nastić AS, Ristić MM, Popović PM (2016) Estimation in a bivariate integer-valued autoregressive process. Commun Stat Theory Methods 45:5660–5678
26. Ng CT, Joe H (2014) Model comparison with composite likelihood information criteria. Bernoulli 20:1738–1764
27. Nunes AN (2012) Regional variability and driving forces behind forest fires in Portugal an overview of the last three decades (1980–2009). Appl Geog 34:576–586
28. Nunes AN, Lourenço L, Castro Meira AC (2016) Exploring spatial patterns and drivers of forest fires in Portugal (1980–2014). Sci Total Environ 573:1190–1202
29. Pedeli X, Karlis D (2011) A bivariate INAR(1) process with application. Stat Modelling 11:325–349
30. Pedeli X, Karlis D (2013a) On composite likelihood estimation of a multivariate INAR(1) model. J Time Ser Anal 34:206–220
31. Pedeli X, Karlis D (2013b) On estimation of the bivariate Poisson INAR process. Commun Stat Simulation Comput 42:514–533
32. Pedeli X, Karlis D (2013c) Some properties of multivariate INAR(1) processes. Comput Statist Data Anal 67:213–225
33. Popović PM (2015) Random coefficient bivariate INAR(1) process. Facta Univ Ser 30:263–280
34. Popović PM (2016) A bivariate INAR(1) model with different thinning parameters. Stat Pap 57:517–538
35. Popović PM, Ristić MM, Nastić AS (2016) A geometric bivariate time series with different marginal parameters. Stat Pap 57:731–753
36. Quoreshi A (2006) Bivariate time series modeling of financial count data. Commun Stat Theory Methods 35:1343–1358
37. Quoreshi A (2008) A vector integer-valued moving average model for high frequency financial count data. Econ Lett 101:258–261
38. Ristić MM, Nastić AS, Jayakumar K, Bakouch HS (2012) A bivariate INAR(1) time series model with geometric marginals. Appl Math Lett 25:481–485
39. Ristić MM, Popović BV (2019) A new bivariate binomial time series model. Markov Processes Relat Fields 25:301–328
40. Ristić MM, Sunecher Y, Mamode Khan N, Jowaheer V (2019) A GQL-based inference in non-stationary BINMA(1) time series. TEST 28:969–998
41. San Miguel-Ayanz J, Camia A (2009) Forest fires at a glance: facts, figures and trends in the EU. Living with wildfires: what science can tell us. A Contribution to the Science-Policy Dialogue, Joensuu: European Forest Institute.
42. Scotto MG, Gouveia S, Carvalho A, Monteiro A, Martins V, Flannigan M, San Miguel-Ayanz J, Miranda AI, Borrego C (2014a) Area burned in Portugal over recent decades: an extreme value analysis. Int J Wildland Fire 23:812–824
43. Scotto MG, Weiß CH, Silva ME, Pereira I (2014b) Bivariate binomial autoregressive models. J Multivar Anal 125:233–251
44. Scotto MG, Weiß CH, Gouveia S (2015) Thinning-based models in the analysis of integer-valued time series: a review. Stat Modelling 15:590–618
45. Sunecher Y, Mamode Khan N, Jowaheer V (2018) BINMA(1) model with COM-Poisson innovations: estimation and application. Commun Stat Simulation Comput DOI: 10.1080/03610918.2018.1502780
46. Sunecher Y, Mamode Khan N, Ristić M, Jowaheer V (2019) BINAR(1) negative binomial model for bivariate non-stationary time series with different over-dispersion indices. Stat Methods Appl DOI:10.1007/s10260-019-00454-0
47. Tonini M, Pereira MG, Parente J, Orozco CV (2017) Evolution of forest fires in Portugal: from spatio-temporal point events to smoothed density maps. Nat Hazards 85:1489–1510
48. Tsay RS (1992) Model checking via parametric bootstraps in time series analysis. Appl Stat 41:1–15
49. Turkman KF, Scotto MG, de Zea Bermudez P (2014) Non-linear time series: extreme events and integer value problems, Springer-Verlag, Switzerland

50. Varin C (2008) On composite marginal likelihoods. AStA–Adv Stat Anal 92:1–28
51. Varin C, Reid NM, Firth D (2011) An overview of composite likelihood methods. Stat Sin 21:5–42
52. Varin C, Vidoni P (2005) A note on composite likelihood inference and model selection. Biometrika 92:519–528
53. Weiß CH (2008) Thinning operations for modelling time series of counts a survey. AStA–Adv Stat Anal 92:319–341
54. Weiß CH (2018a) An introduction to discrete-valued time series. John Wiley & Sons, Inc., Chichester
55. Weiß CH (2018b) Goodness-of-fit testing of a count time series' marginal distribution. Metrika 81:619–651

Table 1 Parameters: $\boldsymbol{\alpha}_j = (\boldsymbol{\alpha}_{j,1}, \boldsymbol{\alpha}_{j,2}, \boldsymbol{\alpha}_{j,3}, \boldsymbol{\alpha}_{j,4})$, $\boldsymbol{\lambda}_j = (\boldsymbol{\lambda}_{j,1}, \boldsymbol{\lambda}_{j,2}, \boldsymbol{\lambda}_{j,3}, \boldsymbol{\lambda}_{j,4}), j = 1, 2, 3$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \beta_4)$.

Set A	$\boldsymbol{\alpha}_1 = (0.53, 0.75, 0.62, 0.83)$;	$\boldsymbol{\lambda}_1 = (4, 2, 3, 5)$	
	$\boldsymbol{\alpha}_2 = (0.72, 0.85, 0.56, 0.91)$;	$\boldsymbol{\lambda}_2 = (5, 3, 1.2, 2)$;	$\boldsymbol{\beta} = (1.6, 0.9, 1.8, 1.2)$
	$\boldsymbol{\alpha}_3 = (0.83, 0.60, 0.41, 0.58)$;	$\boldsymbol{\lambda}_3 = (3, 1.6, 2, 4)$	
Set B	$\boldsymbol{\alpha}_1 = (0.23, 0.45, 0.72, 0.33)$;	$\boldsymbol{\lambda}_1 = (5, 3, 4, 2)$	
	$\boldsymbol{\alpha}_2 = (0.31, 0.54, 0.26, 0.15)$;	$\boldsymbol{\lambda}_2 = (2, 4, 3, 1)$;	$\boldsymbol{\beta} = (1.2, 1.4, 1.8, 1.5)$
	$\boldsymbol{\alpha}_3 = (0.73, 0.16, 0.31, 0.82)$;	$\boldsymbol{\lambda}_3 = (3, 1.6, 2, 4)$	
Set C	$\boldsymbol{\alpha}_1 = (0.20, 0.35, 0.12, 0.43)$;	$\boldsymbol{\lambda}_1 = (5, 3, 4, 2)$	
	$\boldsymbol{\alpha}_2 = (0.30, 0.50, 0.20, 0.12)$;	$\boldsymbol{\lambda}_2 = (2, 4, 3, 1)$;	$\boldsymbol{\beta} = (2, 1.8, 3, 1)$
	$\boldsymbol{\alpha}_3 = (0.44, 0.16, 0.31, 0.22)$;	$\boldsymbol{\lambda}_3 = (3, 1.6, 2, 4)$	

Table 2 YW, CML and CL estimates for parameters in set A: $\boldsymbol{\alpha}_1 = (0.53, 0.75, 0.62, 0.83)$, $\boldsymbol{\alpha}_2 = (0.72, 0.85, 0.56, 0.91)$, $\boldsymbol{\alpha}_3 = (0.83, 0.60, 0.41, 0.58)$; $\boldsymbol{\lambda}_1 = (4, 2, 3, 5)$, $\boldsymbol{\lambda}_2 = (5, 3, 1.2, 2)$, $\boldsymbol{\lambda}_3 = (3, 1.6, 2, 4)$ and $\boldsymbol{\beta} = (1.6, 0.9, 1.8, 1.2)$. MSE in parenthesis.

	n =400			n =1000			n =2000		
	YW	CML	CL	YW	CML	CL	YW	CML	CL
$\hat{\alpha}_{1,1}$	0.521	0.531	0.531	0.528	0.531	0.531	0.528	0.529	0.529
	(0.0056)	(0.0005)	(0.0002)	(0.0021)	(0.0002)	(0.0001)	(0.0016)	(0.0001)	(0.00003)
$\hat{\alpha}_{1,2}$	0.746	0.752	0.751	0.750	0.751	0.751	0.752	0.749	0.748
	(0.0023)	(0.0006)	(0.0008)	(0.0009)	(0.0002)	(0.0007)	(0.0004)	(0.0001)	(0.00002)
$\hat{\alpha}_{1,3}$	0.608	0.618	0.617	0.617	0.621	0.620	0.615	0.620	0.619
	(0.0065)	(0.0008)	(0.0016)	(0.0029)	(0.0003)	(0.0013)	(0.0014)	(0.0001)	(0.0001)
$\hat{\alpha}_{1,4}$	0.789	0.833	0.832	0.826	0.830	0.830	0.825	0.830	0.830
	(0.0086)	(0.0007)	(0.0006)	(0.0050)	(0.0003)	(0.0002)	(0.0023)	(0.0001)	(0.0001)
$\hat{\alpha}_{2,1}$	0.717	0.718	0.717	0.739	0.719	0.719	0.740	0.720	0.720
	(0.0136)	(0.0008)	(0.0038)	(0.0059)	(0.0004)	(0.0003)	(0.0039)	(0.0001)	(0.00006)
$\hat{\alpha}_{2,2}$	0.845	0.854	0.852	0.845	0.851	0.851	0.849	0.851	0.850
	(0.0025)	(0.0005)	(0.0003)	(0.0009)	(0.0002)	(0.0005)	(0.0005)	(0.0001)	(0.00003)
$\hat{\alpha}_{2,3}$	0.552	0.559	0.560	0.559	0.559	0.559	0.561	0.560	0.560
	(0.0017)	(0.0003)	(0.0004)	(0.0007)	(0.0001)	(0.0001)	(0.0003)	(0.0001)	(0.00002)
$\hat{\alpha}_{2,4}$	0.894	0.910	0.910	0.910	0.910	0.911	0.906	0.909	0.910
	(0.0029)	(0.0002)	(0.0005)	(0.0013)	(0.0001)	(0.0003)	(0.0006)	(0.00001)	(0.00002)
$\hat{\alpha}_{3,1}$	0.823	0.832	0.832	0.832	0.831	0.831	0.830	0.830	0.830
	(0.0063)	(0.0005)	(0.0013)	(0.0028)	(0.0002)	(0.0001)	(0.0016)	(0.0001)	(0.00003)
$\hat{\alpha}_{3,2}$	0.596	0.603	0.603	0.601	0.600	0.600	0.599	0.601	0.602
	(0.0021)	(0.0006)	(0.0020)	(0.0008)	(0.0003)	(0.0002)	(0.0004)	(0.0002)	(0.0004)
$\hat{\alpha}_{3,3}$	0.391	0.411	0.410	0.411	0.409	0.409	0.407	0.411	0.411
	(0.0052)	(0.0009)	(0.0110)	(0.0024)	(0.0004)	(0.0041)	(0.0013)	(0.0002)	(0.00002)
$\hat{\alpha}_{3,4}$	0.545	0.587	0.588	0.566	0.580	0.580	0.578	0.580	0.580
	(0.0171)	(0.0020)	(0.0046)	(0.0070)	(0.0007)	(0.0026)	(0.0043)	(0.0005)	(0.0008)
$\hat{\lambda}_{1,1}$	4.182	3.909	3.927	3.986	3.981	3.988	3.995	3.999	4.033
	(1.2512)	(0.3968)	(0.1027)	(0.3894)	(0.1241)	(0.0059)	(0.3336)	(0.0777)	(0.0018)
$\hat{\lambda}_{1,2}$	2.011	2.008	2.012	2.001	1.997	2.005	1.977	2.009	2.007
	(0.2480)	(0.0915)	(0.2645)	(0.1180)	(0.0430)	(0.1500)	(0.0440)	(0.0173)	(0.0057)
$\hat{\lambda}_{1,3}$	3.146	3.024	3.027	3.078	2.980	2.984	3.063	2.978	2.986
	(0.8247)	(0.2198)	(0.0103)	(0.3777)	(0.1018)	(0.0076)	(0.1835)	(0.0487)	(0.0052)
$\hat{\lambda}_{1,4}$	5.289	5.059	5.067	5.011	5.027	5.031	5.031	4.983	4.997
	(1.1822)	(0.3725)	(0.0304)	(0.4751)	(0.1528)	(0.0176)	(0.2622)	(0.0698)	(0.0157)
$\hat{\lambda}_{2,1}$	5.243	4.924	4.954	4.965	5.010	5.022	4.965	5.005	5.042
	(1.8476)	(0.5629)	(0.2708)	(0.7502)	(0.2207)	(0.0247)	(0.4609)	(0.1228)	(0.0108)
$\hat{\lambda}_{2,2}$	3.071	3.005	3.017	3.040	2.970	2.986	3.009	3.005	3.008
	(0.4659)	(0.1812)	(0.0124)	(0.1605)	(0.0622)	(0.0460)	(0.0828)	(0.0356)	(0.0009)
$\hat{\lambda}_{2,3}$	1.324	1.220	1.215	1.221	1.196	1.201	1.203	1.188	1.184
	(0.2941)	(0.0630)	(0.0309)	(0.1310)	(0.0311)	(0.0021)	(0.0600)	(0.0135)	(0.0016)
$\hat{\lambda}_{2,4}$	2.110	2.026	2.029	1.986	1.992	1.993	2.027	1.998	2.001
	(0.3135)	(0.0825)	(0.0143)	(0.1231)	(0.0322)	(0.0089)	(0.0586)	(0.0174)	(0.0007)
$\hat{\lambda}_{3,1}$	3.061	2.929	2.942	2.950	2.986	2.994	2.989	2.992	3.017
	(0.4700)	(0.1983)	(0.0318)	(0.1782)	(0.0720)	(0.0020)	(0.1111)	(0.0438)	(0.0008)
$\hat{\lambda}_{3,2}$	1.619	1.587	1.590	1.591	1.587	1.592	1.598	1.606	1.601
	(0.1874)	(0.0663)	(0.0891)	(0.0750)	(0.0286)	(0.0764)	(0.0358)	(0.0150)	(0.0096)
$\hat{\lambda}_{3,3}$	2.150	2.004	2.008	2.027	2.006	2.003	2.036	1.987	1.989
	(0.2810)	(0.0989)	(0.0204)	(0.1419)	(0.0474)	(0.0100)	(0.0723)	(0.0238)	(0.0084)
$\hat{\lambda}_{3,4}$	4.167	4.009	4.015	4.039	3.981	3.978	4.001	3.997	4.001
	(0.6816)	(0.2625)	(0.2086)	(0.2484)	(0.1016)	(0.0882)	(0.1515)	(0.0566)	(0.0582)
$\hat{\beta}_1$	1.085	1.607	1.609	1.201	1.607	1.614	1.175	1.599	1.611
	(0.4304)	(0.1179)	(0.2239)	(0.1419)	(0.0418)	(0.0211)	(0.1045)	(0.0168)	(0.0007)
$\hat{\beta}_2$	1.481	0.915	0.902	1.529	0.903	0.903	1.554	0.895	0.897
	(0.4640)	(0.0399)	(0.1054)	(0.3124)	(0.0175)	(0.0137)	(0.2970)	(0.0098)	(0.00001)
$\hat{\beta}_3$	2.668	1.844	1.814	2.826	1.839	1.832	2.880	1.793	1.798
	(0.3708)	(0.3042)	(0.5293)	(0.1777)	(0.0955)	(0.2481)	(0.1492)	(0.0446)	(0.0031)
$\hat{\beta}_4$	1.045	1.227	1.231	1.139	1.196	1.205	1.128	1.203	1.202
	(0.1562)	(0.0516)	(0.1194)	(0.1025)	(0.0151)	(0.0051)	(0.0767)	(0.0092)	(0.0015)

Table 3 YW, CML and CL estimates for parameters in set B: $\boldsymbol{\alpha}_1 = (0.23, 0.45, 0.72, 0.33)$, $\boldsymbol{\alpha}_2 = (0.31, 0.54, 0.26, 0.15)$, $\boldsymbol{\alpha}_3 = (0.73, 0.16, 0.31, 0.82)$; $\boldsymbol{\lambda}_1 = (5, 3, 4, 2)$, $\boldsymbol{\lambda}_2 = (2, 4, 3, 1)$, $\boldsymbol{\lambda}_3 = (3, 1.6, 2, 4)$ and $\boldsymbol{\beta} = (1.2, 1.4, 1.8, 1.5)$. MSE in parenthesis.

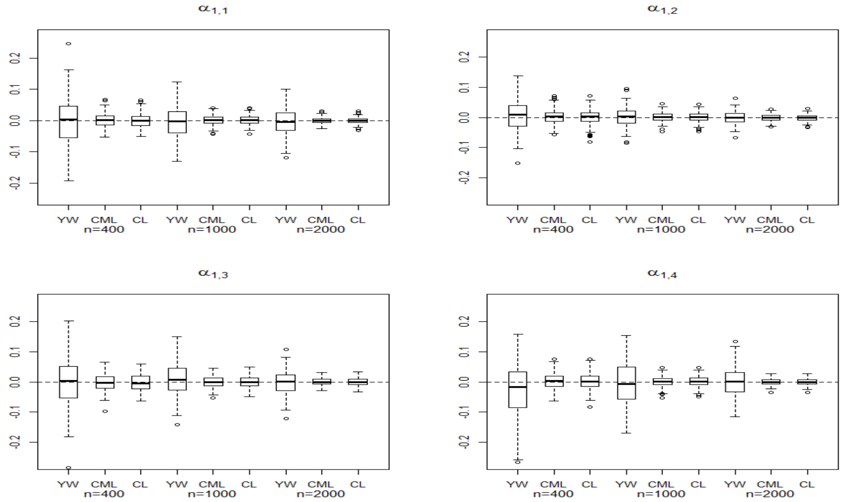
	n =400			n =1000			n =2000		
	YW	CML	CL	YW	CML	CL	YW	CML	CL
$\hat{\alpha}_{1,1}$	0.264	0.234	0.234	0.229	0.231	0.232	0.234	0.229	0.229
	(0.0227)	(0.0025)	(0.0026)	(0.0086)	(0.0010)	(0.0010)	(0.0057)	(0.0005)	(0.0006)
$\hat{\alpha}_{1,2}$	0.447	0.446	0.447	0.452	0.451	0.450	0.448	0.451	0.451
	(0.0057)	(0.0012)	(0.0013)	(0.0022)	(0.0004)	(0.0005)	(0.0012)	(0.0003)	(0.0003)
$\hat{\alpha}_{1,3}$	0.719	0.723	0.724	0.714	0.720	0.720	0.726	0.720	0.719
	(0.0131)	(0.0012)	(0.0012)	(0.0058)	(0.0004)	(0.0004)	(0.0032)	(0.0002)	(0.0002)
$\hat{\alpha}_{1,4}$	0.325	0.330	0.330	0.327	0.331	0.330	0.329	0.330	0.331
	(0.0024)	(0.0007)	(0.0007)	(0.0011)	(0.0002)	(0.0002)	(0.0005)	(0.0001)	(0.0001)
$\hat{\alpha}_{2,1}$	0.332	0.316	0.318	0.312	0.312	0.313	0.307	0.311	0.313
	(0.0235)	(0.0047)	(0.0047)	(0.0086)	(0.0019)	(0.0020)	(0.0050)	(0.0010)	(0.0010)
$\hat{\alpha}_{2,2}$	0.517	0.532	0.533	0.513	0.535	0.534	0.534	0.540	0.540
	(0.0285)	(0.0053)	(0.0055)	(0.0153)	(0.0016)	(0.0017)	(0.0074)	(0.0008)	(0.0008)
$\hat{\alpha}_{2,3}$	0.261	0.261	0.261	0.256	0.260	0.260	0.263	0.258	0.258
	(0.0065)	(0.0009)	(0.0009)	(0.0029)	(0.0004)	(0.0004)	(0.0016)	(0.0002)	(0.0002)
$\hat{\alpha}_{2,4}$	0.146	0.149	0.149	0.149	0.151	0.151	0.150	0.150	0.150
	(0.0020)	(0.0007)	(0.0008)	(0.0007)	(0.0003)	(0.0003)	(0.0004)	(0.0001)	(0.0001)
$\hat{\alpha}_{3,1}$	0.733	0.735	0.736	0.73	0.732	0.732	0.729	0.729	0.729
	(0.0061)	(0.0008)	(0.0009)	(0.0021)	(0.0003)	(0.0003)	(0.0009)	(0.0002)	(0.0002)
$\hat{\alpha}_{3,2}$	0.152	0.157	0.156	0.160	0.160	0.160	0.161	0.160	0.160
	(0.0026)	(0.0006)	(0.0006)	(0.0009)	(0.0002)	(0.0002)	(0.0004)	(0.0001)	(0.0001)
$\hat{\alpha}_{3,3}$	0.311	0.311	0.313	0.310	0.313	0.315	0.312	0.310	0.309
	(0.0151)	(0.0032)	(0.0031)	(0.0067)	(0.0008)	(0.0009)	(0.0031)	(0.0004)	(0.0004)
$\hat{\alpha}_{3,4}$	0.779	0.817	0.816	0.797	0.821	0.821	0.819	0.820	0.820
	(0.0203)	(0.0019)	(0.0019)	(0.0093)	(0.0007)	(0.0008)	(0.0058)	(0.0004)	(0.0005)
$\hat{\lambda}_{1,1}$	4.829	4.961	4.961	4.948	4.937	4.925	5.033	5.056	5.055
	(0.8452)	(0.4022)	(0.4076)	(0.3379)	(0.1728)	(0.1734)	(0.1750)	(0.0692)	(0.0685)
$\hat{\lambda}_{1,2}$	3.027	3.029	3.022	2.981	2.985	2.997	3.028	3.013	3.016
	(0.3279)	(0.1775)	(0.1805)	(0.1157)	(0.0660)	(0.0667)	(0.0660)	(0.0338)	(0.0345)
$\hat{\lambda}_{1,3}$	4.047	4.028	4.033	4.020	3.989	3.983	3.943	3.980	3.987
	(0.7396)	(0.3361)	(0.3460)	(0.2564)	(0.1248)	(0.1318)	(0.1789)	(0.0781)	(0.0764)
$\hat{\lambda}_{1,4}$	2.034	1.993	1.993	2.030	2.003	2.008	1.996	1.983	1.980
	(0.2273)	(0.1156)	(0.1171)	(0.0913)	(0.0361)	(0.0379)	(0.0487)	(0.0197)	(0.0208)
$\hat{\lambda}_{2,1}$	1.972	1.982	1.980	1.986	1.981	1.977	2.021	2.012	2.010
	(0.1087)	(0.0716)	(0.0716)	(0.0474)	(0.0312)	(0.0334)	(0.0252)	(0.0135)	(0.0133)
$\hat{\lambda}_{2,2}$	4.056	4.017	4.012	4.047	3.996	4.010	4.043	4.027	4.033
	(0.3943)	(0.2576)	(0.2574)	(0.1913)	(0.1080)	(0.1073)	(0.0892)	(0.0532)	(0.0540)
$\hat{\lambda}_{2,3}$	2.987	2.990	2.991	3.007	2.987	2.982	2.973	2.997	2.999
	(0.3245)	(0.1839)	(0.1879)	(0.1415)	(0.0770)	(0.0818)	(0.0781)	(0.0417)	(0.0407)
$\hat{\lambda}_{2,4}$	1.011	1.000	1.001	1.005	0.996	0.999	0.999	1.001	1.000
	(0.0589)	(0.0389)	(0.0401)	(0.0239)	(0.0146)	(0.0142)	(0.0117)	(0.0055)	(0.0055)
$\hat{\lambda}_{3,1}$	2.949	2.946	2.941	2.938	2.935	2.925	3.024	3.027	3.023
	(0.3563)	(0.1546)	(0.1577)	(0.1354)	(0.0664)	(0.0704)	(0.0632)	(0.0316)	(0.0312)
$\hat{\lambda}_{3,2}$	1.670	1.638	1.637	1.594	1.596	1.596	1.600	1.610	1.613
	(0.1955)	(0.0769)	(0.0764)	(0.0709)	(0.0279)	(0.0281)	(0.0295)	(0.0128)	(0.0137)
$\hat{\lambda}_{3,3}$	1.982	1.981	1.979	2.005	1.998	1.990	1.981	1.986	1.988
	(0.1993)	(0.0919)	(0.0924)	(0.0860)	(0.0403)	(0.0414)	(0.0422)	(0.0194)	(0.0196)
$\hat{\lambda}_{3,4}$	4.123	4.019	4.014	4.058	3.992	3.997	3.995	3.989	3.986
	(0.5285)	(0.3366)	(0.3426)	(0.2113)	(0.1129)	(0.1098)	(0.0943)	(0.0470)	(0.0485)
$\hat{\beta}_1$	1.310	1.217	1.216	1.229	1.200	1.208	1.201	1.188	1.186
	(0.2967)	(0.0490)	(0.0523)	(0.0801)	(0.0192)	(0.0188)	(0.0335)	(0.0136)	(0.0138)
$\hat{\beta}_2$	1.334	1.366	1.359	1.351	1.402	1.396	1.348	1.404	1.405
	(0.1501)	(0.0803)	(0.0836)	(0.0562)	(0.0327)	(0.0342)	(0.0416)	(0.0152)	(0.0151)
$\hat{\beta}_3$	1.695	1.809	1.809	1.724	1.819	1.825	1.728	1.798	1.791
	(0.3098)	(0.1462)	(0.1579)	(0.1398)	(0.0592)	(0.0586)	(0.0783)	(0.0252)	(0.0258)
$\hat{\beta}_4$	1.505	1.492	1.477	1.534	1.502	1.502	1.538	1.487	1.488
	(0.2289)	(0.0627)	(0.0625)	(0.0969)	(0.0336)	(0.0327)	(0.0594)	(0.0147)	(0.0152)

Table 4 YW, CML and CL estimates for parameters in set C: $\boldsymbol{\alpha}_1 = (0.20, 0.35, 0.12, 0.43)$, $\boldsymbol{\alpha}_2 = (0.30, 0.50, 0.20, 0.12)$, $\boldsymbol{\alpha}_3 = (0.44, 0.16, 0.31, 0.22)$; $\boldsymbol{\lambda}_1 = (5, 3, 4, 2)$, $\boldsymbol{\lambda}_2 = (2, 4, 3, 1)$, $\boldsymbol{\lambda}_3 = (3, 1.6, 2, 4)$ and $\boldsymbol{\beta} = (2, 1.8, 3, 1)$. MSE in parenthesis.

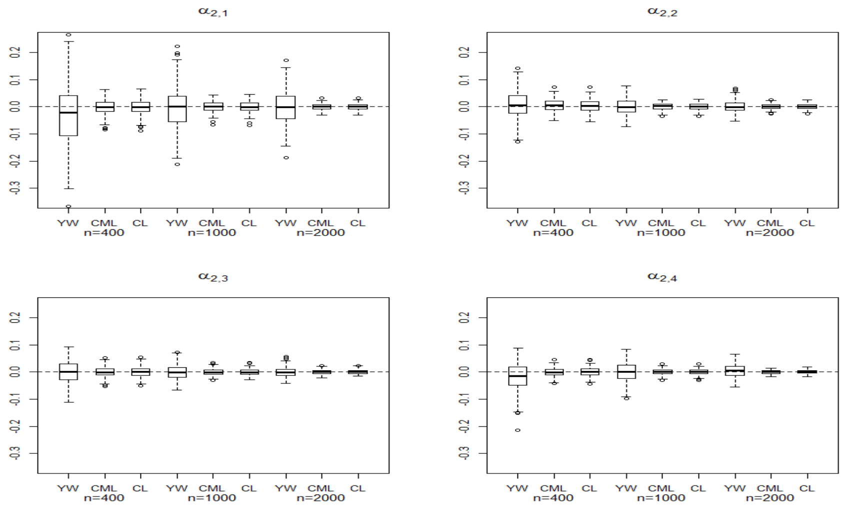
	$n=400$			$n=1000$			$n=2000$		
	YW	CML	CL	YW	CML	CL	YW	CML	CL
$\hat{\alpha}_{1,1}$	0.252	0.200	0.200	0.211	0.202	0.203	0.205	0.202	0.202
	(0.0325)	(0.0022)	(0.0025)	(0.0126)	(0.0006)	(0.0007)	(0.0079)	(0.0004)	(0.0004)
$\hat{\alpha}_{1,2}$	0.345	0.349	0.349	0.347	0.352	0.352	0.351	0.351	0.351
	(0.0047)	(0.0012)	(0.0013)	(0.0016)	(0.0004)	(0.0004)	(0.0009)	(0.0002)	(0.0002)
$\hat{\alpha}_{1,3}$	0.162	0.127	0.125	0.130	0.120	0.120	0.125	0.121	0.121
	(0.0177)	(0.0009)	(0.0009)	(0.0063)	(0.0003)	(0.0003)	(0.0035)	(0.0002)	(0.0002)
$\hat{\alpha}_{1,4}$	0.434	0.432	0.433	0.433	0.433	0.433	0.430	0.431	0.431
	(0.0025)	(0.0012)	(0.0013)	(0.0011)	(0.0006)	(0.0007)	(0.0006)	(0.0002)	(0.0003)
$\hat{\alpha}_{2,1}$	0.351	0.301	0.303	0.305	0.304	0.305	0.304	0.301	0.301
	(0.0370)	(0.0046)	(0.0050)	(0.0141)	(0.0014)	(0.0014)	(0.0088)	(0.0008)	(0.0008)
$\hat{\alpha}_{2,2}$	0.498	0.492	0.491	0.489	0.500	0.499	0.492	0.499	0.500
	(0.0303)	(0.0035)	(0.0037)	(0.0134)	(0.0016)	(0.0016)	(0.0056)	(0.0008)	(0.0009)
$\hat{\alpha}_{2,3}$	0.222	0.204	0.203	0.206	0.201	0.201	0.201	0.201	0.201
	(0.0101)	(0.0009)	(0.0009)	(0.0034)	(0.0003)	(0.0003)	(0.0017)	(0.0001)	(0.0002)
$\hat{\alpha}_{2,4}$	0.122	0.123	0.123	0.121	0.121	0.121	0.119	0.120	0.120
	(0.0015)	(0.0009)	(0.0009)	(0.0005)	(0.0003)	(0.0004)	(0.0003)	(0.0001)	(0.0001)
$\hat{\alpha}_{3,1}$	0.457	0.436	0.437	0.440	0.439	0.439	0.446	0.443	0.443
	(0.0143)	(0.0013)	(0.0014)	(0.0049)	(0.0005)	(0.0005)	(0.0025)	(0.0002)	(0.0003)
$\hat{\alpha}_{3,2}$	0.155	0.158	0.158	0.159	0.161	0.161	0.159	0.160	0.160
	(0.0039)	(0.0008)	(0.0008)	(0.0012)	(0.0003)	(0.0003)	(0.0006)	(0.0001)	(0.0001)
$\hat{\alpha}_{3,3}$	0.331	0.315	0.315	0.315	0.312	0.312	0.312	0.309	0.310
	(0.0216)	(0.0020)	(0.0021)	(0.0077)	(0.0009)	(0.0008)	(0.0040)	(0.0004)	(0.0004)
$\hat{\alpha}_{3,4}$	0.238	0.231	0.232	0.219	0.224	0.223	0.225	0.223	0.223
	(0.0155)	(0.0046)	(0.0047)	(0.0058)	(0.0020)	(0.0021)	(0.0036)	(0.0011)	(0.0010)
$\hat{\lambda}_{1,1}$	4.838	5.049	5.051	4.945	4.986	4.990	4.968	4.981	4.976
	(1.0134)	(0.6116)	(0.6205)	(0.3983)	(0.2376)	(0.2410)	(0.2398)	(0.1192)	(0.1201)
$\hat{\lambda}_{1,2}$	3.023	3.011	2.998	3.052	3.025	3.024	2.970	2.973	2.974
	(0.3318)	(0.2036)	(0.2022)	(0.1231)	(0.0852)	(0.0842)	(0.0599)	(0.0411)	(0.0410)
$\hat{\lambda}_{1,3}$	3.791	3.992	3.967	3.920	3.970	3.977	3.996	4.015	4.010
	(0.8439)	(0.5827)	(0.5914)	(0.3598)	(0.2218)	(0.2211)	(0.1739)	(0.0993)	(0.1036)
$\hat{\lambda}_{1,4}$	1.989	1.998	1.986	1.991	1.991	1.992	2.001	1.994	1.993
	(0.1049)	(0.0723)	(0.0749)	(0.0403)	(0.0289)	(0.0287)	(0.0221)	(0.0155)	(0.0163)
$\hat{\lambda}_{2,1}$	1.960	2.025	2.023	1.995	1.992	1.995	1.997	1.998	1.999
	(0.1475)	(0.1055)	(0.1039)	(0.0635)	(0.0389)	(0.0399)	(0.0377)	(0.0204)	(0.0209)
$\hat{\lambda}_{2,2}$	3.993	4.009	3.993	4.044	4.019	4.020	3.990	3.974	3.976
	(0.5393)	(0.3643)	(0.3616)	(0.1779)	(0.1297)	(0.1284)	(0.0832)	(0.0597)	(0.0600)
$\hat{\lambda}_{2,3}$	2.891	3.001	2.980	2.938	2.968	2.971	3.006	3.007	3.005
	(0.5205)	(0.3283)	(0.3301)	(0.2100)	(0.1222)	(0.1236)	(0.1010)	(0.0592)	(0.0618)
$\hat{\lambda}_{2,4}$	1.002	0.997	0.993	1.006	1.005	1.007	0.998	0.995	0.995
	(0.0353)	(0.0246)	(0.0248)	(0.0155)	(0.0113)	(0.0113)	(0.0073)	(0.0053)	(0.0052)
$\hat{\lambda}_{3,1}$	2.939	3.037	3.038	2.988	2.997	3.007	2.962	2.975	2.97
	(0.4135)	(0.2516)	(0.2516)	(0.1848)	(0.0866)	(0.0914)	(0.0925)	(0.0464)	(0.0462)
$\hat{\lambda}_{3,2}$	1.615	1.602	1.592	1.618	1.611	1.610	1.591	1.588	1.59
	(0.1523)	(0.0701)	(0.0714)	(0.0435)	(0.0249)	(0.0244)	(0.0223)	(0.0108)	(0.0108)
$\hat{\lambda}_{3,3}$	1.952	2.004	1.982	1.966	1.976	1.977	2.003	2.011	2.007
	(0.2338)	(0.1735)	(0.1738)	(0.0963)	(0.0567)	(0.0558)	(0.0443)	(0.0257)	(0.0270)
$\hat{\lambda}_{3,4}$	3.957	3.966	3.957	3.991	3.977	3.983	3.982	3.988	3.985
	(0.2641)	(0.2071)	(0.2243)	(0.1152)	(0.0821)	(0.0809)	(0.0653)	(0.0455)	(0.0455)
$\hat{\beta}_1$	2.122	1.960	1.954	2.062	1.994	2.003	2.007	2.008	2.005
	(0.6502)	(0.1164)	(0.1180)	(0.3060)	(0.0515)	(0.0531)	(0.1328)	(0.0282)	(0.0289)
$\hat{\beta}_2$	1.813	1.797	1.794	1.819	1.801	1.795	1.822	1.823	1.820
	(0.3010)	(0.1272)	(0.1389)	(0.1315)	(0.0525)	(0.0553)	(0.0565)	(0.0250)	(0.0254)
$\hat{\beta}_3$	3.271	3.100	3.079	3.029	2.996	2.991	2.947	2.979	2.983
	(1.8467)	(0.4217)	(0.4428)	(0.6518)	(0.1485)	(0.1432)	(0.2928)	(0.0605)	(0.0618)
$\hat{\beta}_4$	1.537	1.025	1.021	1.450	0.973	0.972	1.467	1.005	1.004
	(0.5490)	(0.0377)	(0.0401)	(0.2966)	(0.0171)	(0.0173)	(0.2689)	(0.0077)	(0.0078)

Table 5 CL and CML estimates from fitting the periodic trivariate INAR(1) model with trivariate negative binomial innovations. Standard errors in parenthesis.

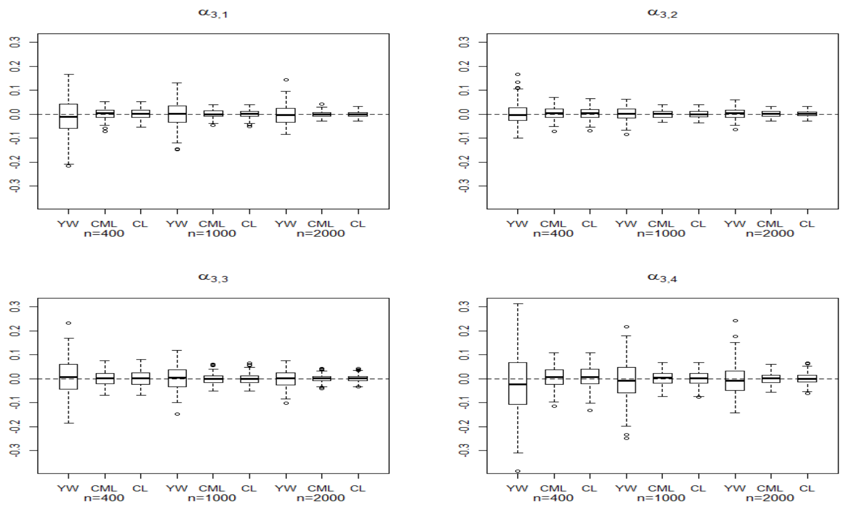
	Composite Likelihood (CL)							Conditonal Maximum Likelihood (CML)						
	Anadia		O.Bairro		Vagos			Anadia		O.Bairro		Vagos		
	$\boldsymbol{\alpha}_1$	$\boldsymbol{\lambda}_1$	$\boldsymbol{\alpha}_2$	$\boldsymbol{\lambda}_2$	$\boldsymbol{\alpha}_3$	$\boldsymbol{\lambda}_3$	$\boldsymbol{\beta}$	$\boldsymbol{\alpha}_1$	$\boldsymbol{\lambda}_1$	$\boldsymbol{\alpha}_2$	$\boldsymbol{\lambda}_2$	$\boldsymbol{\alpha}_3$	$\boldsymbol{\lambda}_3$	$\boldsymbol{\beta}$
January	0.3138	0.4284	0.0004	0.2003	0.0018	0.2331	4.0031	0.3296	0.4098	0.0003	0.1936	0.0002	0.2258	4.1623
	(0.1533)	(0.1360)	(0.0079)	(0.0738)	(0.0079)	(0.0820)	(1.4481)	(0.1321)	(0.0815)	(0.0063)	(0.0658)	(0.0006)	(0.1034)	(0.1751)
February	2.4×10^{-04}	0.8423	0.5238	0.3813	0.1598	0.3504	1.9601	0.0008	0.8125	0.5360	0.3683	0.1041	0.3522	2.0677
	(1.3×10^{-05})	(0.1799)	(0.1954)	(0.1021)	(0.1942)	(0.1007)	(0.5981)	(0.0028)	(0.2727)	(0.2689)	(0.1489)	(0.2816)	(0.1462)	(0.8986)
March	0.0137	2.5256	0.3015	0.9517	0.3617	1.0902	1.8017	0.0185	2.4692	0.0544	1.0368	0.2816	1.0818	1.9611
	(0.1358)	(0.4403)	(0.1547)	(0.1958)	(0.2165)	(0.2200)	(0.3567)	(0.3261)	(0.7041)	(0.2527)	(0.3328)	(0.3218)	(0.3427)	(0.6019)
April	0.3775	1.6122	0.1198	0.7392	0.0918	1.0211	1.4199	0.3826	1.5867	0.1242	0.7118	0.1031	0.9713	1.4330
	(0.0552)	(0.2875)	(0.0638)	(0.1563)	(0.0735)	(0.2034)	(0.4152)	(0.0767)	(0.0854)	(0.0981)	(0.4241)	(0.2226)	(0.2877)	(0.6154)
May	0.1608	2.0588	0.4898	0.6386	0.1343	1.0460	1.5132	0.1638	2.0228	0.5005	0.6090	0.0830	1.0654	1.4160
	(0.0602)	(0.3645)	(0.0930)	(0.1442)	(0.0881)	(0.2121)	(0.3824)	(0.0872)	(0.1262)	(0.1201)	(0.5312)	(0.2004)	(0.3133)	(0.5206)
June	0.1651	4.0059	0.4291	2.5060	0.0546	3.2998	0.6851	0.1565	3.8998	0.3723	2.5536	0.0164	3.2935	0.6456
	(0.0877)	(0.4894)	(0.1541)	(0.3380)	(0.1673)	(0.4270)	(0.1436)	(0.1164)	(0.2238)	(0.2262)	(0.7065)	(0.5070)	(0.6242)	(0.2072)
July	0.3829	5.5467	0.7075	4.1229	0.3899	3.5249	0.4738	0.3702	5.5399	0.6336	4.1700	0.3204	3.6262	0.3761
	(0.0844)	(0.6256)	(0.0792)	(0.4577)	(0.0788)	(0.4292)	(0.0972)	(0.1097)	(0.1127)	(0.1169)	(0.8504)	(0.6504)	(0.6287)	(0.1186)
August	0.4761	6.5477	0.3116	6.4485	0.0539	6.6102	0.5053	0.3966	7.4866	0.2949	6.5650	0.1043	6.2613	0.4062
	(0.0614)	(0.7256)	(0.0648)	(0.7007)	(0.0783)	(0.7115)	(0.0992)	(0.0911)	(0.0881)	(0.0954)	(1.1461)	(0.9976)	(0.9379)	(0.1270)
September	0.3359	3.6745	0.0725	4.2058	0.2783	2.2748	0.5909	0.3592	3.6391	0.0955	3.9529	0.2757	2.2953	0.5274
	(0.0491)	(0.6011)	(0.0496)	(0.6019)	(0.0389)	(0.3500)	(0.1443)	(0.0647)	(0.0720)	(0.0546)	(0.8382)	(0.8483)	(0.5024)	(0.1976)
October	0.0242	2.9897	0.0408	2.2900	0.0620	2.3881	2.5848	0.0287	2.9770	0.0330	2.2808	0.0851	2.2089	2.6112
	(0.0194)	(0.5716)	(0.0351)	(0.4632)	(0.0510)	(0.4961)	(0.5708)	(0.0385)	(0.9424)	(0.0448)	(0.7316)	(0.0623)	(0.7235)	(0.9358)
November	5.1×10^{-04}	0.4189	0.0116	0.6807	0.0034	0.7104	6.0760	3.6×10^{-04}	0.4063	0.0152	0.6506	0.0004	0.6875	6.1412
	(0.0003)	(0.1378)	(0.0178)	(0.2127)	(0.0025)	(0.2174)	(1.6077)	(1.4×10^{-04})	(0.2195)	(0.0247)	(0.3224)	(0.0295)	(0.3304)	(2.7802)
December	0.5025	0.1442	0.1072	0.1185	0.0187	0.0638	6.3424	0.4937	0.1434	0.1175	0.1067	0.0293	0.0625	6.9802
	(0.1053)	(0.0663)	(0.0545)	(0.0595)	(0.0388)	(0.0369)	(2.8320)	(0.1506)	(0.1002)	(0.0761)	(0.0817)	(0.1322)	(0.0651)	(2.6052)



(a)

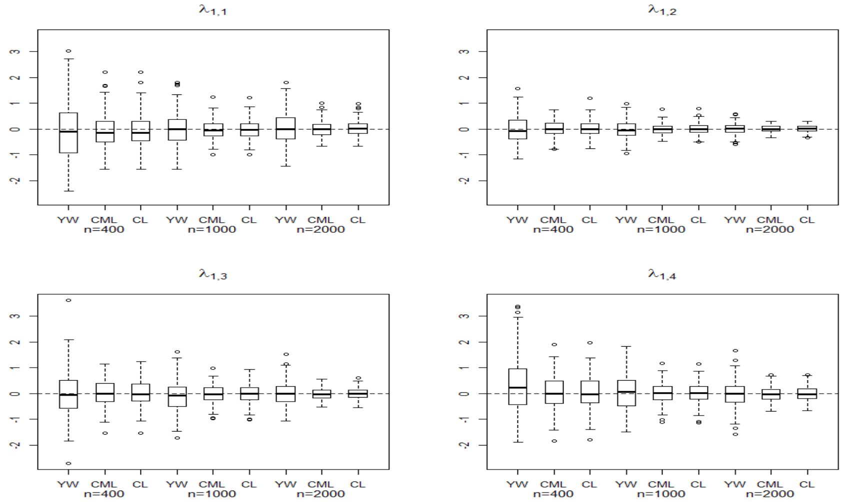


(b)

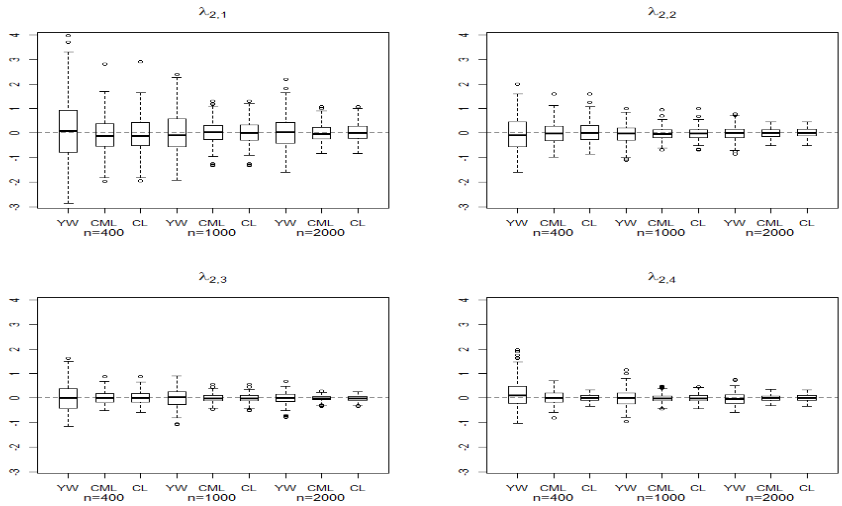


(c)

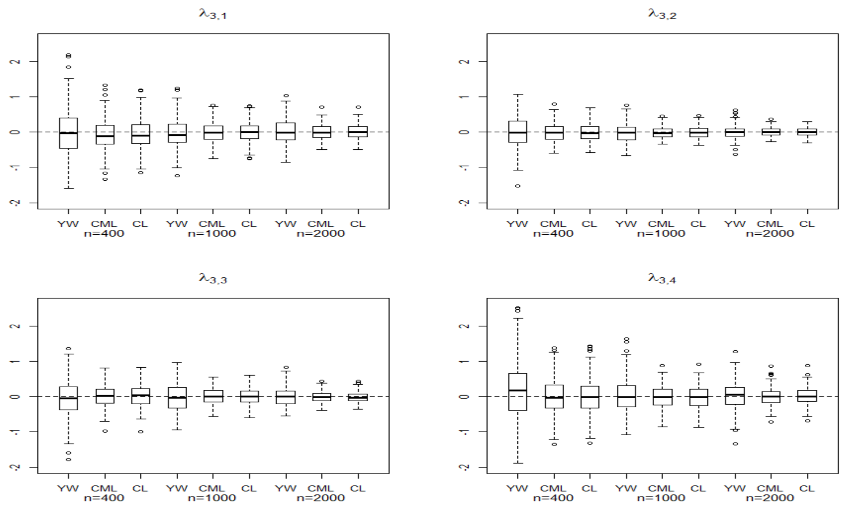
Fig. 1 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\alpha}_1$ (a), $\boldsymbol{\alpha}_2$ (b) and $\boldsymbol{\alpha}_3$ (c) in set A. From left to right, the first three boxplots display the biases of $\hat{\alpha}_{j,1}$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\alpha}_{j,2}$, $\hat{\alpha}_{j,3}$ and $\hat{\alpha}_{j,4}$ $(j = 1, 2, 3)$, respectively.



(a)



(b)



(c)

Fig. 2 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\lambda}_1$ (a), $\boldsymbol{\lambda}_2$ (b) and $\boldsymbol{\lambda}_3$ (c) in Set A. From left to right, the first three boxplots display the biases of $\hat{\lambda}_{j,1}$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\lambda}_{j,2}$, $\hat{\lambda}_{j,3}$ and $\hat{\lambda}_{j,4}$ $(j = 1, 2, 3)$, respectively.

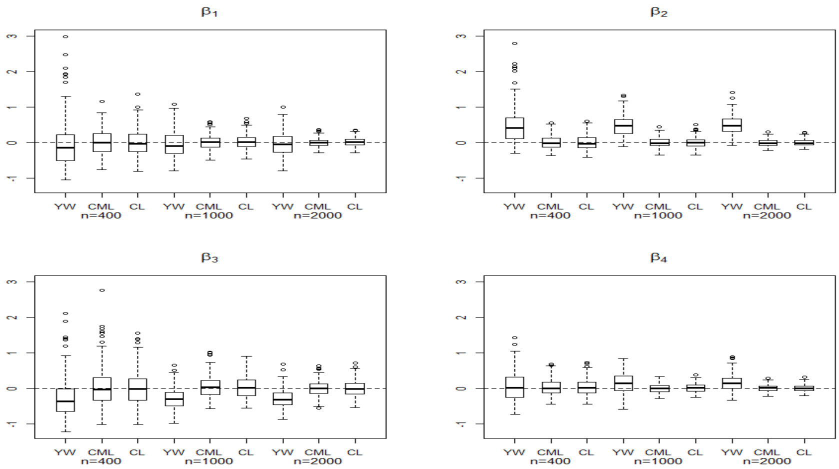
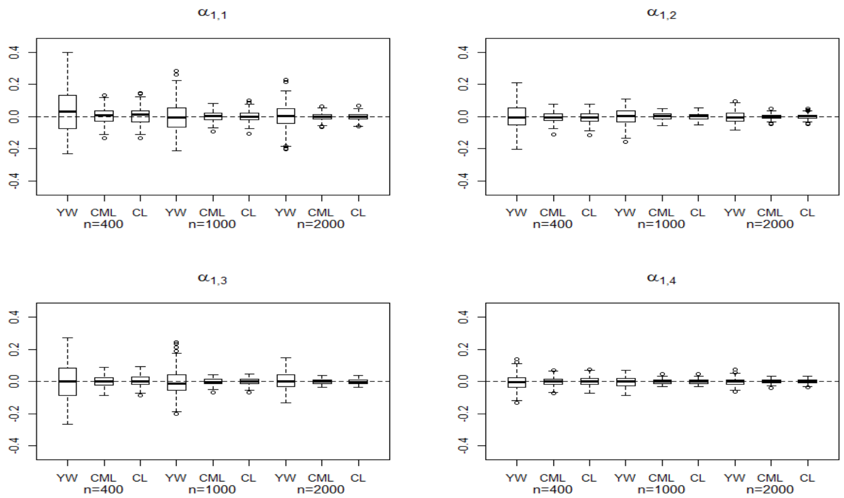
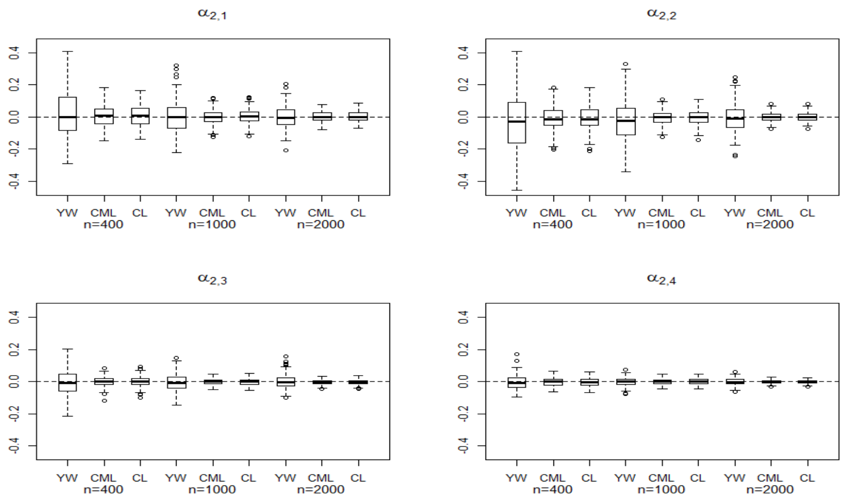


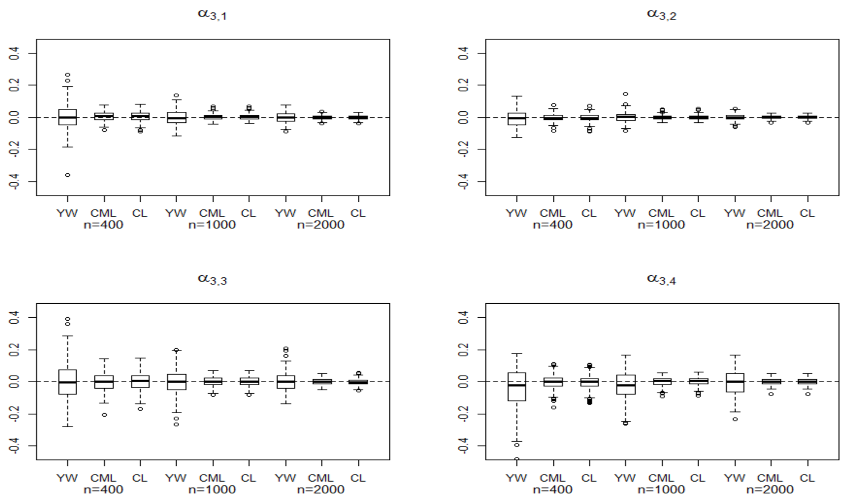
Fig. 3 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\beta}$ in set A. From left to right, the first three boxplots display the biases of $\hat{\beta}_1$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\beta}_2$, $\hat{\beta}_3$ and $\hat{\beta}_4$, respectively.



(a)

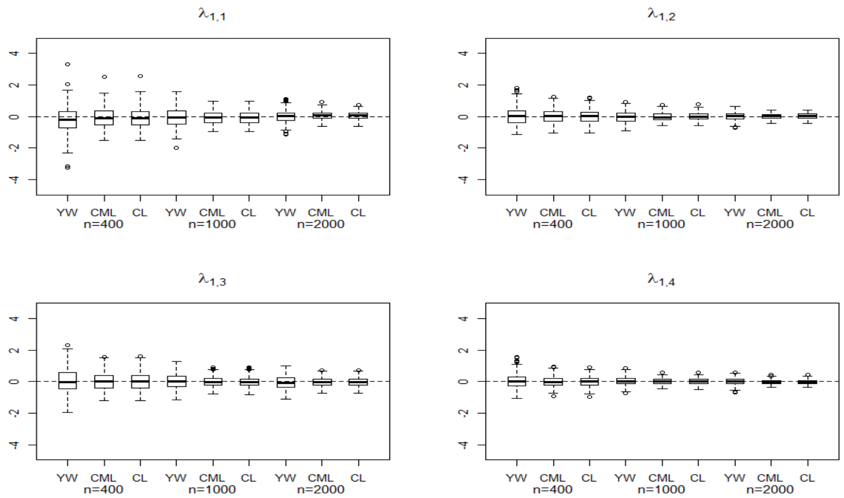


(b)

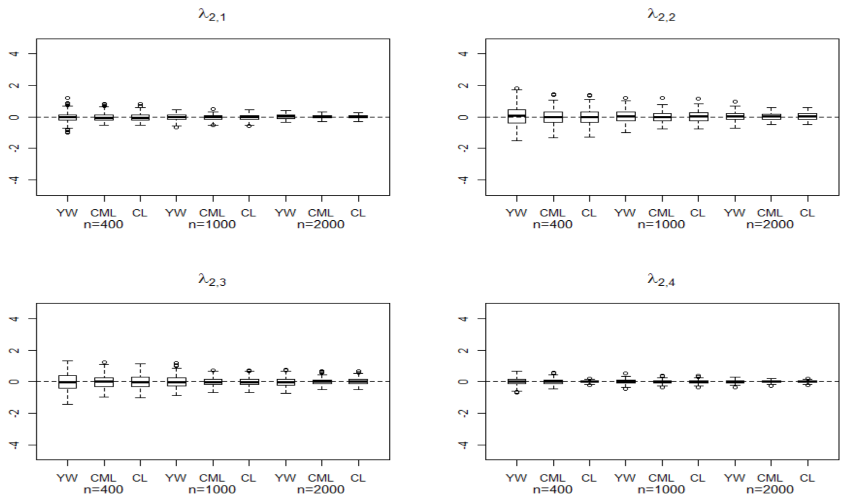


(c)

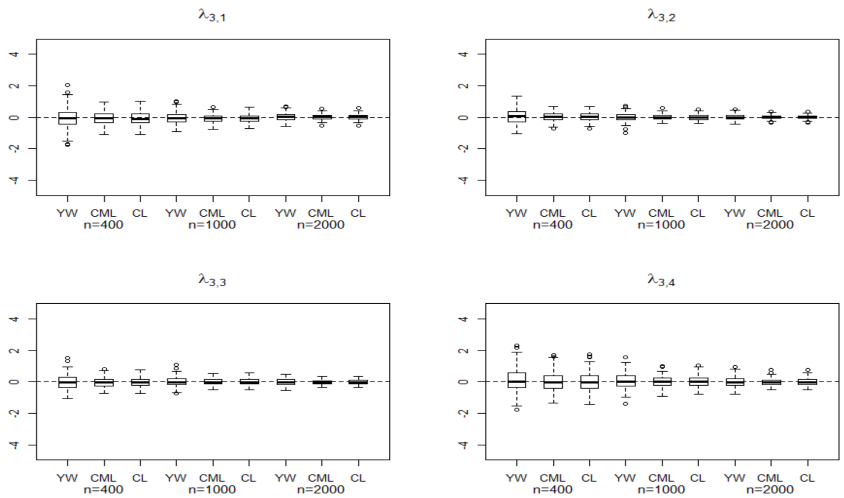
Fig. 4 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\alpha}_1$ (a), $\boldsymbol{\alpha}_2$ (b) and $\boldsymbol{\alpha}_3$ (c) in Set B. From left to right, the first three boxplots display the biases of $\hat{\alpha}_{j,1}$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\alpha}_{j,2}$, $\hat{\alpha}_{j,3}$ and $\hat{\alpha}_{j,4}$ $(j = 1, 2, 3)$, respectively.



(a)



(b)



(c)

Fig. 5 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\lambda}_1$ (a), $\boldsymbol{\lambda}_2$ (b) and $\boldsymbol{\lambda}_3$ (c) in set B. From left to right, the first three boxplots display the biases of $\hat{\lambda}_{j,1}$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\lambda}_{j,2}$, $\hat{\lambda}_{j,3}$ and $\hat{\lambda}_{j,4}$ $(j = 1, 2, 3)$, respectively.

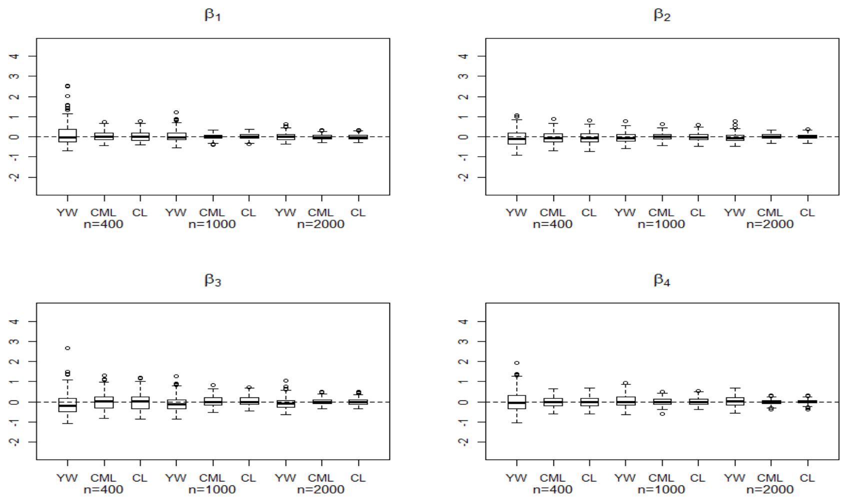
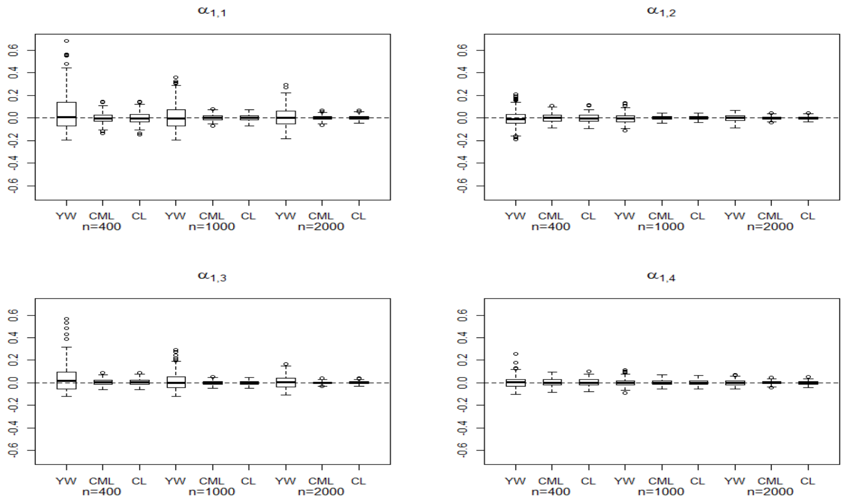
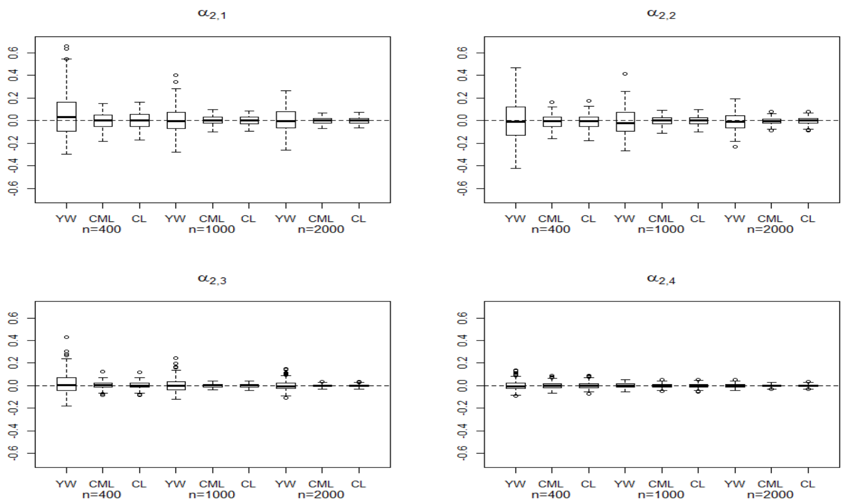


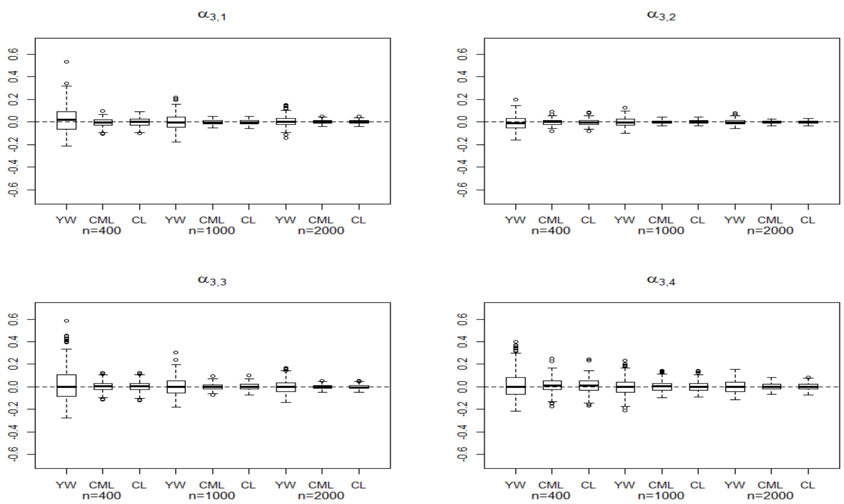
Fig. 6 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\beta}$ in set B. From left to right, the first three boxplots display the biases of $\hat{\beta}_1$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\beta}_2$, $\hat{\beta}_3$ and $\hat{\beta}_4$, respectively.



(a)

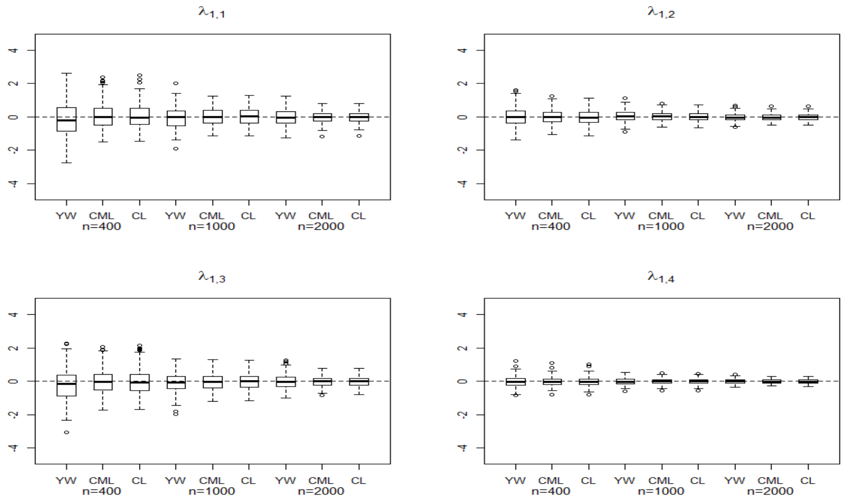


(b)

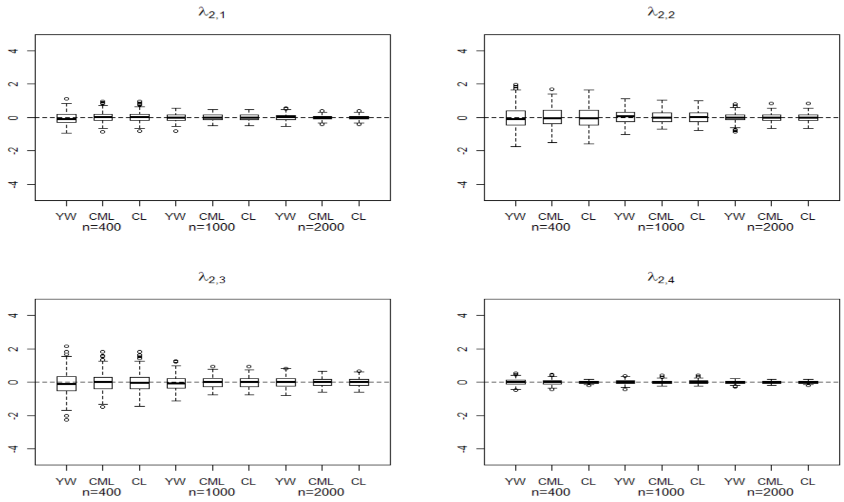


(c)

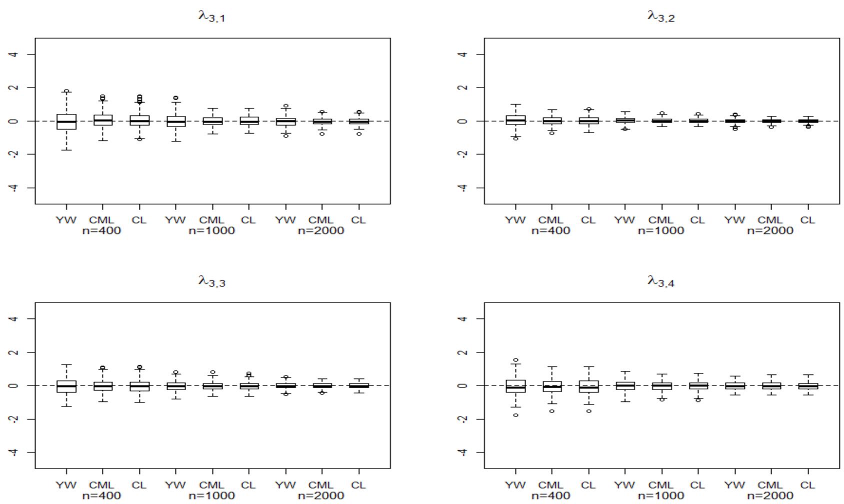
Fig. 7 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\alpha}_1$ (a), $\boldsymbol{\alpha}_2$ (b) and $\boldsymbol{\alpha}_3$ (c) in set C. From left to right, the first three boxplots display the biases of $\hat{\alpha}_{j,1}$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\alpha}_{j,2}$, $\hat{\alpha}_{j,3}$ and $\hat{\alpha}_{j,4}$ $(j = 1, 2, 3)$, respectively.



(a)



(b)



(c)

Fig. 8 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\lambda}_1$ (a), $\boldsymbol{\lambda}_2$ (b) and $\boldsymbol{\lambda}_3$ (c) in set C. From left to right, the first three boxplots display the biases of $\hat{\lambda}_{j,1}$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\lambda}_{j,2}$, $\hat{\lambda}_{j,3}$ and $\hat{\lambda}_{j,4}$ $(j = 1, 2, 3)$, respectively.

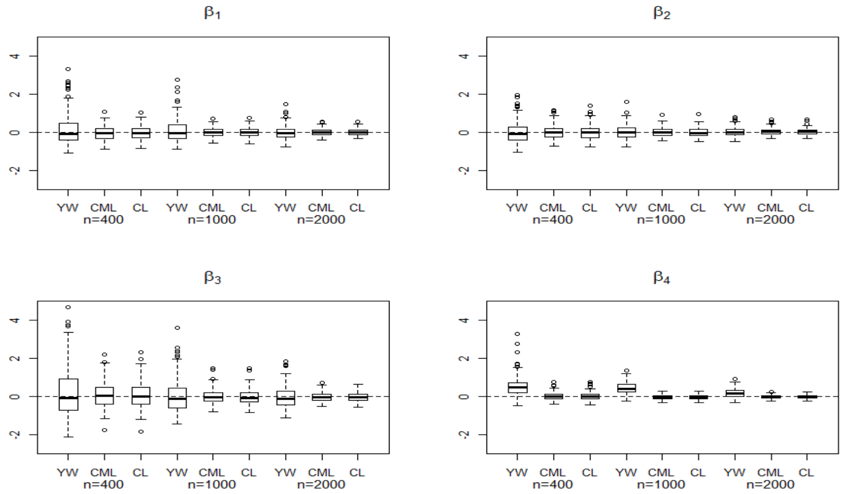


Fig. 9 Boxplots for the biases of the YW, CML and CL estimates of parameter $\boldsymbol{\beta}$ in set C. From left to right, the first three boxplots display the biases of $\hat{\beta}_1$ for the three methods with $n = 400, 1000, 2000$. The same information follows for $\hat{\beta}_2$, $\hat{\beta}_3$ and $\hat{\beta}_4$, respectively.

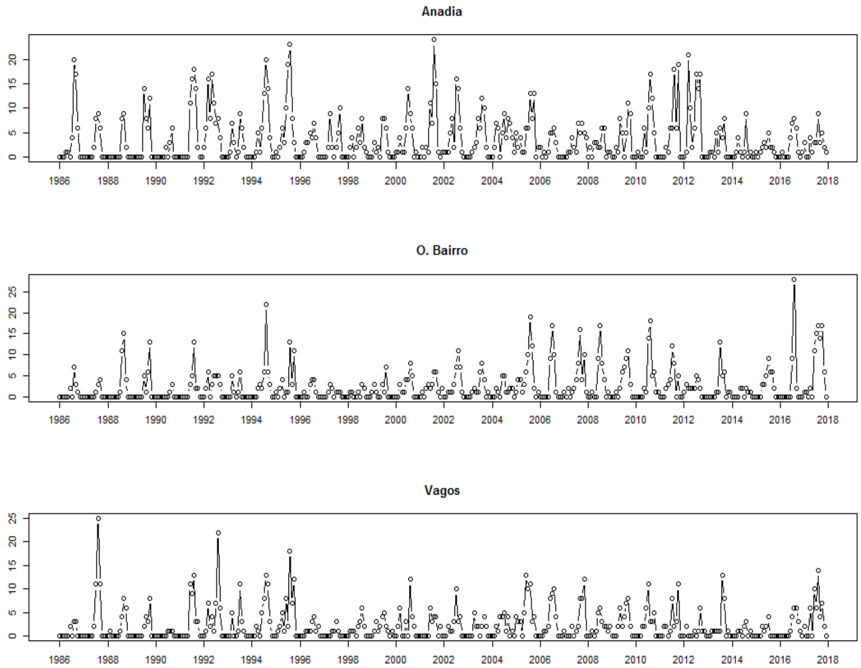


Fig. 10 Number of monthly fires in Anadia, O.Bairro and Vagos counties in Portugal.

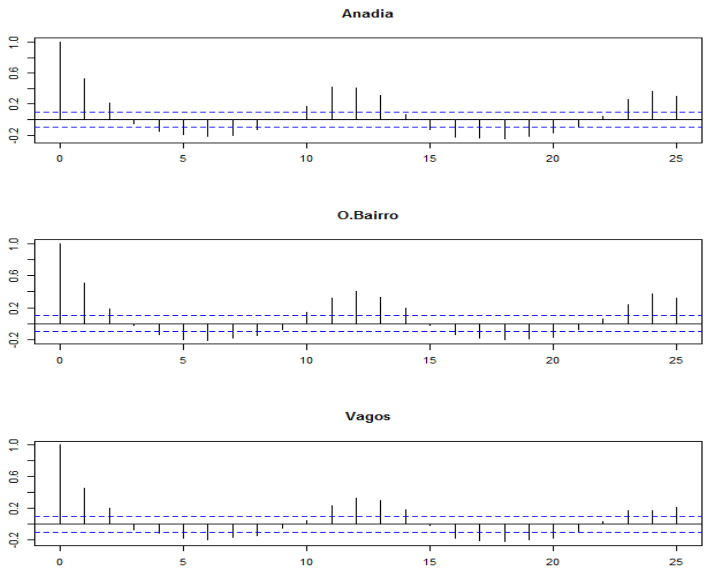


Fig. 11 Sample ACF for the number of monthly fires in Anadia, O.Bairro and Vagos.

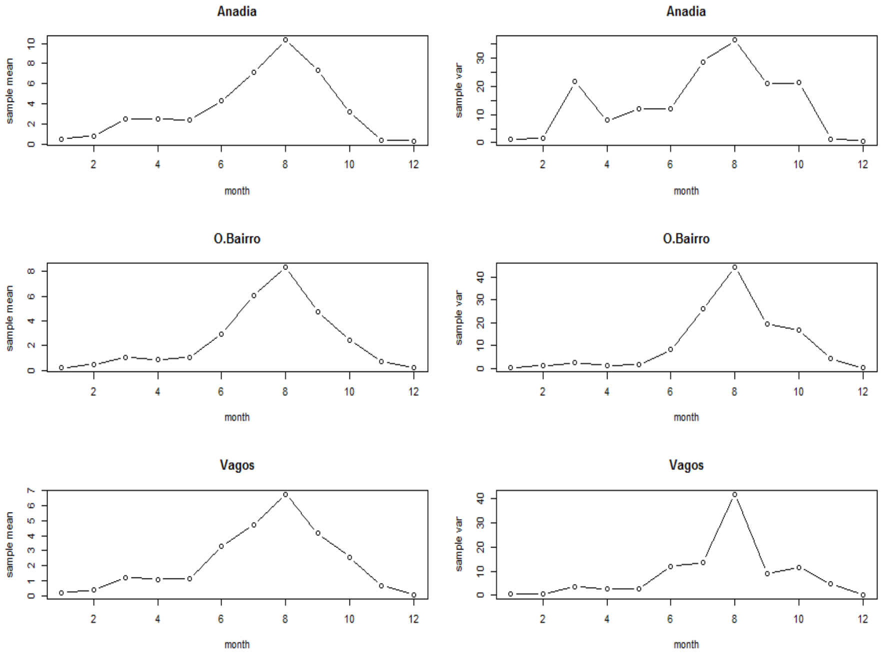


Fig. 12 Sample mean and variance for the number of monthly fires in Anadia, O.Bairro and Vagos.

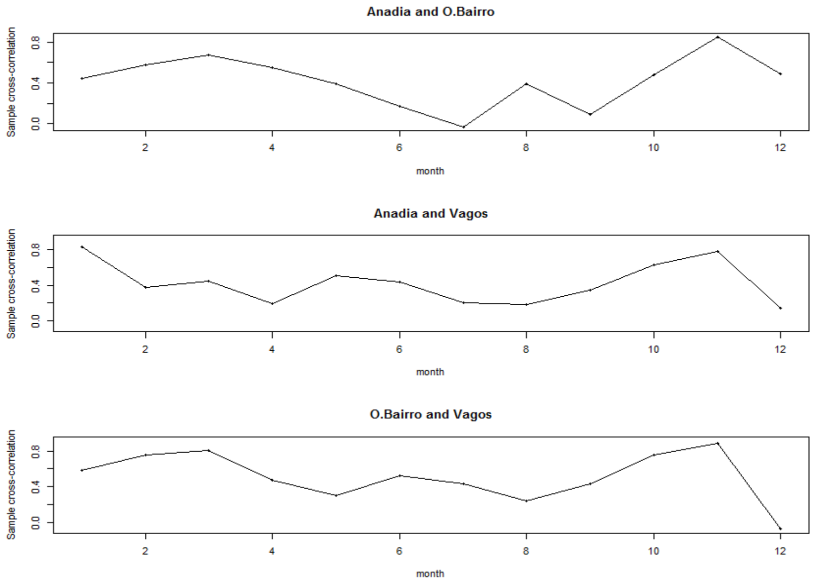


Fig. 13 Sample cross-correlations for the number of monthly fires in Anadia, O.Bairro and Vagos.

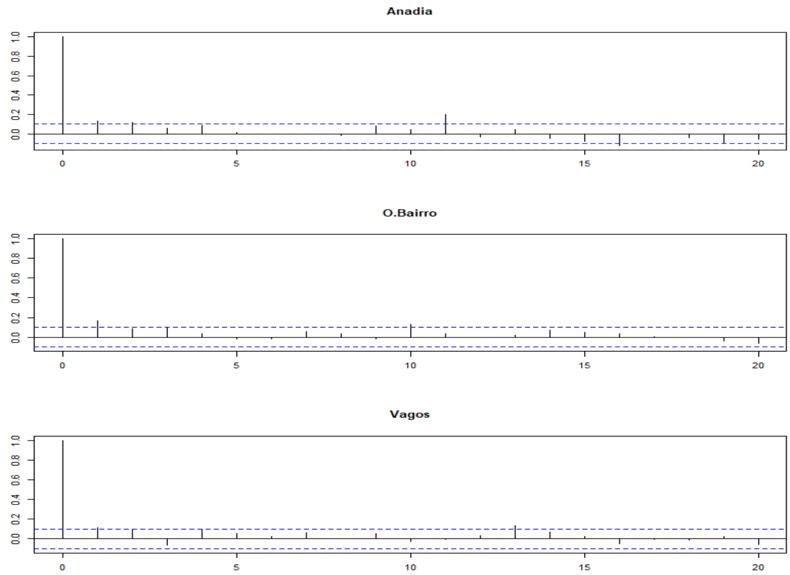


Fig. 14 ACF of Pearson residuals for Anadia, O.Bairro and Vagos counties in Portugal.

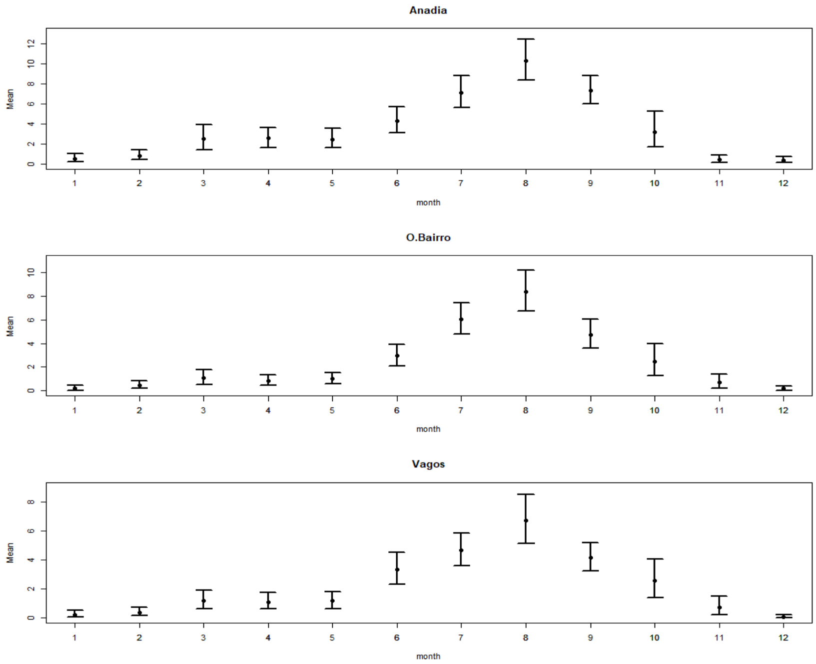


Fig. 15 Empirical mean with 95% bootstrap confidence intervals for Anadia, O.Bairro and Vagos.

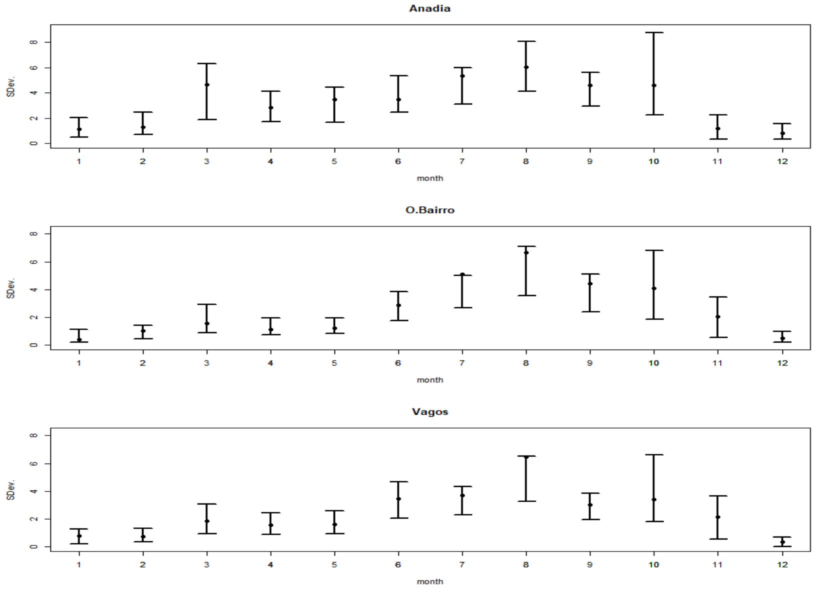


Fig. 16 Empirical standard deviation with 95% bootstrap confidence intervals for Anadia, O.Bairro and Vagos.

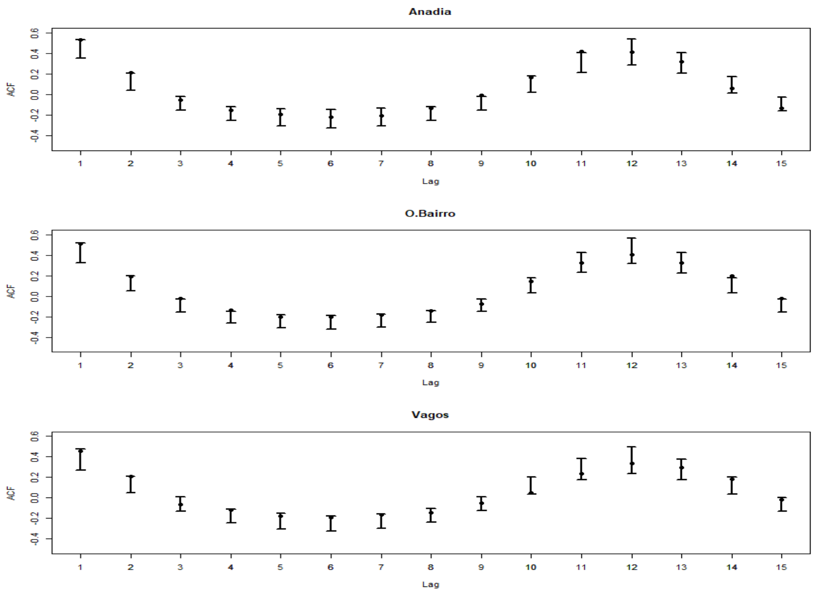


Fig. 17 Empirical ACF with 95% bootstrap confidence intervals for Anadia, O.Bairro and Vagos.

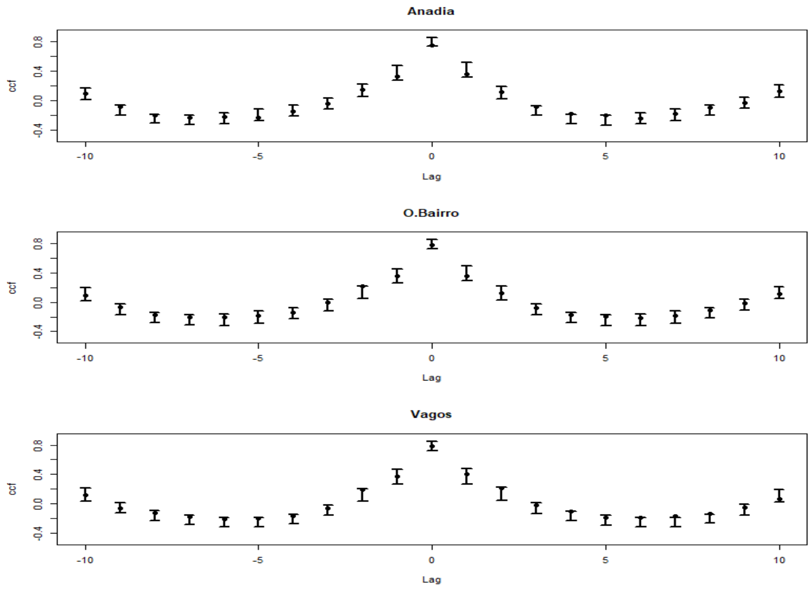


Fig. 18 Empirical cross-correlation function with 95% bootstrap confidence intervals for Anadia, O.Bairro and Vagos.