Journal of Multivariate Analysis 125 (2014) 233–251

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Bivariate binomial autoregressive models

Manuel G. Scotto [a,*], Christian H. Weiß [b], Maria Eduarda Silva [c], Isabel Pereira [a]

[a] CIDMA - Center for Research and Development in Mathematics and Applications, Department of Mathematics, University of Aveiro, Portugal

[b] Department of Mathematics and Statistics, Helmut Schmidt University Hamburg, Germany

[c] CIDMA - Center for Research and Development in Mathematics and Applications, Faculty of Economics, University of Porto, Portugal

ARTICLE INFO

Article history:
Received 31 May 2013
Available online 13 January 2014

AMS subject classifications:
60G10
62M10
60H99

Keywords:
Bivariate binomial distribution
Binomial AR(1) model
INAR(1) model
INGARCH model
Thinning operation

ABSTRACT

This paper introduces new classes of bivariate time series models being useful to fit count data time series with a finite range of counts. Motivation comes mainly from the comparison of schemes for monitoring tourism demand, stock data, production and environmental processes. All models are based on the bivariate binomial distribution of Type II. First, a new family of bivariate integer-valued GARCH models is proposed. Then, a new bivariate thinning operation is introduced and explained in detail. The new thinning operation has a number of advantages including the fact that marginally it behaves as the usual binomial thinning operation and also that allows for both positive and negative cross-correlations. Based upon this new thinning operation, a bivariate extension of the binomial autoregressive model of order one is introduced. Basic probabilistic and statistical properties of the model are discussed. Parameter estimation and forecasting are also covered. The performance of these models is illustrated through an empirical application to a set of rainy days time series collected from 2000 up to 2010 in the German cities of Bremen and Cuxhaven.

1. Introduction

Time series of (low) counts play an important role in the analysis of data sets ranging from economy and finance [20,9,13] to medicine [27,34,3,1] and biology [45]. It is worth to mention that a large part of the literature on this topic is devoted to the analysis of time series having an infinite range of counts. In particular, INteger-valued AutoRegressive-type (INAR) models based on the binomial thinning operation of Steutel and van Harn [36], defined as $\alpha \circ X := Y_1 + \cdots + Y_X$ if $X > 0$, and 0 otherwise, where the Y_i's are independent and identically distributed (i.i.d.) Bernoulli random variables with success probability $\alpha \in (0; 1)$, play a central role. For example, the INAR model of order one [26] is defined by the recursion

$$X_t = \alpha \circ X_{t-1} + \varepsilon_t \equiv \sum_{i=1}^{X_{t-1}} Y_{t,i} + \varepsilon_t, \quad t \in \mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}, \tag{1}$$

where (ε_t) is an i.i.d. process with range $\mathbb{N}_0 = \{0, 1, \ldots\}$, and where all thinning operations are performed independently of each other and of (ε_t). Furthermore, the thinning operations at each time t and ε_t are independent of $(X_s)_{s<t}$. Note that the thinning operation ensures the integer discreteness of the process. More general INAR processes of order $p > 1$ were introduced by Alzaid and Al-Osh [2].

* Correspondence to: Department of Mathematics, University of Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal.
E-mail address: mscotto@ua.pt (M.G. Scotto).

0047-259X/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmva.2013.12.014

A different approach to handle time series of counts is to consider Generalized AutoRegressive Conditional Heteroscedastic (GARCH) models, where the autoregressive structure is incorporated via a link function. A commonly used model is the INteger-valued GARCH (INGARCH) process of order (p, q) of Heinen [17], defined as

$$\begin{cases} X_t|\mathscr{F}_{t-1} : P(\lambda_t); \quad \forall t \in \mathbb{Z} \\ \lambda_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i} + \sum_{j=1}^{q} \beta_j \lambda_{t-j}, \end{cases} \tag{2}$$

where $\mathscr{F}_{t-1} := \sigma(X_s,\ s \le t-1)$, $\alpha_0 > 0$, $\alpha_i \ge 0$, and $\beta_j > 0$. Ferland et al. [12] showed that (X_t) is strictly stationary with finite first- and second-order moments provided that $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$. Weiß [39] derived the variance and autocorrelation function for the INGARCH models with $p, q > 1$. Further properties have been obtained by Zhu and Wang [47,46]. The particular case $p = q = 1$ was analyzed by Fokianos and Tjøstheim [14] and Fokianos et al. [13] under the designation of Poisson Autoregression. We refer the reader to the survey of Tjøstheim [37] and the references therein for further details.

In contrast, however, the analysis of integer-valued time series with a finite range of counts has not received much attention in the literature. The origins of the use of models based on thinning operations applied to time series with a finite range of counts, say $\{0, 1, \ldots, n\}$, can be traced back to McKenzie [26] who gave a remarkable contribution by suggesting to replace the INAR(1) recursion in (1) by

$$X_t = \alpha \circ X_{t-1} + \beta \circ (n - X_{t-1}), \quad t \in \mathbb{Z} \tag{3}$$

with $\alpha = \beta + \rho$, $\beta = \pi(1-\rho)$ for $\pi \in (0; 1)$ and $\rho \in [\max(-\pi/(1-\pi), -(1-\pi)/\pi);\ 1]$, where all thinnings are performed independently of each other, and being the thinnings at time t independent of $(X_s)_{s<t}$. Note that the representation for X_t in (3) guarantees that the range of X_t is given by $\{0, 1, \ldots, n\}$. Furthermore, the condition on ρ guarantees that $\alpha, \beta \in (0; 1)$. The process in (3) used to be referred to as *binomial* AR(1) *process* and is a stationary Markov chain with $n+1$ states and binomial marginal distribution $\mathrm{Bi}(n, \pi)$. The binomial AR(1) process shares some properties with the conventional AR(1) process, namely $\rho(k) := \rho(X_t, X_{t-k}) = \rho^k$, where $\rho(Y, Z)$ abbreviates the correlation between Y and Z. Other important features of the binomial AR(1) process are that both the conditional mean and variance of X_t given X_{t-1} are linear in X_{t-1}, and the fact that is time-reversible. For further properties see [41,43,11,40]. For binomial AR(p) processes with order $p > 1$ see [38]. Further enhancements of the basic binomial AR(1) model are proposed by Weiß and Kim [42] and Weiß and Pollett [44].

The literature on bivariate (and also multivariate) time series with finite or infinite range of counts is still in its infancy. There have been only few attempts to model bivariate/multivariate time series of counts via multivariate INGARCH models. A notable exception is the work of Heinen and Rengifo [18] who introduced the multivariate autoregressive conditional double Poisson model generalizing previous results by Heinen [17] for the univariate case. Another generalization being based on the bivariate Poisson distribution is considered by Liu [24]. Also multivariate models based upon thinning ideas have received little attention in the literature. An important contribution was made by Franke and Rao [15] who introduced the multivariate integer-valued autoregressive (MINAR, in short) model of order one based upon independent binomial thinning operations. Extensions of MINAR models with order $p > 1$ were introduced in [23] in which matrices operate on vectors using the generalized thinning operation. More recently, Pedeli and Karlis [29] introduced the bivariate INAR model of order one with Poisson and negative binomial innovations. The authors illustrated the performance of the model through an empirical application to the joint modeling of the number of daytime and nighttime road accidents in the Netherlands for the year 2001. It is important to refer that in Pedeli and Karlis' model the autoregression matrix is diagonal which means that it causes no cross-correlation in the counts; see also [30,31] for further details. The bivariate INAR model considered by Boudreault and Charpentier [7], in contrast, is the one of Franke and Rao [15], and therefore accounts for cross-correlation in the counts. An important limitation of Pedeli and Karlis' model and also Boudreault and Charpentier's model is that they only allow for positive correlations between the two time series. In order to also account for negative correlation between the time series, Karlis and Pedeli [21] introduced a family of bivariate INAR(1) processes where negative cross-correlation is introduced through the innovations, by defining the distribution of the innovations in terms of appropriate bivariate copulas. Extensions for bivariate INAR(1) models with positively correlated geometric marginals can be found in [33]. Bivariate INMA models based on the binomial thinning operation and contemporaneous only cross-correlation in the counts was proposed by Quoreshi [32] who reports an application to the number of transactions in intra-day data of stocks.

Applied to the bivariate case with $\boldsymbol{X} := [X_1\ X_2]'$, the thinning concept of Franke and Rao [15] and Boudreault and Charpentier [7] leads to the operation

$$\mathbf{A} \circ \boldsymbol{X} = \begin{bmatrix} a_{11} \circ X_1 + a_{12} \circ X_2 \\ a_{21} \circ X_1 + a_{22} \circ X_2 \end{bmatrix} \quad \text{with } \mathbf{A} \in [0; 1]^{2\times 2}, \tag{4}$$

where the thinnings are performed independently of each other. Karlis and Pedeli [21] and Pedeli and Karlis [30,31,29] restrict to the case where $a_{12} = a_{21} = 0$ such that $(\mathbf{A} \circ \boldsymbol{X})_i$ has the same distribution as $a_{ii} \circ X_i$, i.e., the marginals behave like the univariate thinning operation. However, this nice feature is obtained at the cost of no additional cross-correlation between $(\mathbf{A} \circ \boldsymbol{X})_1$ and $(\mathbf{A} \circ \boldsymbol{X})_2$, in the sense that

$$\mathrm{Cov}\big((\mathbf{A} \circ \boldsymbol{X})_1,\ (\mathbf{A} \circ \boldsymbol{X})_2\big) = \mathrm{Cov}\big(\mathrm{E}(a_{11} \circ X_1 \mid \boldsymbol{X}),\ \mathrm{E}(a_{22} \circ X_2 \mid \boldsymbol{X})\big) = a_{11} a_{22} \cdot \mathrm{Cov}(X_1, X_2). \tag{5}$$

Thus, cross-correlation is introduced in a bivariate INAR(1) model based on such a diagonal matrix thinning only through the innovations. If, in contrast, $a_{12}, a_{21} \neq 0$ is allowed as in [15,7], then the marginals of $\mathbf{A} \circ \boldsymbol{X}$ do not behave like univariate thinnings. In particular, this also implies that the marginals of a bivariate INAR(1) model being defined by using this operation do not behave like univariate INAR(1) models.

Concerning a possible bivariate extension of the binomial AR(1) model according to (3), say

$$\boldsymbol{X}_t = \mathbf{A} \circ \boldsymbol{X}_{t-1} + \mathbf{B} \circ (\boldsymbol{n} - \boldsymbol{X}_{t-1})$$

with $\boldsymbol{n} := [n_1\ n_2]' \in \mathbb{N}^2$, neither type of matrix thinning may be used also for another reason. Note that, if the matrices **A**, **B** are diagonal ones, we will not observe cross-correlation at all because no innovations are available for the model recursion; see (5). On the other hand if the matrices **A** and **B** are non-diagonal, then $(\mathbf{A}\circ\boldsymbol{X}_{t-1})_i \leq X_{t-1,i}$ and $\big(\mathbf{B}\circ(\boldsymbol{n}-\boldsymbol{X}_{t-1})\big)_i \leq n_i - X_{t-1,i}$ do not necessarily hold anymore, so it may happen that $\mathbf{A} \circ \boldsymbol{X}_{t-1} + \mathbf{B} \circ (\boldsymbol{n} - \boldsymbol{X}_{t-1}) \notin \{0, \ldots, n_1\} \times \{0, \ldots, n_2\}$, leading to a violation of the range.

For these reasons, we shall introduce a new bivariate thinning operation (referred to as *bivariate binomial* thinning operation) in Section 4, which has, among others, the following properties: (a) marginally, it behaves like the usual binomial thinning operation, but it induces additional cross-correlation compared to (5); and (b) this cross-correlation might be both positive and negative. This new thinning operation is based on the Bivariate Binomial distribution of Type II (BVB_{II}, in short), which is briefly surveyed in Section 2. In Section 3, a new INARCH(1) model based on the BVB_{II} is introduced. In Section 5, a new class of bivariate binomial AR(1) models (BVB_{II}-AR) based on the bivariate binomial thinning operation is introduced and studied in some detail. Properties concerning transition probabilities and cross-correlation are discussed. Section 6 deals with parameter estimation for both BVB_{II}-INARCH and BVB_{II}-AR processes. Forecasting is covered in Section 7. An empirical application to rainy days time series collected from 2000 up to 2010 in the German cities of Bremen and Cuxhaven is presented in Section 8. Finally, conclusions and likely directions of future work are discussed in Section 9.

2. Bivariate binomial distributions

In this section, we briefly discuss some notation and background results about bivariate Bernoulli and Type II binomial distributions. Further details can be found in the surveys by Kocherlakota and Kocherlakota [22] and Johnson et al. [19].

2.1. Bivariate Bernoulli distribution

Let $\boldsymbol{Y} := [Y_1\ Y_2]'$ be a random variable with bivariate Bernoulli distribution taking the four possible outcomes $(1, 1)$, $(1, 0)$, $(0, 1)$, $(0, 0)$ with probabilities $p_{11}, p_{10}, p_{01}, p_{00}$, respectively. These probabilities are determined by the parameters $0 < \alpha_1, \alpha_2 < 1$ and $0 < \alpha < \min(\alpha_1, \alpha_2)$ by setting

$$p_{11} = \alpha, \qquad p_{11} + p_{10} = \alpha_1, \qquad p_{11} + p_{01} = \alpha_2.$$

The marginals Y_1 and Y_2 are univariately Bernoulli-distributed with success probabilities α_1 and α_2, respectively. Furthermore, it is also well known that the correlation between Y_1 and Y_2 is given by

$$\rho(Y_1, Y_2) = \frac{\alpha - \alpha_1\alpha_2}{\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}} =: \phi_\alpha. \tag{6}$$

An alternative parametrization is obtained by replacing α by the "correlation parameter" ϕ_α according to (6). The range of ϕ_α is restricted to

$$\max\left\{-\sqrt{\frac{\alpha_1\alpha_2}{(1-\alpha_1)(1-\alpha_2)}}, -\sqrt{\frac{(1-\alpha_1)(1-\alpha_2)}{\alpha_1\alpha_2}}\right\} < \phi_\alpha < \min\left\{\sqrt{\frac{\alpha_1(1-\alpha_2)}{(1-\alpha_1)\alpha_2}}, \sqrt{\frac{(1-\alpha_1)\alpha_2}{\alpha_1(1-\alpha_2)}}\right\}. \tag{7}$$

To prove (7) note first that any of the probabilities p_{ij} with $i, j \in \{0, 1\}$ has to satisfy $0 < p_{ij} < \min(p_{i\bullet}, p_{\bullet j})$. In addition, note that

$$\alpha = \alpha_1\alpha_2 + \phi_\alpha \cdot \sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}.$$

From $\alpha > 0$, it follows that $\phi_\alpha > -\sqrt{\frac{\alpha_1\alpha_2}{(1-\alpha_1)(1-\alpha_2)}}$ has to hold, while $\alpha < \min(\alpha_1, \alpha_2)$ implies

$$\phi_\alpha < \min\left\{\sqrt{\frac{\alpha_1(1-\alpha_2)}{(1-\alpha_1)\alpha_2}}, \sqrt{\frac{(1-\alpha_1)\alpha_2}{\alpha_1(1-\alpha_2)}}\right\}.$$

From $p_{10} = \alpha_1 - p_{11} < 1 - \alpha_2 = p_{\bullet 0}$, it also follows that

$$\alpha_1(1-\alpha_2) - \phi_\alpha \cdot \sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)} < 1 - \alpha_2$$

and hence $\phi_\alpha > -\sqrt{\frac{(1-\alpha_1)(1-\alpha_2)}{\alpha_1\alpha_2}}$.

The probability generating function (pgf) equals

$$G_{\boldsymbol{Y}}(z_1,z_2) \equiv E(z_1^{Y_1}z_2^{Y_2}) = p_{00}+p_{10}z_1+p_{01}z_2+p_{11}z_1z_2$$
$$= (1-\alpha_1+\alpha_1 z_1)(1-\alpha_2+\alpha_2 z_2)+\phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}\cdot(1-z_1)(1-z_2), \quad (8)$$

so $\phi_\alpha = 0$ implies that Y_1 and Y_2 are independent Bernoulli random variables.

If $\boldsymbol{Y}_1,\ldots,\boldsymbol{Y}_k$ are $k \geq 1$ i.i.d. bivariately Bernoulli-distributed random variables, then the sample sum $\boldsymbol{W} := \boldsymbol{Y}_1+\cdots+\boldsymbol{Y}_k$ is said to follow a *bivariate binomial distribution of Type* I, abbreviated as $\mathrm{BVB_I}(k;\alpha_1,\alpha_2,\phi_\alpha)$ or $\mathrm{BVB_I}(k;\alpha_1,\alpha_2,\alpha)$, respectively. Its marginals are univariately binomially distributed, $W_1 \sim \mathrm{Bi}(k,\alpha_1)$ and $W_2 \sim \mathrm{Bi}(k,\alpha_2)$. For further details, see [22,25].

2.2. Bivariate binomial distributions of Type II

To obtain a bivariate distribution with marginals $\mathrm{Bi}(n_1,\alpha_1)$ and $\mathrm{Bi}(n_2,\alpha_2)$, where both $n_1 \neq n_2$ and $\alpha_1 \neq \alpha_2$ may happen, the following construction can be done [22,25].

Definition 2.1. Let $n_1, n_2 > 0$ and $0 \leq k < \min(n_1,n_2)$, and $\alpha_1,\alpha_2,\alpha,\phi_\alpha$ as defined in Section 2.1. Let $\boldsymbol{W}, U, V$ be independent random variables, where $\boldsymbol{W} \sim \mathrm{BVB_I}(k;\alpha_1,\alpha_2,\phi_\alpha)$, $U \sim \mathrm{Bi}(n_1-k,\alpha_1)$ and $V \sim \mathrm{Bi}(n_2-k,\alpha_2)$. Then

$$\boldsymbol{X} := [W_1+U\ W_2+V]', \quad (9)$$

is said to follow a bivariate binomial distribution of Type II, abbreviated as $\mathrm{BVB_{II}}(n_1,n_2,k;\alpha_1,\alpha_2,\alpha)$ or $\mathrm{BVB_{II}}(n_1,n_2,k;\alpha_1,\alpha_2,\phi_\alpha)$, respectively.

Let $\boldsymbol{X} := [X_1\ X_2]'$ be such a bivariate random variable, then

$$X_1 \sim \mathrm{Bi}(n_1,\alpha_1) \quad \text{and} \quad X_2 \sim \mathrm{Bi}(n_2,\alpha_2), \quad (10)$$

$$\mathrm{Cov}(X_1,X_2) = k(\alpha-\alpha_1\alpha_2) = k\phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}, \quad (11)$$

$$\rho(X_1,X_2) = \frac{k}{\sqrt{n_1n_2}}\cdot\frac{\alpha-\alpha_1\alpha_2}{\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}} = \frac{k}{\sqrt{n_1n_2}}\cdot\phi_\alpha \quad (12)$$

and

$$p_{(n_1,n_2,k;\alpha_1,\alpha_2,\alpha)}(x_1,x_2) := \mathrm{P}(X_1=x_1,X_2=x_2)$$
$$= \sum_{j_1=0}^{\min(x_1,n_1-k)}\sum_{j_2=0}^{\min(x_2,n_2-k)}\binom{n_1-k}{j_1}\binom{n_2-k}{j_2}\alpha_1^{j_1}(1-\alpha_1)^{n_1-k-j_1}\alpha_2^{j_2}(1-\alpha_2)^{n_2-k-j_2}$$
$$\cdot\sum_{i=\max(0,x_1-j_1+x_2-j_2-k)}^{\min(x_1-j_1,x_2-j_2)}\binom{k}{i,\ x_1-j_1-i,\ x_2-j_2-i,\ k+i+j_1+j_2-x_1-x_2}$$
$$\times\alpha^i(\alpha_1-\alpha)^{x_1-j_1-i}(\alpha_2-\alpha)^{x_2-j_2-i}(1+\alpha-\alpha_1-\alpha_2)^{k+i+j_1+j_2-x_1-x_2}. \quad (13)$$

It is worth to mention that (13) corresponds to (3.4.2) in [22], or to (3.4) in [25].

If the parameter ϕ_α according to (6) is used instead of α, then formula (13) is modified by considering that

$$\begin{aligned}
\alpha &= \alpha_1\alpha_2+\phi_\alpha\cdot\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)},\\
\alpha_1-\alpha &= \alpha_1(1-\alpha_2)-\phi_\alpha\cdot\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)},\\
\alpha_2-\alpha &= (1-\alpha_1)\alpha_2-\phi_\alpha\cdot\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)},\\
1+\alpha-\alpha_1-\alpha_2 &= (1-\alpha_1)(1-\alpha_2)+\phi_\alpha\cdot\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}.
\end{aligned} \quad (14)$$

The pgf satisfies (see the result in (8))

$$G_{\boldsymbol{X}}(z_1,z_2) = (1-\alpha_1+\alpha_1z_1)^{n_1-k}(1-\alpha_2+\alpha_2z_2)^{n_2-k}\cdot G_{\boldsymbol{Y}}^k(z_1,z_2). \quad (15)$$

This shows that $k=0$ or $\phi_\alpha=0$ implies that X_1 and X_2 are independent binomial random variables.

Remark 2.2. We extend the above definition to $n_1=0$ or $n_2=0$ (both also implying that $k=0$) as follows:

- $n_1=0, n_2>0$: $\boldsymbol{X}=[0\ X_2]'$ with $X_2 \sim \mathrm{Bi}(n_2,\alpha_2)$;
- $n_1>0, n_2=0$: $\boldsymbol{X}=[X_1\ 0]'$ with $X_1 \sim \mathrm{Bi}(n_1,\alpha_1)$;
- $n_1=n_2=0$: $\boldsymbol{X}=[0\ 0]'$.

Henceforth, we usually set $k=\min(n_1,n_2)$ as also done by Biswas and Hwang [5]. In this case, we denote the bivariate probability mass function (pmf) according to formula (13) by $p_{(n_1,n_2;\alpha_1,\alpha_2,\phi_\alpha)}(x_1,x_2)$.

3. The $\mathrm{BVB_{II}}$-INARCH(1) model

The $\mathrm{BVB_{II}}$-distribution as discussed in Section 2.2 can be used to adapt the INGARCH approach to bivariate processes of counts with a finite range. Here, we consider a first-order autoregressive model, which we shall refer to as the $\mathrm{BVB_{II}}$-INARCH(1) model.

Definition 3.1. Let $\boldsymbol{n} := [n_1\, n_2]' \in \mathbb{N}^2$ denote the vector of upper limits for the bivariate range. A $\mathrm{BVB_{II}}$-INARCH(1) process is a bivariate process $(\boldsymbol{X}_t)$ of bivariate random variables $\boldsymbol{X}_t := [X_{t,1}\, X_{t,2}]'$ satisfying the following recursion

$$\boldsymbol{X}_t|\mathscr{F}_{t-1} \sim \mathrm{BVB_{II}}\left(n_1, n_2, k; \alpha_{0,1} + \alpha_{1,1}\frac{X_{t-1,1}}{n_1}, \alpha_{0,2} + \alpha_{1,2}\frac{X_{t-1,2}}{n_2}, \phi\right), \quad t \in \mathbb{Z}, \tag{16}$$

where $\mathscr{F}_{t-1} := \sigma(\boldsymbol{X}_s,\ s \le t-1)$, $k := \min(n_1, n_2)$, $|\phi| < 1$ and $\alpha_{0,1}, \alpha_{0,2}, \alpha_{0,1} + \alpha_{1,1}, \alpha_{0,2} + \alpha_{1,2} \in (0; 1)$.

Note that relation (10) implies that the marginals of the $\mathrm{BVB_{II}}$-INARCH(1) process follow the univariate binomial INARCH(1) model, which was briefly considered in Section 4.3 of Weiß and Pollett [44].

The transition probabilities of the $\mathrm{BVB_{II}}$-INARCH(1) model according to Definition 3.1 are computed by using the bivariate pmf in (13) of the $\mathrm{BVB_{II}}$-distribution as introduced in Section 2, that is

$$p(\boldsymbol{x}|\boldsymbol{y}) := \mathrm{P}\Big(\boldsymbol{X}_t = \boldsymbol{x} \;\Big|\; \boldsymbol{X}_{t-1} = \boldsymbol{y}\Big) = p_{(n_1,n_2;\alpha_{0,1}+\alpha_{1,1}y_1/n_1,\alpha_{0,2}+\alpha_{1,2}y_2/n_2,\phi)}(x_1, x_2). \tag{17}$$

Since these transition probabilities at lag 1 are truly positive, the $\mathrm{BVB_{II}}$-INARCH(1) process is a primitive and finite-state Markov chain, which, in turn, implies irreducibility and aperiodicity [35, Section 4.2]. Hence, a uniquely determined stationary marginal distribution exists. As an ergodic finite-state Markov chain, the $\mathrm{BVB_{II}}$-INARCH(1) process is also ψ- and φ-mixing; see [6]. If $\boldsymbol{Q} := \big(p(\boldsymbol{x}|\boldsymbol{y})\big)$ abbreviates the transition matrix (where the bivariate states are ordered in a certain way), then the unique stationary marginal distribution, expressed as a vector $\boldsymbol{p}$, is obtained as the solution of the linear equation $\boldsymbol{Q}\boldsymbol{p} = \boldsymbol{p}$.

The parameter ϕ of the $\mathrm{BVB_{II}}$-INARCH(1) model according to Definition 3.1 takes values in the interval $[-1; 1]$, although it has to satisfy the restrictions imposed by relation (7) as well. In fact, this relation has to be satisfied for any of the conditional distributions $\mathrm{BVB_{II}}\big(n_1, n_2, k; \alpha_{0,1} + \alpha_{1,1}y_1/n_1, \alpha_{0,2} + \alpha_{1,2}y_2/n_2, \phi\big)$ with $y_1 \in \{0, \dots, n_1\}$ and $y_2 \in \{0, \dots, n_2\}$. Hence we obtain

$$\max\left\{-\sqrt{\frac{\alpha_{0,1}\alpha_{0,2}}{(1-\alpha_{0,1})(1-\alpha_{0,2})}},\ -\sqrt{\frac{(1-\alpha_{0,1}-\alpha_{1,1})(1-\alpha_{0,2}-\alpha_{1,2})}{(\alpha_{0,1}+\alpha_{1,1})(\alpha_{0,2}+\alpha_{1,2})}}\right\}$$
$$< \phi < \min\left\{\sqrt{\frac{\alpha_{0,1}(1-\alpha_{0,2}-\alpha_{1,2})}{(1-\alpha_{0,1})(\alpha_{0,2}+\alpha_{1,2})}}, \sqrt{\frac{(1-\alpha_{0,1}-\alpha_{1,1})\alpha_{0,2}}{(\alpha_{0,1}+\alpha_{1,1})(1-\alpha_{0,2})}}\right\}. \tag{18}$$

Next we investigate the first- and second-order moments of the $\mathrm{BVB_{II}}$-INARCH(1) model. Since the conditional distribution of the marginals is a binomial one according to relations (10) and (16), i.e., $X_{t,i}|\mathscr{F}_{t-1} \sim \mathrm{Bi}\big(n_i,\ \alpha_{0,i}+\alpha_{1,i}X_{t-1,i}/n_i\big)$ for $i = 1, 2$, it immediately follows that

$$\begin{aligned} \mathrm{E}[X_{t,i}|\mathscr{F}_{t-1}] &= n_i\,\alpha_{0,i} + \alpha_{1,i}X_{t-1,i}, \\ \mathrm{V}[X_{t,i}|\mathscr{F}_{t-1}] &= n_i\alpha_{0,i}(1-\alpha_{0,i}) + (1-2\alpha_{0,i})\alpha_{1,i}X_{t-1,i} - \alpha_{1,i}^2X_{t-1,i}^2/n_i. \end{aligned} \tag{19}$$

Hence, the following properties of the unconditional moments of the uniquely determined stationary $\mathrm{BVB_{II}}$-INARCH(1) model can be established.

Theorem 3.2. *The stationary* $\mathrm{BVB_{II}}$-*INARCH*(1) *model has mean value*

$$\mathrm{E}[X_{t,i}] = \frac{n_i\,\alpha_{0,i}}{(1-\alpha_{1,i})} \tag{20}$$

and variance

$$\mathrm{V}[X_{t,i}] = \frac{n_i\alpha_{0,i}(1-\alpha_{0,i}-\alpha_{1,i})}{(1-\alpha_{1,i})^2\big(1-(1-1/n_i)\,\alpha_{1,i}^2\big)}, \tag{21}$$

for $i = 1, 2$. *Furthermore, the cross-covariance function takes the form*

$$\mathrm{Cov}(X_{t,1}, X_{t,2}) = \frac{k\phi}{1-\alpha_{1,1}\alpha_{1,2}}h(X_{t,1}, X_{t,2}) \tag{22}$$

with

$$h(X_{t,1}, X_{t,2}) := \mathrm{E}\left[\sqrt{\left(\alpha_{0,1}+\alpha_{1,1}\frac{X_{t,1}}{n_1}\right)\left(1-\alpha_{0,1}-\alpha_{1,1}\frac{X_{t,1}}{n_1}\right)\left(\alpha_{0,2}+\alpha_{1,2}\frac{X_{t,2}}{n_2}\right)\left(1-\alpha_{0,2}-\alpha_{1,2}\frac{X_{t,2}}{n_2}\right)}\right].$$

For $i \neq j$ and $s > 0$, it follows that

$$\mathrm{Cov}(X_{t,i}, X_{t-s,j}) = \alpha_{1,i}^{s} \cdot \mathrm{Cov}(X_{t,1}, X_{t,2}). \tag{23}$$

Remark 3.3. Theorem 3.2 implies that the marginal distribution of $X_{t,i}$ is overdispersed with regard to a binomial distribution (extra-binomial variation), since

$$\frac{\mathrm{V}[X_{t,i}]}{\mathrm{E}[X_{t,i}]\left(1 - \mathrm{E}[X_{t,i}]/n_i\right)} = \frac{1}{1 - (1 - 1/n_i)\,\alpha_{1,i}^2} > 1. \tag{24}$$

Proof. The stationary mean μ_i has to satisfy

$$\mu_i := \mathrm{E}[X_{t,i}] = \mathrm{E}\big[\mathrm{E}[X_{t,i}|\mathscr{F}_{t-1}]\big] \overset{(19)}{=} n_i\,\alpha_{0,i} + \alpha_{1,i}\,\mathrm{E}[X_{t-1,i}] = n_i\,\alpha_{0,i} + \alpha_{1,i}\,\mu_i,$$

which implies the formula for the mean. For the stationary variance σ_i^2, we obtain

$$\begin{aligned}
\sigma_i^2 &:= \mathrm{V}[X_{t,i}] = \mathrm{V}\big[\mathrm{E}[X_{t,i}|\mathscr{F}_{t-1}]\big] + \mathrm{E}\big[\mathrm{V}[X_{t,i}|\mathscr{F}_{t-1}]\big] \\
&\overset{(19)}{=} \alpha_{1,i}^2\,\sigma_i^2 + n_i\alpha_{0,i}(1-\alpha_{0,i}) + (1-2\alpha_{0,i})\alpha_{1,i}\,\mu_i - \alpha_{1,i}^2\,(\sigma_i^2 + \mu_i^2)/n_i \\
&= (1 - 1/n_i)\,\alpha_{1,i}^2\,\sigma_i^2 + n_i\alpha_{0,i}(1-\alpha_{0,i}) + n_i\,\alpha_{0,i}(1-2\alpha_{0,i})\,\frac{\alpha_{1,i}}{1-\alpha_{1,i}} - n_i\,\alpha_{0,i}^2\,\frac{\alpha_{1,i}^2}{(1-\alpha_{1,i})^2}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\sigma_i^2 &= \frac{n_i\alpha_{0,i}}{(1-\alpha_{1,i})^2} \cdot \frac{(1-\alpha_{0,i})(1-\alpha_{1,i})^2 + (1-2\alpha_{0,i})\,\alpha_{1,i}(1-\alpha_{1,i}) - \alpha_{0,i}\,\alpha_{1,i}^2}{1 - (1-1/n_i)\,\alpha_{1,i}^2} \\
&= \frac{n_i\alpha_{0,i}}{(1-\alpha_{1,i})^2} \cdot \frac{(1-\alpha_{1,i})^2 + \alpha_{1,i}(1-\alpha_{1,i}) - \alpha_{0,i}\left((1-\alpha_{1,i})^2 + 2\alpha_{1,i}(1-\alpha_{1,i}) + \alpha_{1,i}^2\right)}{1 - (1-1/n_i)\,\alpha_{1,i}^2} \\
&= \frac{n_i\alpha_{0,i}}{(1-\alpha_{1,i})^2} \cdot \frac{1-\alpha_{0,i}-\alpha_{1,i}}{1-(1-1/n_i)\,\alpha_{1,i}^2}.
\end{aligned}$$

Moreover, to compute $\mathrm{Cov}(X_{t,1}, X_{t,2})$ note that

$$\begin{aligned}
\mathrm{Cov}(X_{t,1}, X_{t,2}) &= \mathrm{E}\big[\mathrm{Cov}(X_{t,1}, X_{t,2}|\mathscr{F}_{t-1})\big] + \mathrm{Cov}\big(\mathrm{E}[X_{t,1}|\mathscr{F}_{t-1}], \mathrm{E}[X_{t,2}|\mathscr{F}_{t-1}]\big) \\
&= k\phi \cdot h(X_{t,1}, X_{t,2}) + \alpha_{1,1}\alpha_{1,2} \cdot \mathrm{Cov}(X_{t,1}, X_{t,2}),
\end{aligned}$$

where the last step is justified by stationarity and the results in (11) and (19). Finally, in proving (23) note that

$$\begin{aligned}
\mathrm{Cov}(X_{t,i}, X_{t-s,j}) &= \mathrm{E}\big[\mathrm{Cov}(X_{t,i}, X_{t-s,j} \mid \mathscr{F}_{t-1})\big] + \mathrm{Cov}\big(\mathrm{E}[X_{t,i} \mid \mathscr{F}_{t-1}],\ \mathrm{E}[X_{t-s,j} \mid \mathscr{F}_{t-1}]\big) \\
&= \alpha_{1,i} \cdot \mathrm{Cov}(X_{t-1,i}, X_{t-s,j}) \\
&\ \ \vdots \\
&= \alpha_{1,i}^{s} \cdot \mathrm{Cov}(X_{t-s,i}, X_{t-s,j}).
\end{aligned}$$

Again, by stationarity the proof of (23) is completed. □

4. Bivariate binomial thinning operation

Let $\boldsymbol{X} := [X_1\ X_2]'$ be a bivariate random variable, abbreviate $\boldsymbol{\alpha} := (\alpha_1, \alpha_2, \phi_\alpha)$ with $0 < \alpha_1, \alpha_2 < 1$ and ϕ satisfying (7). We define the *bivariate binomial thinning operation* as follows:

$$\boldsymbol{\alpha} \otimes \boldsymbol{X} \mid \boldsymbol{X} \sim \mathrm{BVB_{II}}\,(X_1, X_2, \min(X_1, X_2); \alpha_1, \alpha_2, \phi_\alpha)\,. \tag{25}$$

In analogy to the case of the diagonal matrix thinning used by Karlis and Pedeli [21] and Pedeli and Karlis [30,31,29], relation (10) implies that

$$\begin{aligned}
(\boldsymbol{\alpha} \otimes \boldsymbol{X})_1 \mid \boldsymbol{X} &\sim \mathrm{Bi}(X_1, \alpha_1), \\
(\boldsymbol{\alpha} \otimes \boldsymbol{X})_2 \mid \boldsymbol{X} &\sim \mathrm{Bi}(X_2, \alpha_2),
\end{aligned} \tag{26}$$

i.e., the marginals are thinned according to the usual binomial thinning operation of Steutel and van Harn [36]. Lemma 4.1 summarizes some important properties of the bivariate binomial thinning operation.

Lemma 4.1. *The bivariate binomial thinning operation in (25) has the following properties*

- *Mean value:* $\mathrm{E}\Big(\big(\boldsymbol{\alpha} \otimes \boldsymbol{X}\big)_i\Big) = \alpha_i \cdot \mathrm{E}(X_i),\ i = 1, 2;$
- *Variance:* $\mathrm{V}\Big(\big(\boldsymbol{\alpha} \otimes \boldsymbol{X}\big)_i\Big) = \alpha_i(1-\alpha_i) \cdot \mathrm{E}(X_i) + \alpha_i^2 \cdot \mathrm{V}(X_i),\ i = 1, 2;$
- *Covariance:*

$$\mathrm{Cov}\Big(\big(\boldsymbol{\alpha} \otimes \boldsymbol{X}\big)_1,\ \big(\boldsymbol{\alpha} \otimes \boldsymbol{X}\big)_2\Big) = \alpha_1\alpha_2 \cdot \mathrm{Cov}(X_1, X_2) + \phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)} \cdot \mathrm{E}\Big(\min(X_1, X_2)\Big). \tag{27}$$

Proof. First note that

$$\begin{aligned}\mathrm{E}\big(\boldsymbol{\alpha} \otimes \boldsymbol{X}\big) &= \mathrm{E}\Big(\mathrm{E}\big(\boldsymbol{\alpha} \otimes \boldsymbol{X} \mid \boldsymbol{X}\big)\Big)\\ &= \mathrm{E}\big([\alpha_1 \cdot X_1 \alpha_2 \cdot X_2]'\big) = \big[\alpha_1 \cdot \mathrm{E}(X_1)\alpha_2 \cdot \mathrm{E}(X_2)\big]'.\end{aligned} \tag{28}$$

Next, define $\boldsymbol{Z} := \boldsymbol{\alpha} \otimes \boldsymbol{X}$, i.e., $[Z_1\ Z_2]' = (\alpha_1, \alpha_2, \phi_\alpha) \otimes [X_1\ X_2]'$. Thus,

$$\mathrm{V}\Big(\boldsymbol{\alpha} \otimes \boldsymbol{X}\Big) = \begin{pmatrix} \mathrm{V}(Z_1) & \mathrm{Cov}(Z_1, Z_2) \\ \mathrm{Cov}(Z_1, Z_2) & \mathrm{V}(Z_2) \end{pmatrix},$$

where

$$\begin{aligned}\mathrm{V}(Z_1) &= \mathrm{E}\Big(\mathrm{V}\big(Z_1 \mid \boldsymbol{X}\big)\Big) + \mathrm{V}\Big(\mathrm{E}\big(Z_1 \mid \boldsymbol{X}\big)\Big)\\ &= \mathrm{E}\Big(\alpha_1(1-\alpha_1) \cdot X_1\Big) + \mathrm{V}\Big(\alpha_1 \cdot X_1\Big)\\ &= \alpha_1(1-\alpha_1) \cdot \mathrm{E}(X_1) + \alpha_1^2 \cdot \mathrm{V}(X_1),\end{aligned}$$

and, similarly

$$\mathrm{V}(Z_2) = \alpha_2(1-\alpha_2) \cdot \mathrm{E}(X_2) + \alpha_2^2 \cdot \mathrm{V}(X_2).$$

Finally, to compute $\mathrm{Cov}(Z_1, Z_2)$ note that

$$\mathrm{Cov}(Z_1, Z_2) = \mathrm{E}\Big(\mathrm{Cov}\big(Z_1, Z_2 \mid \boldsymbol{X}\big)\Big) + \mathrm{Cov}\Big(\mathrm{E}\big(Z_1 \mid \boldsymbol{X}\big),\ \mathrm{E}\big(Z_2 \mid \boldsymbol{X}\big)\Big).$$

By the result in (11), it follows that

$$\begin{aligned}\mathrm{E}\Big(\mathrm{Cov}\big(Z_1, Z_2 \mid \boldsymbol{X}\big)\Big) &= \mathrm{E}\Big(\min(X_1, X_2)\phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}\Big)\\ &= \mathrm{E}\Big(\min(X_1, X_2)\Big)\phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)},\end{aligned}$$

and

$$\mathrm{Cov}\Big(\mathrm{E}(Z_1 \mid \boldsymbol{X}),\ \mathrm{E}(Z_2 \mid \boldsymbol{X})\Big) = \mathrm{Cov}(\alpha_1 \cdot X_1, \alpha_2 \cdot X_2) = \alpha_1\alpha_2 \cdot \mathrm{Cov}(X_1, X_2).$$

Thus,

$$\mathrm{Cov}\Big(Z_1, Z_2\Big) = \phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)} \cdot \mathrm{E}\Big(\min(X_1, X_2)\Big) + \alpha_1\alpha_2 \cdot \mathrm{Cov}(X_1, X_2),$$

which completes the proof of (27). □

Note that, as compared to the case of diagonal matrix thinning in (5), bivariate binomial thinning causes additional cross-correlation as long as $\phi_\alpha \neq 0$. Furthermore, the cross-correlation can even be negative if $\phi_\alpha < 0$. In fact, the diagonal matrix thinning used by Karlis and Pedeli [21] and Pedeli and Karlis [30,31,29] can not only be understood as a special case of the matrix thinning concept (4), but it might also be understood as a special case of bivariate binomial thinning with $\phi_\alpha = 0$.

In addition to formulae (27) and (28) for unconditional moments of $\boldsymbol{\alpha} \otimes \boldsymbol{X}$, the unconditional pmf can be computed through

$$\begin{aligned}P(\boldsymbol{\alpha} \otimes \boldsymbol{X} = \boldsymbol{z}) &= \sum_{x_1=z_1}^{\infty}\sum_{x_2=z_2}^{\infty} P(\boldsymbol{\alpha} \otimes \boldsymbol{X} = \boldsymbol{z} \mid \boldsymbol{X} = \boldsymbol{x}) \cdot P(\boldsymbol{X} = \boldsymbol{x})\\ &= \sum_{x_1=z_1}^{\infty}\sum_{x_2=z_2}^{\infty} p_{(x_1, x_2; \alpha_1, \alpha_2, \phi_\alpha)}(z_1, z_2) \cdot P(\boldsymbol{X} = \boldsymbol{x}),\end{aligned} \tag{29}$$

where $p_{(n_1,n_2;\alpha_1,\alpha_2,\phi_\alpha)}(x_1,x_2)$ denotes the bivariate pmf according to formula (13). For the pgf of $\boldsymbol{\alpha}\otimes\boldsymbol{X}$, however, it is not possible to find a simple closed-form formula. Using (15), we obtain

$$\begin{aligned} G_{\boldsymbol{Z}}(z_1,z_2) &= \mathrm{E}\Big[(1-\alpha_1+\alpha_1 z_1)^{X_1-\min(X_1,X_2)}\cdot(1-\alpha_2+\alpha_2 z_2)^{X_2-\min(X_1,X_2)} \\ &\quad\cdot\Big((1-\alpha_1+\alpha_1 z_1)(1-\alpha_2+\alpha_2 z_2)+\phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}\cdot(1-z_1)(1-z_2)\Big)^{\min(X_1,X_2)}\Big] \\ &= \mathrm{E}\Bigg[(1-\alpha_1+\alpha_1 z_1)^{X_1}\cdot(1-\alpha_2+\alpha_2 z_2)^{X_2} \\ &\quad\cdot\left(1+\phi_\alpha\frac{\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)}\cdot(1-z_1)(1-z_2)}{(1-\alpha_1+\alpha_1 z_1)(1-\alpha_2+\alpha_2 z_2)}\right)^{\min(X_1,X_2)}\Bigg], \end{aligned}$$

which could also be expressed in terms of the bivariate pgf of $(X_1+X_2, |X_1-X_2|)$ by using the relation $2\cdot\min(a,b)=a+b-|a-b|$.

5. The $\mathrm{BVB_{II}}$-AR(1) model

In this section, a bivariate extension of the binomial AR(1) model in (3) based on the bivariate binomial thinning operation (25) is introduced. The definition of the bivariate binomial AR(1) model ($\mathrm{BVB_{II}}$-AR(1)) is given below.

Definition 5.1. Let $\pi_1,\pi_2\in(0;1)$, and

$$\rho_i\in\left(\max\left\{-\frac{\pi_i}{1-\pi_i},-\frac{1-\pi_i}{\pi_i}\right\};1\right),\quad i=1,2.$$

Define $\beta_i:=\pi_i\cdot(1-\rho_i)$ and $\alpha_i:=\beta_i+\rho_i$, for $i=1,2$. Let $\boldsymbol{n}:=[n_1\ n_2]'\in\mathbb{N}^2$ denote the vector of upper limits for the bivariate range and $\boldsymbol{\alpha}:=(\alpha_1,\alpha_2,\phi_\alpha)$, and $\boldsymbol{\beta}:=(\beta_1,\beta_2,\phi_\beta)$ as the thinning parameters. A $\mathrm{BVB_{II}}$-AR(1) process is a bivariate process $(\boldsymbol{X}_t)$ of bivariate random variables $\boldsymbol{X}_t:=[X_{t,1}\ X_{t,2}]'$ satisfying the following recursion

$$\boldsymbol{X}_t=\boldsymbol{\alpha}\otimes\boldsymbol{X}_{t-1}+\boldsymbol{\beta}\otimes(\boldsymbol{n}-\boldsymbol{X}_{t-1}),\quad t\in\mathbb{Z}, \tag{30}$$

where

$$\boldsymbol{\alpha}\otimes\boldsymbol{X}_{t-1}\mid\boldsymbol{X}_{t-1}\sim\mathrm{BVB_{II}}(X_{t-1,1},X_{t-1,2},\min(X_{t-1,1},X_{t-1,2});\alpha_1,\alpha_2,\phi_\alpha)$$

and

$$\boldsymbol{\beta}\otimes(\boldsymbol{n}-\boldsymbol{X}_{t-1})\mid\boldsymbol{X}_{t-1}\sim\mathrm{BVB_{II}}(n_1-X_{t-1,1},n_2-X_{t-1,2},\min(n_1-X_{t-1,1},n_2-X_{t-1,2});\beta_1,\beta_2,\phi_\beta),$$

where the thinnings are performed independently of each other.

Note that condition on ρ_i guarantees that $\alpha_i,\beta_i\in(0;1)$.

The $\mathrm{BVB_{II}}$-AR(1) model according to Definition 5.1 offers the potential for application in many fields of practice. In a metapopulation context as in [43], for instance, the two components of $\boldsymbol{X}_t$ refer to two metapopulations consisting of n_1 and n_2 habitat patches, respectively, where the colonization and extinction mechanisms are possibly cross-correlated because of mutual competition or mutual exchange. Another scenario might be the monitoring along time of two groups of patients with different types of medical treatment, e.g., with regard to the therapeutic outcome or the severity of symptoms. For some applications, it could also be justified to simplify the model parametrization. As an example, setting $\phi_\alpha=0$ in a metapopulation context implies that the survival of the already occupied patches is not affected by the other metapopulation, while such an interaction is possible for the case of the colonization of empty patches provided that $\phi_\beta\neq 0$.

The transition probabilities of the $\mathrm{BVB_{II}}$-AR(1) model according to Definition 5.1 are computed by using Eq. (13) as follows:

$$\begin{aligned} p(\boldsymbol{x}|\boldsymbol{y}) &:= \mathrm{P}\Big(\boldsymbol{X}_t=\boldsymbol{x}\ \Big|\ \boldsymbol{X}_{t-1}=\boldsymbol{y}\Big) \\ &= \mathrm{P}\Big(\boldsymbol{\alpha}\otimes\boldsymbol{X}_{t-1}+\boldsymbol{\beta}\otimes(\boldsymbol{n}-\boldsymbol{X}_{t-1})=\boldsymbol{x}\ \Big|\ \boldsymbol{X}_{t-1}=\boldsymbol{y}\Big) \\ &= \sum_{a_1=0}^{\min(x_1,y_1)}\sum_{a_2=0}^{\min(x_2,y_2)}\mathrm{P}\Big(\boldsymbol{\alpha}\otimes\boldsymbol{X}_{t-1}=\boldsymbol{a}\ \Big|\ \boldsymbol{X}_{t-1}=\boldsymbol{y}\Big)\cdot\mathrm{P}\Big(\boldsymbol{\beta}\otimes(\boldsymbol{n}-\boldsymbol{X}_{t-1})=\boldsymbol{x}-\boldsymbol{a}\ \Big|\ \boldsymbol{X}_{t-1}=\boldsymbol{y}\Big) \\ &= \sum_{a_1=0}^{\min(x_1,y_1)}\sum_{a_2=0}^{\min(x_2,y_2)}p_{(y_1,y_2;\alpha_1,\alpha_2,\phi_\alpha)}(a_1,a_2)\cdot p_{(n_1-y_1,n_2-y_2;\beta_1,\beta_2,\phi_\beta)}(x_1-a_1,x_2-a_2), \end{aligned} \tag{31}$$

where in the last line we used the bivariate pmf of the $\mathrm{BVB_{II}}$-distribution as introduced in Section 2. Likewise, as for the $\mathrm{BVB_{II}}$-INARCH(1) model from Section 3, since these transition probabilities are truly positive the $\mathrm{BVB_{II}}$-AR(1) process is a primitive and finite-state Markov chain (also ψ- and φ-mixing). Thus, a uniquely determined stationary marginal distribution exists, which can be computed from the transition matrix, see the discussion after formula (17) in Section 2.

Since BVB thinning behaves marginally as binomial thinning, i.e., $(\boldsymbol{\alpha} \otimes \boldsymbol{X})_i \overset{D}{=} \alpha_i \circ X_i$, also the marginals of the $\mathrm{BVB_{II}}$-AR(1) process are distributed like the usual binomial AR(1) process according to (3). In particular, for the unique stationary marginal distribution of $\boldsymbol{X}_t$, we have that $X_{t,i} \sim \mathrm{Bi}(n_i, \pi_i)$ and $\rho(X_{t,i}, X_{t-k,i}) = \rho_i^k$; see [26]. Moreover, marginal conditional moments are obtained as

$$\begin{aligned} &\mathrm{E}(X_{t,i}|\boldsymbol{X}_{t-1}) = \rho_i \cdot X_{t-1,i} + n_i\beta_i, \\ &\mathrm{V}(X_{t,i}|\boldsymbol{X}_{t-1}) = \rho_i(1-\rho_i)(1-2\pi_i) \cdot X_{t-1,i} + n_i\beta_i(1-\beta_i). \end{aligned} \tag{32}$$

Concerning the cross-covariance of $X_{t,1}$ and $X_{t,2}$, we obtain the following result.

Theorem 5.2 (*Cross-Covariance*)**.** *If* $(\boldsymbol{X}_t)$ *is a stationary* $\mathrm{BVB_{II}}$-*AR*(1) *process, then*

$$\begin{aligned} \mathrm{Cov}(X_{t,1}, X_{t,2}) = \frac{1}{1-\rho_1\rho_2}\Big(&\phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)} \cdot \mathrm{E}\big(\min(X_{t,1}, X_{t,2})\big) \\ &+ \phi_\beta\sqrt{\beta_1\beta_2(1-\beta_1)(1-\beta_2)} \cdot \mathrm{E}\big(\min(n_1 - X_{t,1}, n_2 - X_{t,2})\big)\Big). \end{aligned} \tag{33}$$

Furthermore, for $i \neq j$ *and* $s > 0$*, we have*

$$\mathrm{Cov}(X_{t,i}, X_{t-s,j}) = \rho_i^s \cdot \mathrm{Cov}(X_{t,1}, X_{t,2}). \tag{34}$$

Proof. Since $X_{t,i} = \big(\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}\big)_i + \big(\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})\big)_i$, it follows that

$$\begin{aligned} \mathrm{Cov}(X_{t,1}, X_{t,2}) = &\,\mathrm{Cov}\Big(\big(\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}\big)_1, \big(\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}\big)_2\Big) + \mathrm{Cov}\Big(\big(\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})\big)_1, \big(\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})\big)_2\Big) \\ &+ \mathrm{Cov}\Big(\big(\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}\big)_1, \big(\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})\big)_2\Big) + \mathrm{Cov}\Big(\big(\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})\big)_1, \big(\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}\big)_2\Big). \end{aligned}$$

The first two summands are computed by using formula (27). For the remaining summands, we have to note that $\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}$ and $\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})$ are performed independently. Hence, by conditioning

$$\begin{aligned} \mathrm{Cov}\Big(\big(\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}\big)_1, \big(\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})\big)_2\Big) &= 0 + \mathrm{Cov}\Big(\mathrm{E}\big[\big(\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}\big)_1 \mid \boldsymbol{X}_{t-1}\big],\ \mathrm{E}\big[\big(\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})\big)_2 \mid \boldsymbol{X}_{t-1}\big]\Big) \\ &= \alpha_1\beta_2 \cdot \mathrm{Cov}\big(X_{t-1,1}, n_2 - X_{t-1,2}\big) = -\alpha_1\beta_2 \cdot \mathrm{Cov}\big(X_{t-1,1}, X_{t-1,2}\big), \end{aligned}$$

and analogously,

$$\mathrm{Cov}\Big(\big(\boldsymbol{\beta} \otimes (\boldsymbol{n} - \boldsymbol{X}_{t-1})\big)_1, \big(\boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1}\big)_2\Big) = -\alpha_2\beta_1 \cdot \mathrm{Cov}\big(X_{t-1,1}, X_{t-1,2}\big).$$

Since

$$\alpha_1\alpha_2 - \alpha_1\beta_2 - \alpha_2\beta_1 + \beta_1\beta_2 = (\alpha_1 - \beta_1)(\alpha_2 - \beta_2) = \rho_1\rho_2,$$

we obtain

$$\begin{aligned} \mathrm{Cov}(X_{t,1}, X_{t,2}) = &\,\rho_1\rho_2 \cdot \mathrm{Cov}\big(X_{t-1,1}, X_{t-1,2}\big) + \phi_\alpha\sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)} \cdot \mathrm{E}\Big(\min(X_{t-1,1}, X_{t-1,2})\Big) \\ &+ \phi_\beta\sqrt{\beta_1\beta_2(1-\beta_1)(1-\beta_2)} \cdot \mathrm{E}\Big(\min(n_1 - X_{t-1,1}, n_2 - X_{t-1,2})\Big). \end{aligned}$$

Thus, the expression in (33) follows by the stationary assumption. Moreover, from (32) we have

$$\begin{aligned} \mathrm{Cov}(X_{t,i}, X_{t-s,j}) &= \mathrm{Cov}\Big(\mathrm{E}\big(X_{t,i} \mid \boldsymbol{X}_{t-1}, \ldots\big), \mathrm{E}\big(X_{t-s,j} \mid \boldsymbol{X}_{t-1}, \ldots\big)\Big) + \mathrm{E}\Big(\mathrm{Cov}\big(X_{t,i}, X_{t-s,j} \mid \boldsymbol{X}_{t-1}, \ldots\big)\Big) \\ &= \mathrm{Cov}\big(\rho_i \cdot X_{t-1,i} + n_i\beta_i,\ X_{t-s,j}\big) + 0 \\ &= \rho_i \cdot \mathrm{Cov}(X_{t-1,i}, X_{t-s,j}) = \cdots = \rho_i^s \cdot \mathrm{Cov}(X_{t-s,i}, X_{t-s,j}). \end{aligned}$$

Again, by the stationarity assumption the proof of (34) is complete. □

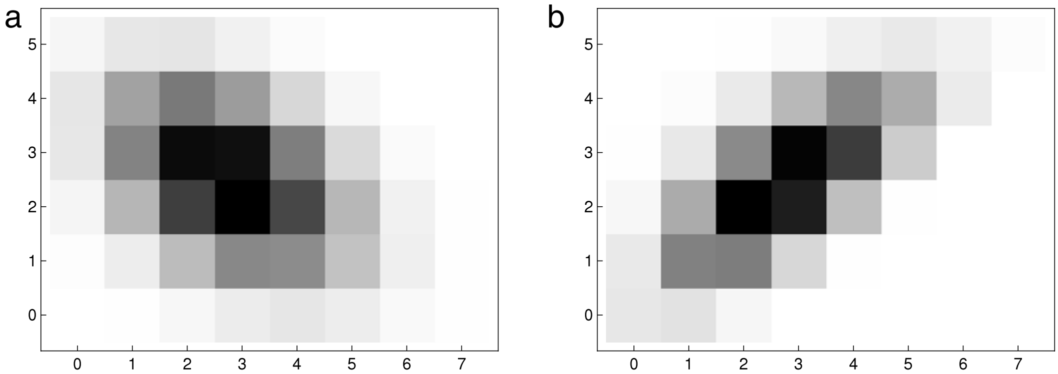

Fig. 1. Stationary marginal distribution of $\mathrm{BVB_{II}}$-AR(1) models with $(n_1, n_2, \pi_1, \pi_2, \rho_1, \rho_2) = (5, 7, 0.5, 0.4, 0.3, 0.3)$ and (a) $(\phi_\alpha, \phi_\beta) = (-0.62, -0.45)$, (b) $(\phi_\alpha, \phi_\beta) = (0.86, 0.84)$.

5.1. Numerical illustrations

Formula (33) in Theorem 5.2 shows that the components $X_{t,1}$ and $X_{t,2}$ are generally cross-correlated, provided that $(\phi_\alpha, \phi_\beta) \neq (0, 0)$. The degree of cross-correlation extends from negative to positive values depending on the signs of $(\phi_\alpha, \phi_\beta)$. A similar behavior is obtained for the $\mathrm{BVB_{II}}$-INARCH(1) model from Section 3, where the sign and degree of cross-correlation are controlled by the parameter ϕ (see (22)). To illustrate the extend of cross-correlation and to compare both types of model with each other, consider a $\mathrm{BVB_{II}}$-AR(1) model with $(n_1, n_2, \pi_1, \pi_2, \rho_1, \rho_2) = (5, 7, 0.5, 0.4, 0.3, 0.3)$ and a $\mathrm{BVB_{II}}$-INARCH(1) model with $(n_1, n_2, \alpha_{0,1}, \alpha_{1,1}, \alpha_{0,2}, \alpha_{1,2}) = (5, 7, 0.35, 0.3, 0.28, 0.3)$. These parameters are chosen so that the marginal processes $(X_{t,i}^{\mathrm{BVB_{II}\text{-}AR}})$ and $(X_{t,i}^{\mathrm{BVB_{II}\text{-}INARCH}})$ have the same mean and autocorrelation function.

For the $\mathrm{BVB_{II}}$-AR(1) model, it follows from Definition 5.1 that $(\alpha_1, \beta_1) = (0.65, 0.35)$ and $(\alpha_2, \beta_2) = (0.58, 0.28)$. Hence, formula (7) implies that we have to choose ϕ_α between -0.6244 and 0.8623, and ϕ_β between -0.4576 and 0.8498. We select two models from the opposite ends of this scale: model (a) with $(\phi_\alpha, \phi_\beta) = (-0.62, -0.45)$, and model (b) with $(\phi_\alpha, \phi_\beta) = (0.86, 0.84)$. Analogously, formula (18) for the $\mathrm{BVB_{II}}$-INARCH(1) model implies to choose ϕ between -0.4576 and 0.4576, so we decided for model (c) with $\phi = -0.45$, and model (d) with $\phi = 0.45$.

The stationary marginal distribution of $\boldsymbol{X}_t$ can be computed from the transition probability matrix according to (31) and (17), respectively. From this distribution, in turn, the required moments $\mathrm{E}\big(f(X_{t,1}, X_{t,2})\big)$ follow. Let us start with the $\mathrm{BVB_{II}}$-AR(1) model. There, we obtain

- $\mathrm{E}\big(\min(X_{t,1}, X_{t,2})\big)$ as 1.851 for model (a) and 2.279 for model (b),
- $\mathrm{E}\big(\min(n_1 - X_{t,1}, n_2 - X_{t,2})\big)$ as 2.282 for model (a) and 2.500 for model (b), and
- $\mathrm{Cov}(X_{t,1}, X_{t,2})$ as -0.539 for model (a) and 1.001 for model (b).

These results are in accordance with expression (33) in Theorem 5.2. In particular, the components of model (a) show a considerable degree of negative cross-correlation (-0.372), while they exhibit a strong positive cross-correlation (0.691) for model (b). This is well illustrated by the density plot of the respective pmf in Fig. 1, where the probability concentrates either on the antidiagonal or the main diagonal.

For the $\mathrm{BVB_{II}}$-INARCH(1) models (c) and (d), the cross-covariances equal -0.595 and 0.595, respectively, leading to the cross-correlations -0.380 and 0.380. A plot of the stationary marginal distributions is shown in Fig. 2.

It is noteworthy that although both models can be used for describing considerable degrees of cross-correlation, they differ regarding the marginal variances $\mathrm{V}(X_{t,1})$ and $\mathrm{V}(X_{t,2})$. In fact, the marginal variances for the above models (a)–(d) are 1.250, 1.680, 1.347 and 1.820, respectively, clearly indicating that the considered $\mathrm{BVB_{II}}$-INARCH(1) models show about 8% of extra-binomial variation. Note that the inequality in (24) implies that the degree of extra-binomial variation is controlled by the dependence parameters $\alpha_{1,1}$ and $\alpha_{1,2}$.

Remark 5.3. The stationary marginal distribution of a $\mathrm{BVB_{II}}$-AR(1) process does not generally belong to the $\mathrm{BVB_{II}}$ family according to Section 2.2. This is easily verified for the models being considered above by comparing their pmfs with the pmfs of all relevant $\mathrm{BVB_{II}}(n_1, n_2, k; \pi_1, \pi_2, \phi_\pi)$ distributions, i.e., where (k, ϕ_π) with $1 \le k \le \min(n_1, n_2)$ is chosen so that the cross-correlation according to (12) equals the above values of about -0.372 and 0.691, respectively.

6. Parameter estimation for $\mathrm{BVB_{II}}$-INARCH(1) and $\mathrm{BVB_{II}}$-AR(1) processes

In this section, we consider the parameter estimation of both $\mathrm{BVB_{II}}$-INARCH(1) and $\mathrm{BVB_{II}}$-AR(1) processes. In particular, the conditional maximum likelihood (CML) method is adopted. For this purpose let $\boldsymbol{X} := (\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_T)$ be a finite time series from either a $\mathrm{BVB_{II}}$-INARCH(1) with vector of unknown parameters $\boldsymbol{\theta} := (\alpha_{0,1}, \alpha_{1,1}, \alpha_{0,2}, \alpha_{1,2}, \phi)$ or a $\mathrm{BVB_{II}}$-AR(1)

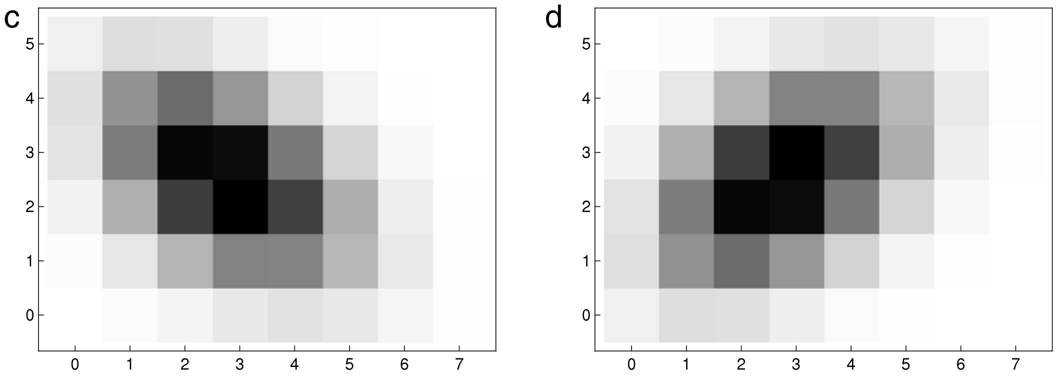

Fig. 2. Stationary marginal distribution of BVB_{II}-INARCH(1) models with $(n_1, n_2, \alpha_{0,1}, \alpha_{1,1}, \alpha_{0,2}, \alpha_{1,2}) = (5, 7, 0.35, 0.3, 0.28, 0.3)$ and (c) $\phi = -0.45$, (d) $\phi = 0.45$.

model with $\boldsymbol{\theta} := (\pi_1, \pi_2, \rho_1, \rho_2, \phi_\alpha, \phi_\beta)$ or $\boldsymbol{\theta} := (\alpha_1, \alpha_2, \phi_\alpha, \beta_1, \beta_2, \phi_\beta)$, respectively. The latter parametrization has the practical advantage that $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ have to satisfy the box constraint $(0; 1)^4$, while the first one is more relevant for theoretical analysis. However, the estimates for $(\alpha_1, \alpha_2, \beta_1, \beta_2)$ are easily transformed into those for $(\pi_1, \pi_2, \rho_1, \rho_2)$ (and vice versa) according to the relations given in Definition 5.1.

The CML-estimators are obtained by maximizing the conditional log-likelihood function

$$l(\boldsymbol{\theta}) := \log \mathscr{L}(\boldsymbol{X}, \boldsymbol{\theta}|\boldsymbol{X}_0) \equiv \sum_{t=1}^{T} \log \mathrm{P}\Big(\boldsymbol{X}_t = \boldsymbol{x}_t \;\Big|\; \boldsymbol{X}_{t-1} = \boldsymbol{y}_{t-1}\Big)$$

with $\mathrm{P}\Big(\boldsymbol{X}_t = \boldsymbol{x}_t \;\Big|\; \boldsymbol{X}_{t-1} = \boldsymbol{y}_{t-1}\Big)$ defined as in (17) and subject to (18) and $0 < \alpha_{0,i} + \alpha_{1,i} < 1,\ i = 1, 2$, for the BVB_{II}-INARCH(1) model or as in (31), subject to (7) and

$$\max\left\{-\sqrt{\frac{\beta_1\beta_2}{(1-\beta_1)(1-\beta_2)}}, -\sqrt{\frac{(1-\beta_1)(1-\beta_2)}{\beta_1\beta_2}}\right\} < \phi_\beta < \min\left\{\sqrt{\frac{\beta_1(1-\beta_2)}{(1-\beta_1)\beta_2}}, \sqrt{\frac{(1-\beta_1)\beta_2}{\beta_1(1-\beta_2)}}\right\}, \tag{35}$$

for the BVB_{II}-AR(1) model.

Concerning the asymptotic behavior of the CML-estimators, it is important to note that both processes are finite-state Markov chains, so it suffices to check if condition 5.1 of Billingsley [4] holds, i.e., if the transition probabilities (17) and (31) are three times continuously differentiable with respect to $\boldsymbol{\theta}$, and the $((n_1+1)\cdot(n_2+1)) \times s$ matrix (being $s = 5$ for the BVB_{II}-INARCH(1) and $s = 6$ for the BVB_{II}-AR(1) model) with entries $\frac{\partial}{\partial\theta_j} p(\boldsymbol{x}|\boldsymbol{y})$, where θ_j represents the jth element of the parameter vector $\boldsymbol{\theta}$, has rank s. Note that both models are primitive and hence ergodic Markov chains without transient states. The differentiability of the transition probabilities according to (17) and (31) is determined through the differentiability of the pmf of the BVB_{II}-distribution, see (13) and (14). Obviously, this pmf has continuous partial derivatives up to any order within the allowed parameter range. Finally, the rank of the Jacobian matrix in condition 5.1 of Billingsley [4] is checked with the same arguments as in Sections 4.2 and 5.2 in [44]. Straightforward (although tedious) algebraic calculations lead to conclude that the matrix of first-order partial derivatives contains s independent rows (since its corresponding determinant takes a non-zero value) implying that the rank of the matrix is s. Therefore, we can apply Theorems 2.1 and 2.2 of Billingsley [4] and conclude that the CML-estimators exist in both cases and are consistent and asymptotically normal, with an (asymptotic) covariance matrix given by the inverse of the expected Fisher information.

Note that neither analytical estimates nor closed-form expressions for the expected Fisher information can be found and, thus, numerical procedures have to be employed. The initial estimates required by such numerical procedures can be obtained by the method of moments, see the details below. If a Newton-type optimization method is used, then one obtains the Hessian of the log-likelihood at the optimum and, hence, the observed Fisher information. This matrix is used to approximate the expected Fisher information such that approximate standard errors are obtained for the CML-estimates. The approximation of the expected Fisher information by the observed one is justified by Theorems 1.3 and 1.1 of Billingsley [4], which, in turn, are applicable since we are concerned with primitive and finite-state Markov chains.

For the BVB_{II}-INARCH(1) model, the estimates obtained from the method of moments (which can be determined through expressions (20)–(22)) can be used as starting values for $\alpha_{0,1}, \alpha_{1,1}, \alpha_{0,2}, \alpha_{1,2}$ and ϕ to initialize the algorithm. For the BVB_{II}-AR(1) model the starting values for $\alpha_1, \alpha_2, \beta_1$ and β_2 are the Yule–Walker estimators

$$\hat{\beta}_i^{(YW)} = \hat{\pi}_i^{(YW)}(1 - \hat{\rho}_i^{(YW)}), \qquad \hat{\alpha}_i^{(YW)} = \hat{\beta}_i^{(YW)} + \hat{\rho}_i^{(YW)}, \quad i = 1, 2,$$

Table 1
Set of parameters from the BVB_{II}-INARCH(1) and BVB_{II}-AR(1) processes.

Model	BVB_{II}-INARCH(1)					Model	BVB_{II}-AR(1)					
	$\alpha_{0,1}$	$\alpha_{1,1}$	$\alpha_{0,2}$	$\alpha_{1,2}$	ϕ		α_1	α_2	ϕ_α	β_1	β_2	ϕ_β
S1	0.35	0.30	0.28	0.30	0.20	M1	0.47	0.74	0.30	0.17	0.14	0.50
S2	0.50	0.12	0.60	0.15	0.40	M2	0.20	0.53	0.15	0.15	0.63	−0.20
S3	0.50	0.20	0.60	0.10	0.20	M3	0.78	0.32	−0.10	0.23	0.48	0.34

Table 2
Maximum likelihood estimates for $\boldsymbol{\theta}$ in the BVB_{II}-INARCH(1) model. Standard errors in parentheses.

Model	T	$\alpha_{0,1}$	$\alpha_{1,1}$	$\alpha_{0,2}$	$\alpha_{1,2}$	ϕ
S1	100	0.34 (0.05)	0.30 (0.10)	0.28 (0.04)	0.29 (0.09)	0.26 (0.09)
	500	0.35 (0.03)	0.29 (0.05)	0.28 (0.02)	0.29 (0.04)	0.21 (0.05)
	1000	0.35 (0.02)	0.29 (0.03)	0.28 (0.01)	0.29 (0.03)	0.20 (0.03)
S2	100	0.47 (0.07)	0.16 (0.11)	0.57 (0.09)	0.17 (0.12)	0.40 (0.21)
	500	0.50 (0.02)	0.12 (0.04)	0.60 (0.03)	0.14 (0.04)	0.38 (0.08)
	1000	0.50 (0.02)	0.12 (0.03)	0.60 (0.02)	0.14 (0.03)	0.38 (0.06)
S3	100	0.48 (0.08)	0.23 (0.12)	0.54 (0.11)	0.18 (0.12)	0.27 (0.12)
	500	0.49 (0.05)	0.22 (0.07)	0.58 (0.06)	0.13 (0.08)	0.24 (0.07)
	1000	0.50 (0.03)	0.20 (0.04)	0.59 (0.03)	0.11 (0.04)	0.24 (0.04)

Table 3
Maximum likelihood estimates for $\boldsymbol{\theta}$ in the BVB_{II}-AR(1) model. Standard errors in parentheses.

Model	T	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\phi}_\alpha$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\phi}_\beta$
M1	100	0.46 (0.06)	0.73 (0.04)	0.23 (0.24)	0.17 (0.02)	0.14 (0.02)	0.52 (0.15)
	500	0.47 (0.03)	0.74 (0.02)	0.30 (0.12)	0.17 (0.01)	0.14 (0.01)	0.50 (0.07)
	1000	0.47 (0.02)	0.74 (0.01)	0.30 (0.09)	0.17 (0.08)	0.14 (0.01)	0.50 (0.06)
M2	100	0.21 (0.10)	0.54 (0.06)	0.02 (0.39)	0.15 (0.02)	0.61 (0.09)	−0.27 (0.26)
	500	0.20 (0.04)	0.52 (0.02)	0.13 (0.26)	0.15 (0.01)	0.63 (0.03)	−0.22 (0.12)
	1000	0.20 (0.04)	0.53 (0.02)	0.12 (0.23)	0.15 (0.01)	0.63 (0.02)	−0.20 (0.11)
M3	100	0.77 (0.03)	0.31 (0.05)	−0.01 (0.19)	0.23 (0.05)	0.47 (0.04)	0.37 (0.21)
	500	0.78 (0.02)	0.32 (0.03)	−0.06 (0.14)	0.23 (0.02)	0.48 (0.02)	0.33 (0.11)
	1000	0.78 (0.01)	0.31 (0.02)	−0.08 (0.10)	0.23 (0.01)	0.47 (0.16)	0.33 (0.08)

where

$$\begin{cases} \hat{\pi}_i^{(YW)} = \dfrac{1}{n_i} \dfrac{\sum_{t=0}^{T} X_{t,i}}{T+1} \\ \hat{\rho}_i^{(YW)} = \dfrac{\sum_{t=2}^{T} (X_{t,i} - \bar{X}_i)(X_{t-1,j} - \bar{X}_j)}{\sum_{t=1}^{T} (X_{t,i} - \bar{X}_i)(X_{t,j} - \bar{X}_j)}. \end{cases}$$

Furthermore, initial estimates for ϕ_α and ϕ_β are randomly selected within the intervals defined in (7) and (35).

The simulation study contemplates the following combinations of model parameters for the BVB_{II}-INARCH(1) and BVB_{II}-AR(1) models (see Table 1).

Times series from the BVB_{II}-INARCH(1) and BVB_{II}-AR(1) processes in (16) and (30) of length 100, 500 and 1000 with 10,000 independent replicates were generated. In all cases $n_1 = 5$ and $n_2 = 7$. The results are summarized in Tables 2 and 3. Furthermore, Figs. 3 and 4 display boxplots of the biases for the estimates of models S1 and M1. Results for the other models are similar.

The results in Tables 2 and 3 illustrate the consistency of the estimators, since the standard errors of the estimators rapidly decrease to zero as T increases. Furthermore, Figs. 3 and 4 illustrate the small sample properties of the estimators: the estimates, componentwise, tend to be unbiased and consistent for the six sets of parameters. However, among these scenarios, it becomes clear that the estimation of ϕ or ϕ_α and ϕ_β is most challenging, since the bias and standard errors are rather large for $T = 100$, i.e., for short time series.

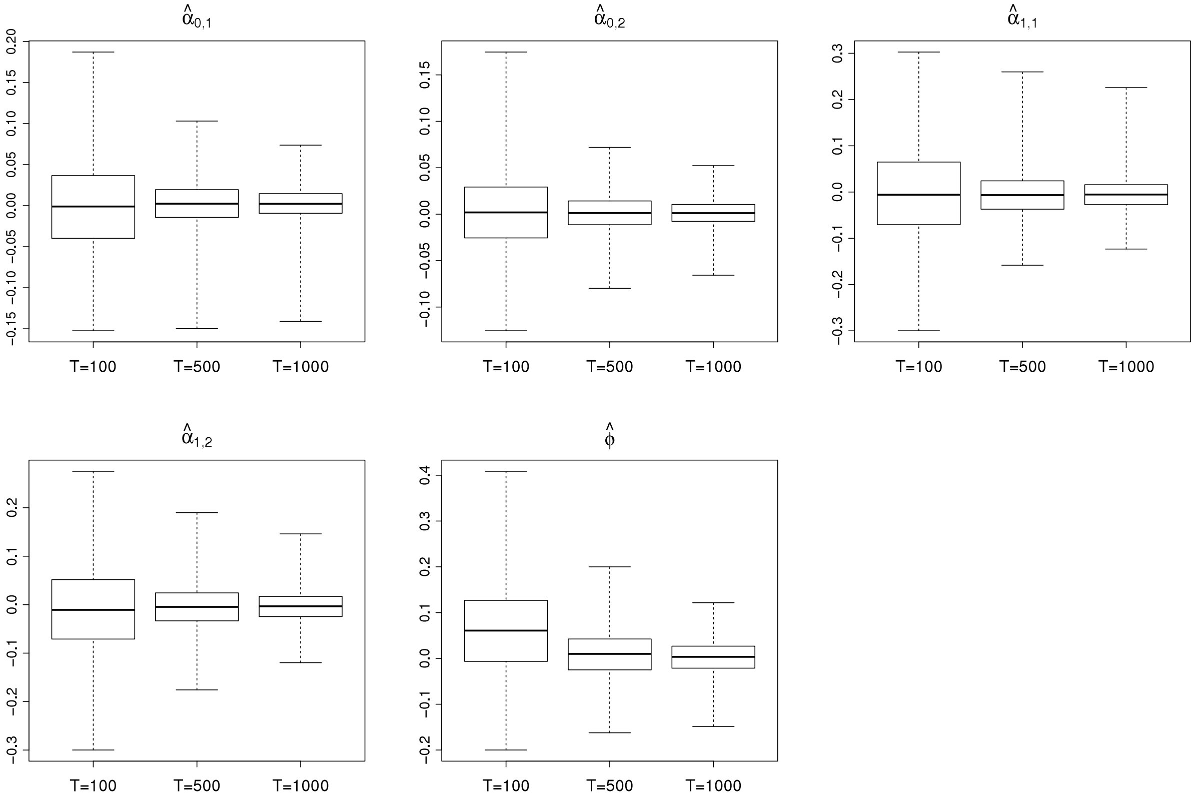

Fig. 3. Biases of the estimates for model S1.

7. Forecasting for BVB_{II}-INARCH(1) and BVB_{II}-AR(1) processes

In this section we consider the problem of predicting the values of $\boldsymbol{X}_{T+h},\ h \in \mathbb{N}$, for both BVB_{II}-INARCH(1) and the BVB_{II}-AR(1) processes based on the observed series up to time T. The usual way of producing forecasts is via the conditional forecast distribution. The following result establishes the h-step-ahead conditional distribution of $\boldsymbol{X}_{T+h}$ given $\boldsymbol{X}_T$.

Proposition 7.1. *The h-step-ahead conditional distribution of $\boldsymbol{X}_{T+h}$ given $\boldsymbol{X}_T$ is given by*

$$\mathrm{P}(\boldsymbol{X}_{T+h} = \boldsymbol{x}_{T+h} | \boldsymbol{X}_T = \boldsymbol{x}_T) = \boldsymbol{Q}^h \tag{36}$$

with $\boldsymbol{x}_{T+h} := (x_{T+h,1}, x_{T+h,2})'$, $\boldsymbol{x}_T := (x_{T,1}, x_{T,2})'$ *and*

$$\boldsymbol{Q} = \begin{bmatrix} p(0,0|0,0) & p(0,1|0,0) & \cdots & p(n_1,n_2|0,0) \\ p(0,0|0,1) & p(0,1|0,1) & \cdots & p(n_1,n_2|0,1) \\ \vdots & \vdots & \ddots & \vdots \\ p(0,0|n_1,n_2) & p(0,1|n_1,n_2) & \cdots & p(n_1,n_2|n_1,n_2) \end{bmatrix},$$

where $p(x_{T+1,1}, x_{T+1,2}|x_{T,1}, x_{T,2}) := \mathrm{P}(\boldsymbol{X}_{T+1} = \boldsymbol{x}_{T+1}|\boldsymbol{X}_T = \boldsymbol{x}_T)$ *is the* 1*-step-ahead conditional distribution given in* (17) *for the* BVB_{II}*-INARCH*(1) *process and equals* (31) *in the* BVB_{II}*-AR*(1) *case. Furthermore, for the* BVB_{II}*-AR*(1) *process*

$$\mathrm{E}[\boldsymbol{X}_{T+h}|\boldsymbol{X}_T] = \begin{bmatrix} \rho_1^h \cdot X_{T,1} + \dfrac{1-\rho_1^h}{1-\rho_1} n_1 \beta_1 \\ \rho_2^h \cdot X_{T,2} + \dfrac{1-\rho_2^h}{1-\rho_2} n_2 \beta_2 \end{bmatrix} \tag{37}$$

and

$$\mathrm{V}[\boldsymbol{X}_{T+h}|\boldsymbol{X}_T] = \begin{bmatrix} \rho_1^h(1-\rho_1^h)(1-2\beta_1/(1-\rho_1)) \cdot X_{T,1} + \dfrac{1-\rho_1^h}{1-\rho_1} n_1\beta_1 \left(1 - \dfrac{1-\rho_1^h}{1-\rho_1}\beta_1\right) \\ \rho_2^h(1-\rho_2^h)(1-2\beta_2/(1-\rho_2)) \cdot X_{T,2} + \dfrac{1-\rho_2^h}{1-\rho_2} n_2\beta_2 \left(1 - \dfrac{1-\rho_2^h}{1-\rho_2}\beta_2\right) \end{bmatrix}. \tag{38}$$

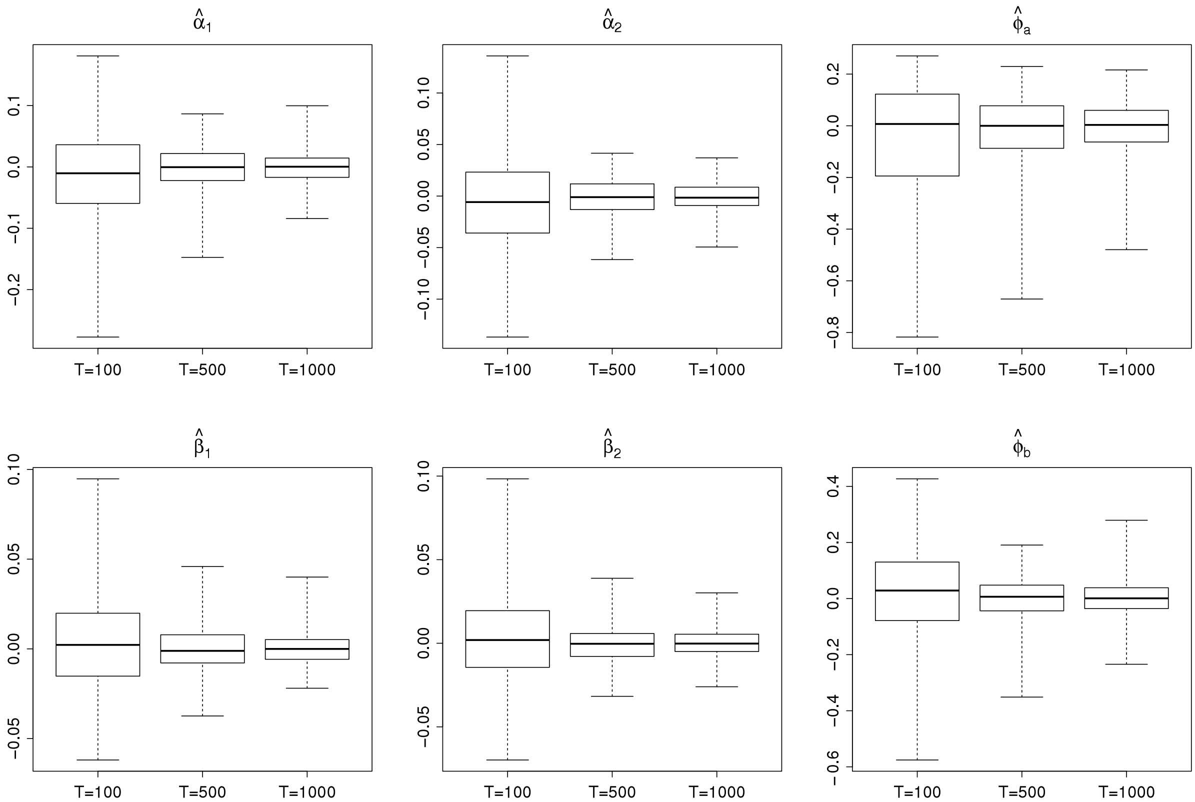

Fig. 4. Biases of the estimates for model M1.

Proof. Note that the conditional distribution of $\boldsymbol{X}_{T+h}$ given $\boldsymbol{X}_T$ satisfies the Chapman–Kolmogorov equations, i.e.

$$\begin{aligned} \mathrm{P}(\boldsymbol{X}_{t+h}=\boldsymbol{x}_{t+h}|\boldsymbol{X}_t=\boldsymbol{x}_t) &= \sum_{x_{t+h-1}} \mathrm{P}(\boldsymbol{X}_{t+h}=\boldsymbol{x}_{t+h}|\boldsymbol{X}_{t+h-1}=\boldsymbol{x}_{t+h-1})\cdot \mathrm{P}(\boldsymbol{X}_{t+h-1}=\boldsymbol{x}_{t+h-1}|\boldsymbol{X}_t=\boldsymbol{x}_t) \\ &= \boldsymbol{Q}^h, \end{aligned}$$

where the last equality is justified by stationarity. Moreover, in proving (37) the h-step-ahead conditional expectation

$$\begin{aligned} \mathrm{E}[\boldsymbol{X}_{T+h}|\boldsymbol{X}_T] &= \mathrm{E}[\boldsymbol{\alpha}\otimes\boldsymbol{X}_{T+h-1}+\boldsymbol{\beta}\otimes(\boldsymbol{n}-\boldsymbol{X}_{T+h-1})|\boldsymbol{X}_T] \\ &= \begin{bmatrix} \alpha_1\cdot \mathrm{E}(X_{T+h-1,1}|\boldsymbol{X}_T)+\beta_1 n_1-\beta_1\cdot \mathrm{E}(X_{T+h-1,1}|\boldsymbol{X}_T) \\ \alpha_2\cdot \mathrm{E}(X_{T+h-1,2}|\boldsymbol{X}_T)+\beta_2 n_2-\beta_2\cdot \mathrm{E}(X_{T+h-1,2}|\boldsymbol{X}_T) \end{bmatrix} \\ &= \begin{bmatrix} \alpha_1\rho_1^{h-1}\cdot X_{T,1}+\beta_1\rho_1^h(n_1-X_{T,1})+\dfrac{1-\rho_1^h}{1-\rho_1}n_1\beta_1 \\ \alpha_2\rho_2^{h-1}\cdot X_{T,2}+\beta_2\rho_2^h(n_2-X_{T,2})+\dfrac{1-\rho_2^h}{1-\rho_2}n_2\beta_2 \end{bmatrix} \\ &= \begin{bmatrix} \rho_1^h\cdot X_{T,1}+\dfrac{1-\rho_1^h}{1-\rho_1}n_1\beta_1 \\ \rho_2^h\cdot X_{T,2}+\dfrac{1-\rho_2^h}{1-\rho_2}n_2\beta_2 \end{bmatrix}. \end{aligned}$$

The proof of (38) follows by similar arguments. We skip the details. □

The univariate counterparts of (37) and (38) can be found in [43]. Note that (32) is included in these formulae as the special case for $h = 1$.

Forecasting can be pursued within the setting of conditional expectations in which case the h-step ahead predictor yielding minimum mean squared error forecasts is $\hat{\boldsymbol{X}}_{T+h} := \mathrm{E}[\boldsymbol{X}_{T+h}|\boldsymbol{X}_T]$. It is worth to mention that such an approach cannot be directly applied in the $\mathrm{BVB_{II}}$-INARCH(1) case since no closed-expressions for the conditional moments of the h-step-ahead conditional distribution are known. However, from the distribution in (36) such moments can be estimated numerically. The

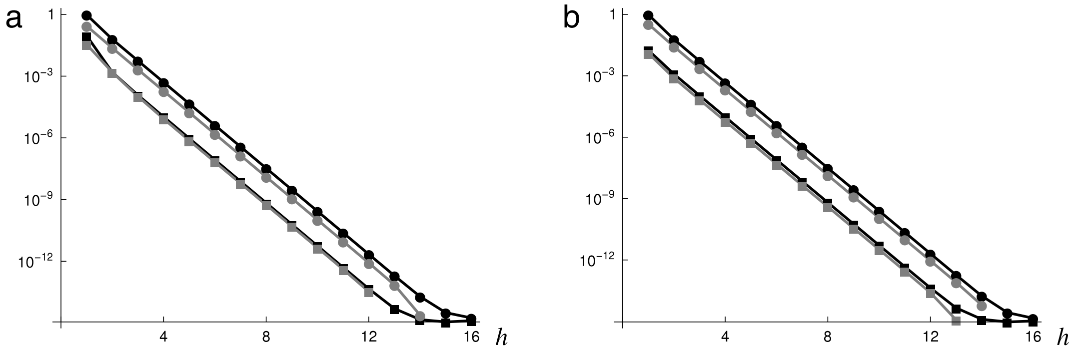

Fig. 5. Kullback–Leibler divergences, $d_{\mathrm{KL}}\left(\mathrm{P}_{\mathrm{marg}}, \mathrm{P}_{\mathrm{cond}}(h|\boldsymbol{y})\right)$, between the stationary marginal distribution and the lag-h conditional distributions for the $\mathrm{BVB_{II}}$-AR(1) model (a) and the $\mathrm{BVB_{II}}$-INARCH(1) model (b) from Section 5.1. Black/gray lines represent models with negative and positive correlation parameter ϕ, conditioned on $\boldsymbol{y} = (0, 0)$ (circles) or $\boldsymbol{y} = (3, 3)$ (squares), respectively.

major drawback of forecasting based on the conditional expectation is that it hardly produces *coherent* (i.e., integer-valued) predictions. In order to generate data coherent predictions, the median of the distribution in (36) (which minimizes the expected absolute error) or its corresponding mode can be employed as a point forecast; see [28,16] for details. Furthermore, prediction intervals can be obtained by taking advantage of the asymptotic normality of the conditional maximum likelihood estimators and the use of the δ-method applied to $h(\hat{\boldsymbol{\theta}}) := \mathrm{P}(\boldsymbol{X}_{T+h} = \boldsymbol{x}_{T+h}|\boldsymbol{X}_T = \boldsymbol{x}_T; \hat{\boldsymbol{\theta}})$.

Remark 7.2. As argued in Sections 3 and 5, both types models are ergodic (and finite-state) Markov chains, so we know from Section 4.2 in [35] that the rate of convergence in $\mathrm{P}(\boldsymbol{X}_{T+h} = \boldsymbol{x}|\boldsymbol{X}_T = \boldsymbol{y}) \to \mathrm{P}(\boldsymbol{X}_t = \boldsymbol{x})$ is geometric. For the $\mathrm{BVB_{II}}$-AR(1) model, this geometric rate also becomes obvious from formulae (37) and (38) for conditional moments, where the mean forecast decays geometrically from $[X_{T,1},\ X_{T,2}]'$ towards $[n_1\beta_1/(1-\rho_1),\ n_2\beta_2/(1-\rho_2)]'$ as the forecast horizon h increases, whereas the variance converges geometrically to the fixed values $[n_1\beta_1(1-\beta_1/(1-\rho_1))/(1-\rho_1),\ n_2\beta_2(1-\beta_2/(1-\rho_2))/(1-\rho_2)]'$. As an illustration, let us look back to the models considered in Section 5.1. We computed the lag-h conditional distributions, conditioned on either $\boldsymbol{y} = (0, 0)'$ or $\boldsymbol{y} = (3, 3)'$, and compared them to their respective stationary marginal distribution in terms of the Kullback–Leibler divergence,

$$d_{\mathrm{KL}}\left(\mathrm{P}_{\mathrm{marg}}, \mathrm{P}_{\mathrm{cond}}(h|\boldsymbol{y})\right) := \sum_{\boldsymbol{x}} \mathrm{P}(\boldsymbol{X}_t = \boldsymbol{x})\ \ln\left(\frac{\mathrm{P}(\boldsymbol{X}_t = \boldsymbol{x})}{\mathrm{P}(\boldsymbol{X}_{T+h} = \boldsymbol{x}|\boldsymbol{X}_T = \boldsymbol{y})}\right).$$

The result is shown in Fig. 5 where the geometric rate of convergence is obvious. While the choice of the condition $\boldsymbol{y}$ certainly has a strong effect on the actual divergence value, it is interesting to note that the shape of the distribution (positive or negative cross-correlation) is nearly without effect on the divergence values.

8. An application to rainy days time series

In this section, the results given above are applied to the analysis of the number of rainy days per week at several locations in Germany, collected from the year 2000 to 2010 by the German Weather Service (DWD = "Deutscher WetterDienst", http://www.dwd.de/). The DWD online database offers (among others) daily data from 44 measuring stations about a number of climate measurements (temperature, wind speed, precipitation, etc.). From these data, the number of rainy days per week and per station was obtained after some data pre-processing by the authors; each count data time series $(x_{t,i})_{t=1,\ldots,T}$ has the range $\{0, \ldots, n\}$ with $n = 7$ and is of length $T = 574$. A similar application, but related to Key West in Florida with its tropical savanna climate, was discussed by Cui and Lund [10].

Generally, the 44 time series exhibit similar characteristics, namely, a mean rate of rainy days around 0.5 and a significant empirical ACF(1) value close to 0.15. Accordingly, any pair of time series showed a strong positive cross-correlation, with the values typically ranging between 0.5 and 0.8. For illustrative purposes, let us now consider the time series of Bremen and Cuxhaven, which we denote by $(x_{t,1})_{t=1,\ldots,T}$ and $(x_{t,2})_{t=1,\ldots,T}$, respectively. Both cities are located close to the North Sea and exhibit an oceanic climate. The corresponding time series are displayed in Fig. 6. The mean number of rainy days per week is $\bar{x}_1 \approx 3.65$ and $\bar{x}_2 \approx 3.84$ with variances $s_1^2 \approx 3.99$ and $s_2^2 \approx 3.88$ in Bremen and Cuxhaven, respectively. Empirical ACFs and cross-correlation of both time series are plotted in Figs. 7 and 8, respectively. Both ACFs show a first-order autocorrelation around 0.15–0.20, indicating that a first-order model should be adequate to model the serial (weak) dependence in both cases. Furthermore, the correlation between the two time series is around 0.83, clearly indicating strong dependence between the series.

Next, we fitted both the $\mathrm{BVB_{II}}$-INARCH(1) and $\mathrm{BVB_{II}}$-AR(1) models to the data. CML-estimates and their corresponding standard errors are shown in Table 4. For the $\mathrm{BVB_{II}}$-AR(1) model both types of parametrization (α_i, β_i or π_i, ρ_i) were considered. Looking at the values of the information criteria AIC and BIC the $\mathrm{BVB_{II}}$-INARCH(1) model (with only 5 instead of

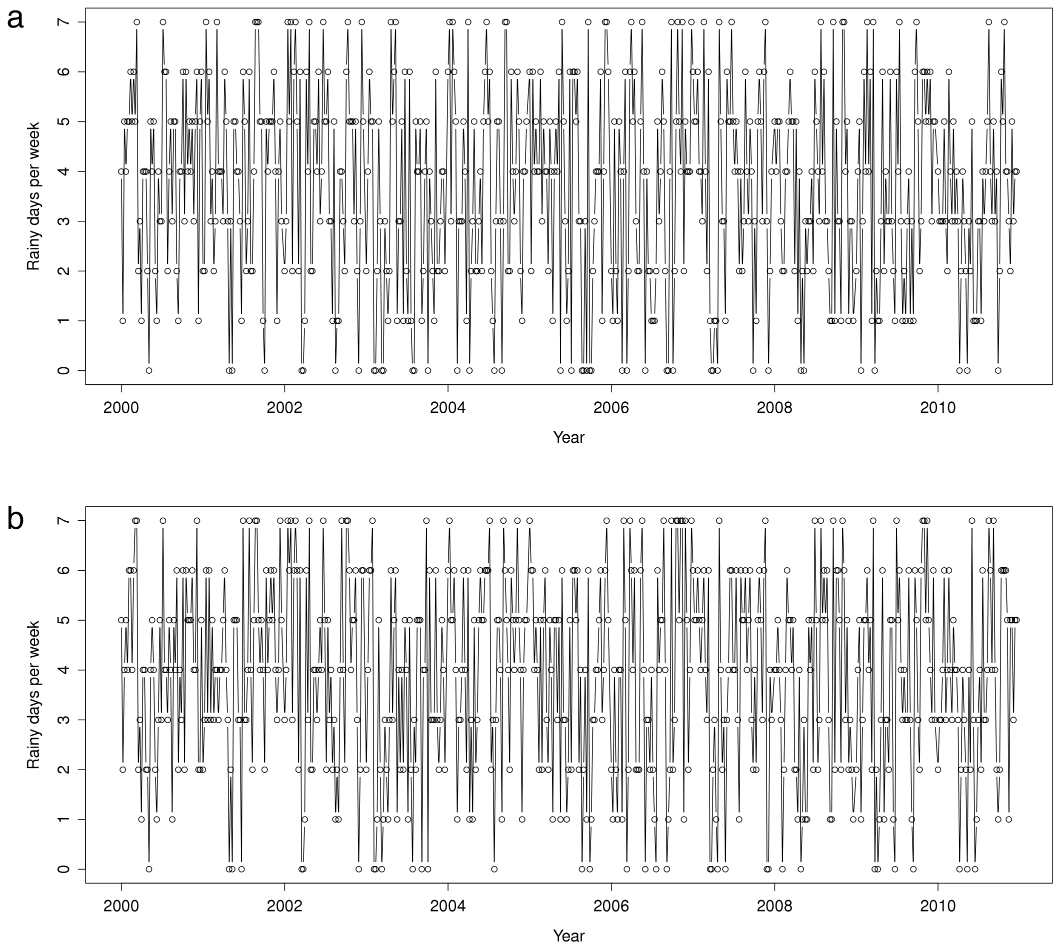

Fig. 6. Time series of number of weekly rainy days from 2000 to 2010 in (a) Bremen and (b) Cuxhaven (Germany).

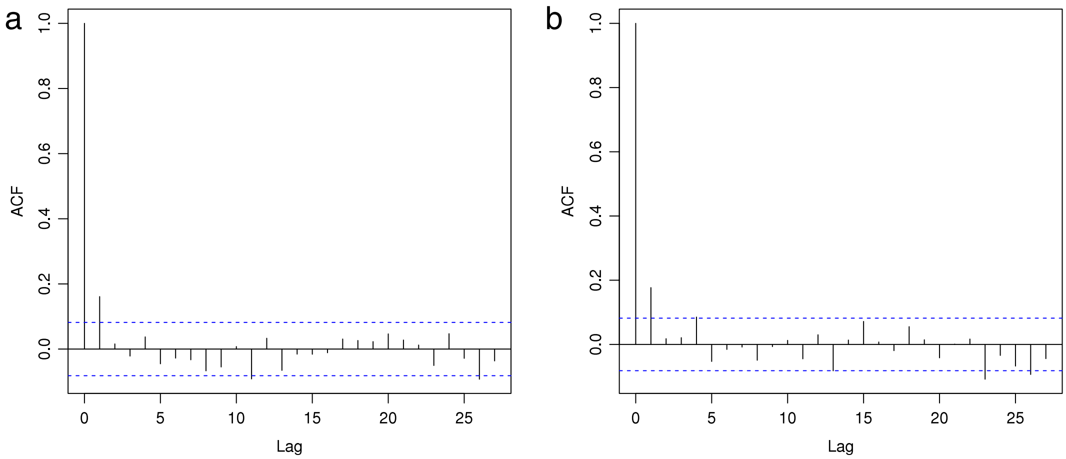

Fig. 7. ACF for the time series in (a) Bremen and (b) Cuxhaven.

6 parameters) appears preferable. This preference is also supported by the fact that the values for ACF(1) within the fitted INARCH model are more close to the respective empirical ACF(1) value than those within the fitted AR model ($\hat{\alpha}_{1,i}$ vs. $\hat{\rho}_i$). A similar conclusion holds with regard to the cross-correlations: 0.696 within fitted BVB_{II}-INARCH(1) model versus 0.672 within fitted BVB_{II}-AR(1) model. Finally, from the values for empirical mean and variance above it becomes clear that both marginal time series exhibit empirical extra-binomial variation, a feature being ignored by the BVB_{II}-AR(1) model.

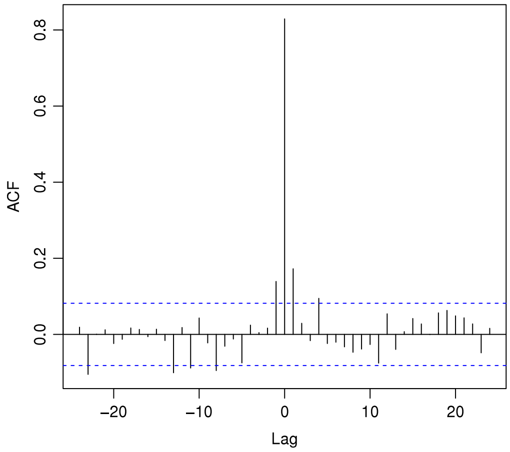

Fig. 8. Cross-correlation between Bremen and Cuxhaven.

Table 4
Parameter estimates for the parameters of the BVB_{II}-INARCH(1) model and BVB_{II}-AR(1) model. Standard errors in parentheses.

Bremen		Cuxhaven		Cross-corr.	
BVB_{II}-INARCH(1) model, AIC $\approx$ 4414, BIC $\approx$ 4435					
$\hat{\alpha}_{0,1}$	$\hat{\alpha}_{1,1}$	$\hat{\alpha}_{0,2}$	$\hat{\alpha}_{1,2}$	$\hat{\phi}$	
0.447 (0.015)	0.141 (0.023)	0.473 (0.016)	0.138 (0.024)	0.696 (0.015)	
BVB_{II}-AR(1) model, AIC $\approx$ 4456, BIC $\approx$ 4482					
$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\alpha}_2$	$\hat{\beta}_2$	$\hat{\phi}_\alpha$	$\hat{\phi}_\beta$
0.574 (0.013)	0.462 (0.014)	0.596 (0.013)	0.491 (0.014)	0.735 (0.028)	0.780 (0.028)
$\hat{\pi}_1$	$\hat{\rho}_1$	$\hat{\pi}_2$	$\hat{\rho}_2$	$\hat{\phi}_\alpha$	$\hat{\phi}_\beta$
0.520 (0.009)	0.113 (0.022)	0.549 (0.009)	0.105 (0.022)	0.735 (0.028)	0.780 (0.028)

Remark 8.1. The ACF plots in Fig. 7 indicate a first-order autocorrelation structure. Looking for further alternatives to model our data, one may first think of using another bivariate autoregressive model for counts. However, we are concerned with a finite range here, being $\{0, \ldots, 7\}^2$. As far as we are aware, there are no other autoregressive models for bivariate time series of counts with a finite range. Certainly, an obvious alternative would be to use a non-parametric Markov model (i.e., where each transition probability is estimated directly by the corresponding conditional frequency). However, Markov chains tend to be over-parameterized for practical purposes. In fact, considering that our state space consists of the $8^2 = 64$ pairs from $\{0, \ldots, 7\}^2$, such a model would have an extremely large number of parameters. Note that even for the first-order model we would have to estimate 4032 transition probabilities. As a consequence, we get much larger values for AIC and BIC, AIC $\approx$ 10 880 and BIC $\approx$ 28 430, than the corresponding values for the BVB_{II}-INARCH(1) and BVB_{II}-AR(1) models in Table 4.

9. Discussion

The bivariate binomial distribution of Type II constitutes a powerful device for developing models for bivariate time series of counts. This was exemplified in the present paper by our novel BVB_{II}-INARCH(1) model (utilizing the BVB_{II}-distribution as a conditional distribution) and our novel BVB_{II}-AR(1) model (using the probabilistic operation of BVB_{II}-thinning); both models are applied to stationary bivariate count data processes with a finite and time-independent range. But our BVB_{II}-approach is certainly not limited to this scenario. As an example, our novel bivariate binomial thinning operation (25) can also be used to extend the univariate INAR(1) model (1) for the analysis of bivariate time series with an *infinite* range of counts

$$\boldsymbol{X}_t = \boldsymbol{\alpha} \otimes \boldsymbol{X}_{t-1} + \boldsymbol{\varepsilon}_t, \quad t \in \mathbb{N}_0. \tag{39}$$

Such a BVB_{II}-INAR(1) model would be attractive for several reasons: first, it includes the bivariate models of Karlis and Pedeli [21] and Pedeli and Karlis [30,31,29] as a special case for $\phi_\alpha = 0$; see the discussion in Section 4. Secondly, in analogy to the case of the BVB_{II}-AR(1) model, also the marginals $(X_{t,i})$ of the BVB_{II}-INAR(1) model (39) are distributed like usual

INAR(1) processes. Finally, $(\boldsymbol{X}_t)$ constitutes a Markov chain with transition probabilities

$$p(\boldsymbol{x}|\boldsymbol{y}) = \sum_{a_1=0}^{\min(x_1,y_1)} \sum_{a_2=0}^{\min(x_2,y_2)} p_{(y_1,y_2;\alpha_1,\alpha_2,\phi_\alpha)}(a_1,a_2) \cdot \mathrm{P}\big(\boldsymbol{\varepsilon}_t = \boldsymbol{x} - \boldsymbol{a}\big) \tag{40}$$

and with an AR(1)-like autocorrelation structure, where the cross-covariance

$$\mathrm{Cov}(X_{t,1}, X_{t,2}) = \frac{\mathrm{Cov}(\varepsilon_{t,1}, \varepsilon_{t,2}) + \phi_\alpha \sqrt{\alpha_1\alpha_2(1-\alpha_1)(1-\alpha_2)} \cdot \mathrm{E}\big(\min(X_{t,1}, X_{t,2})\big)}{1-\alpha_1\alpha_2}, \tag{41}$$

might be both positive or negative (even if the innovations' components are independent). All these features make the $\mathrm{BVB}_{\mathrm{II}}$-INAR(1) model an attractive issue for future research. Besides establishing conditions to ensure the existence of a stationary process satisfying (39), also parameter estimation and forecasting in connection with this family of models are still open problems.

Beyond this adaption to model bivariate time series with an infinite range of counts, there are a number of further possibilities for future research in this area. In order to make the bivariate INAR-type models more flexible with respect to real data applications, it may be of interest to include explanatory covariates in the model to account for dependence through the thinning operations on several factors. Another possible extension is to allow $\boldsymbol{n}$ to vary in time as a realization of a sequence of i.i.d. random variables or a realization of some Markov process with support on $\mathbb{N}$. This approach should allow the $\mathrm{BVB}_{\mathrm{II}}$-AR(1) model to be more attractive for economic applications [8]. Also the $\mathrm{BVB}_{\mathrm{II}}$-INARCH model in (16) can be generalized in a number of ways, namely by considering $\mathrm{BVB}_{\mathrm{II}}$-INARCH models of higher-order with covariates, or by assuming other bivariate discrete distributions such as the bivariate negative binomial. Within the framework of the $\mathrm{BVB}_{\mathrm{II}}$-INARCH and the $\mathrm{BVB}_{\mathrm{II}}$-AR model, extensions to the multivariate case for modeling more than two correlated time series of count data should also be highly desirable.

Acknowledgments

The authors thank the referee for carefully reading the article and for the comments, which greatly improved the article. This research was partially supported by FEDER funds through COMPETE Operational Programme Factors of Competitiveness ("Programa Operacional Factores de Competitividade") and by Portuguese funds through the Center for Research and Development in Mathematics and Applications and the Portuguese Foundation for Science and Technology ("FCT-Fundação para a Ciência e a Tecnologia"), within project PEst-C/MAT/UI4106/2011 with COMPETE number FCOMP-01-0124-FEDER-022690.

References

[1] M. Alosh, The impact of missing data in a generalized integer-valued autoregression model for count data, J. Biopharm. Statist. 19 (2009) 1039–1054.
[2] A.A. Alzaid, M.A. Al-Osh, An integer-valued *p*th-order autoregressive structure (INAR(*p*)) process, J. Appl. Probab. 27 (1990) 314–324.
[3] J. Anderson, D. Karlis, Treating missing values in INAR(1) models: an application to syndromic survillance data, J. Time Ser. Anal. 31 (2010) 12–19.
[4] P. Billingsley, Statistical Inference for Markov Processes, University of Chicago Press, 1961.
[5] A. Biswas, J.S. Hwang, A new bivariate binomial distribution, Statist. Probab. Lett. 60 (2002) 231–240.
[6] J.R. Blum, D.L. Hanson, L.H. Koopmans, On the strong law of large numbers for a class of stochastic processes, Z. Wahrscheinlichkeitstheor. 2 (1963) 1–11.
[7] M. Boudreault, A. Charpentier, Multivariate integer-valued autoregressive models applied to earthquake counts, 2011, submitted for publication.
[8] K. Brännäs, J. Nordström, Tourist accommodation effects of festivals, Tourism Econ. 12 (2006) 291–302.
[9] K. Brännäs, A.M.M.S. Quoreshi, Integer-valued moving average modelling of the number of transactions in stocks, Appl. Financ. Econ. 20 (2010) 1429–1440.
[10] Y. Cui, R. Lund, A new look at time series of counts, Biometrika 96 (2009) 781–792.
[11] Y. Cui, R. Lund, Inference in binomial AR(1) models, Statist. Probab. Lett. 80 (2010) 1985–1990.
[12] R. Ferland, A. Latour, D. Oraichi, Integer-valued GARCH processes, J. Time Ser. Anal. 27 (2006) 923–942.
[13] K. Fokianos, A. Rahbek, D. Tjøstheim, Poisson autoregression, J. Amer. Statist. Assoc. 104 (2009) 1430–1439.
[14] K. Fokianos, D. Tjøstheim, Nonlinear Poisson autoregression, Ann. Inst. Statist. Math. 64 (2012) 1205–1225.
[15] J. Franke, T.S. Rao, Multivariate First-Order Integer-Valued Autoregression, Technical Report, Universität Kaiserslautern, 1993.
[16] R.K. Freeland, B.P.M. McCabe, Forecasting discrete valued low count time series, Int. J. Forecast. 20 (2004) 427–434.
[17] A. Heinen, Modeling time series count data: an autoregressive conditional Poisson model, CORE Discussion Paper 2003/62, 2003.
[18] A. Heinen, E. Rengifo, Multivariate autoregressive modeling of time series count data using copulas, J. Empir. Financ. 14 (2007) 564–583.
[19] N.L. Johnson, S. Kotz, N. Balakrishnan, Discrete Multivariate Distributions, John Wiley & Sons, Inc., New York, 1997.
[20] R.C. Jung, A.R. Tremayne, Useful models for time series of counts or simply wrong ones? Adv. Stat. Anal. 95 (2011) 59–91.
[21] D. Karlis, X. Pedeli, Flexible bivariate INAR(1) processes using copulas, Comm. Statist. Theory Methods 42 (2013) 723–740.
[22] S. Kocherlakota, K. Kocherlakota, Bivariate Discrete Distributions, Marcel Dekker, Inc., 1992.
[23] A. Latour, The multivariate GINAR(*p*) process, Adv. Appl. Probab. 29 (1997) 228–248.
[24] H. Liu, Some models for time series of counts, Ph.D. Thesis, Columbia University, 2012.
[25] A.W. Marshall, I. Olkin, A family of bivariate distributions generated by the bivariate Bernoulli distribution, J. Amer. Statist. Assoc. 80 (1985) 332–338.
[26] E. McKenzie, Some simple models for discrete variate time series, Water Resour. Bull. 21 (1985) 645–650.
[27] D. Morina, P. Puig, J. Ríos, A. Vilella, A. Trilla, A statistical model for hospital admissions caused by seasonal diseases, Stat. Med. 30 (2011) 3125–3136.
[28] H. Pavlopoulos, D. Karlis, INAR(1) modeling of overdispersed count series with an environmental application, Environmetrics 19 (2008) 369–393.
[29] X. Pedeli, D. Karlis, A bivariate INAR(1) process with application, Stat. Model. 11 (2011) 325–349.
[30] X. Pedeli, D. Karlis, On composite likelihood estimation of a multivariate INAR(1) model, J. Time Ser. Anal. 34 (2013) 206–220.
[31] X. Pedeli, D. Karlis, On estimation of the bivariate Poisson INAR process, Comm. Statist. Simulation Comput. 42 (2013) 514–533.
[32] A.M.M.S. Quoreshi, Bivariate time series modelling of financial count data, Comm. Statist. Theory Methods 35 (2006) 1343–1358.

[33] M.M. Ristic, A.S. Nastic, K. Jayakumar, H.S. Bakouch, A bivariate INAR(1) time series model with geometric marginals, Appl. Math. Lett. 25 (2012) 481–485.
[34] A.M. Schmidt, J.B.M. Pereira, Modelling time series of counts in epidemiology, Internat. Statist. Rev. 79 (2011) 48–69.
[35] E. Seneta, Non-Negative Matrices and Markov Chains, second ed., Springer-Verlag, New York, Heidelberg, Berlin, 1981.
[36] F.W. Steutel, K. van Harn, Discrete analogues of self-decomposability and stability, Ann. Probab. 7 (1979) 893–899.
[37] D. Tjøstheim, Some recent theory for autoregressive count time series, TEST 21 (2012) 413–438.
[38] C.H. Weiß, A new class of autoregressive models for time series of binomial counts, Comm. Statist. Theory Methods 38 (2009) 447–460.
[39] C.H. Weiß, Modelling time series of counts with overdispersion, Stat. Methods Appl. 18 (2009) 507–519.
[40] C.H. Weiß, Monitoring correlated processes with binomial marginals, J. Appl. Stat. 36 (2009) 399–414.
[41] C.H. Weiß, H.Y. Kim, Binomial AR(1) processes: moments, cumulants, and estimation, Statistics 47 (2013) 494–510.
[42] C.H. Weiß, H.Y. Kim, Diagnosing and modelling extra-binomial variation for time-dependent counts, Appl. Stoch. Models Bus. Ind. (2013) in press.
[43] C.H. Weiß, P.K. Pollett, Chain binomial models and binomial autoregressive processes, Biometrics 68 (2012) 815–824.
[44] C.H. Weiß, P.K. Pollett, Binomial autoregressive processes with density dependent thinning, J. Time Ser. Anal. (2013) in press.
[45] J. Zhou, I.V. Basawa, Least-squared estimation for bifurcation autoregressive processes, Statist. Probab. Lett. 74 (2005) 77–88.
[46] F. Zhu, D. Wang, Diagnostic checking integer-valued ARCH(p) models using conditional residual autocorrelations, Comput. Statist. Data Anal. 54 (2010) 496–508.
[47] F. Zhu, D. Wang, Estimation and testing for a Poisson autoregressive model, Metrika 73 (2011) 211–230.