

Modelling overdispersion with integer-valued moving average processes*

Maria Eduarda Silva and Isabel Silva and Cristina Torres

Abstract A new first-order integer-valued moving average, INMA(1), model based on the negative binomial thinning operation defined by Ristić et al. [21] is proposed and characterized. It is shown that this model has negative binomial (NB) marginal distribution when the innovations follow a NB distribution and therefore it can be used in situations where the data present overdispersion. Additionally, this model is extended to the bivariate context. The Generalized Method of Moments (GMM) is used to estimate the unknown parameters of the proposed models and the results of a simulation study that intends to investigate the performance of the method show that, in general, the estimates are consistent and symmetric. Finally, the proposed model is fitted to a real dataset and the quality of the adjustment is evaluated.

Key words: Count time series, Overdispersion, Bivariate time series, INAR models, Negative binomial

* Accepted authors manuscript (AAM) published in Stochastic Models, Statistics and Their Applications - SMSA 2019. Springer Proceedings in Mathematics & Statistics, vol 294. DOI: 10.1007/978-3-030-28665-1_21. Final source webpage: https://doi.org/10.1007/978-3-030-28665-1_21

Maria Eduarda Silva
Faculdade de Economia, Universidade do Porto and CIDMA, Portugal e-mail: mesilva@fep.up.pt

Isabel Silva
Faculdade de Engenharia, Universidade do Porto and CIDMA, Portugal e-mail: ims@fe.up.pt

Cristina Torres
ISCAP-IPP, Portugal e-mail: cmptorres@gmail.com

1 Introduction

In the last decades there has been a growing interest in studying non-negative integer-valued time series and, in particular, time series of counts. This kind of data is characterized by low counts, asymmetric distributions, excess of zeros, overdispersion, etc, precluding Gaussian approximations and requiring different modelling strategies. A comprehensive account of models for counting processes can be found in [27] and references therein. One of the approaches to construct models for time series of counts considers a random operation called thinning coupled with innovations following discrete distributions and constitutes the family of integer-valued autoregressive and moving average, INARMA. These models have been extensively studied in the literature, especially with regard to integer-valued autoregressive models (see [22]). On the other hand, integer-valued moving average (INMA) models have attracted less attention, mainly because likelihood-based inference is not straightforward in this context, since the conditional distribution functions are not easily obtained.

The INMA(q) models satisfy the following recursion

$$X_t = \beta_0 \circ_t \varepsilon_t + \beta_1 \circ_t \varepsilon_{t-1} + \cdots + \beta_q \circ_t \varepsilon_{t-q}, \quad t \in \mathbb{Z}, \quad (1)$$

where the innovation process, $\{\varepsilon_t\}$, is an independent and identically distributed (i.i.d.) process with support on \mathbb{N}_0 , finite mean $\mu_\varepsilon > 0$ and variance $\sigma_\varepsilon^2 > 0$; $\beta_0, \beta_1, \dots, \beta_q \in [0, 1]$, $\beta_q \neq 0$ (in general $\beta_0 = 1$) and " \circ_t " denotes the binomial thinning operator (proposed by [23]), defined as $\beta \circ X = \sum_{i=1}^X \xi_i(\beta)$, where $\{\xi_i(\beta)\}$ is a collection of i.i.d. Bernoulli counting random variables with parameter β , independent of X . The sequence $\{\xi_i(\beta) : i \in \mathbb{N}\}$ is designated by counting series and $\beta \circ X | X \sim Bi(X, \alpha)$. The subscript t in the operator ' \circ ' emphasizes the fact that the thinning operations are performed at each time t . The thinning operations $\beta_j \circ_t \varepsilon_{t-j}$, $j = 1, \dots, q$ may be considered independent as in [1] and [18] or not as in [3], giving rise to four different models that can be embedded into a single family of INMA(q) models, [26]. Under the assumption of Poisson innovations, $\{\varepsilon_t\} \sim \text{Po}(\lambda/(1 + \beta_1))$, the INMA(1) process is $\text{Po}(\lambda)$, [18]. Using an alternative parametrization, [1] proved that if $\{\varepsilon_t\}$ has a Poisson distribution with mean λ , then $\{X_t\}$ has Poisson distribution with parameter $\lambda(1 + \beta_1)$. The resulting process is time reversible in the sense that $\{X_1, \dots, X_k\}$ has the same joint distribution as $\{X_k, \dots, X_1\}$ for all $k \geq 2$. Recently, the INMA models based on the binomial thinning operation as defined in (1), have been extended to threshold INMA models, [31], INMA models with structural changes, [28] and Poisson combined INMA(q) models, [29]. Additionally, the INMA processes have been applied in the reinsurance context, namely on discrete risk models (see [5, 6, 10, 17, 30]).

Several modifications of the binomial thinning operator have been proposed in order to make the models more flexible, allowing more variability in the processes and consequently different properties. The case where the counting series are i.i.d. random variables with geometric distribution was analysed by [21], and referred to

as the negative binomial (NB) thinning operator. Thus, the NB thinning operator is defined as

$$\beta * X = \sum_{i=1}^X \omega_i(\beta), \quad (2)$$

where $\beta \in [0, 1[$ and the counting series $\{\omega_i(\beta)\}$ is a sequence of i.i.d. random variables with geometric distribution with parameter $\frac{\beta}{1+\beta}$, i.e., with probability mass function of the form $P(\omega_i(\beta) = k) = \frac{\beta^k}{(1+\beta)^{k+1}}$, $k \geq 0$.

In this work, a first-order INMA model with NB marginal, denoted by INMA-NB(1), based on the NB thinning operator is proposed and its properties are studied. One of the advantages of this model is that the ability to handle overdispersed datasets.

Often, the collected time series are multivariate in the sense that there are counts of several events observed over time and the counts at each time point are correlated. While research on univariate time series of counts continues to grow, the literature on bivariate and also multivariate time series models for count data is less developed. This might be explained by the fact that classical inference in multivariate count data models has proven to be analytically as well as computationally very demanding. In the last 20 years, some interesting attempts have been made but most of them do not arise in the context of thinning-based processes. The first attempt to develop multivariate INMA models came from [4], who proposed a Bivariate INMA(1) model for guest nights in hotels and cottages. Another model was proposed by [19] which extends the univariate INMA model of [18]. Recently, [15, 16, 24] proposed stationary and non-stationary BINMA models based on the binomial thinning operation and where the innovation series follow the bivariate Poisson distribution under time-varying moment assumptions and constant cross-correlations and used generalized quasi-likelihood estimation method.

In this work, the INMA-NB(1) model is extended to the bivariate case, with the cross-correlation generated through a bivariate negative binomial-type II (BNB_{II}) process.

The remainder of this work is organized as follows. The INMA-NB(1) model is defined and its properties are established in Section 2. Then, in Section 3, this model is extended to the bivariate case and its characterization is presented. Parameter estimation is accomplished by Generalized Method of Moments (GMM), and the finite sample behaviour of GMM are examined using Monte Carlo methods in Section 4. Furthermore, the proposed model is applied to real dataset in Section 5. Finally, Section 6 concludes the paper.

2 First-Order Negative Binomial INteger-Valued Moving Average Model

Let $\{X_t\}, t \in \mathbb{Z}$, be a sequence of non-negative integer-valued random variables. Then $\{X_t\}$ is called a first-order negative binomial integer-valued moving average, INMA-NB(1), model if it satisfies the equation

$$X_t = \varepsilon_t + \beta * \varepsilon_{t-1}, \quad t \in \mathbb{Z}, \quad (3)$$

where $\beta \in]0, 1[$; '*' is the NB thinning operator defined by (2) and $\{\varepsilon_t\}$ is a sequence of i.i.d. non-negative random variables following a NB distribution, $\varepsilon_t \sim \text{NB}\left(\kappa, \frac{1}{1+\beta}\right)$, with $E[\varepsilon_t] = \kappa\beta$, $\text{Var}[\varepsilon_t] = \kappa\beta(1+\beta)$ and probability generating function (pgf) given by $G_\varepsilon(s) = \frac{1}{(1-\beta(s-1))^\kappa}$ (see [12]). Note that, conditioned on ε_{t-1} , $\beta * \varepsilon_{t-1}$ is NB distributed, i.e., $\beta * \varepsilon_{t-1} | \varepsilon_{t-1} \sim \text{NB}\left(\varepsilon_{t-1}, \frac{1}{1+\beta}\right)$.

Proposition 1. Let $G_X(s)$ and $G_\varepsilon(s)$ denote the pgf of the random variables $\{X_t\}$ and $\{\varepsilon_t\}$, respectively. Then the pgf of the INMA-NB(1) model is given by

$$G_X(s) = G_\varepsilon(s) G_\varepsilon\left(\frac{1}{1-\beta(s-1)}\right) = \left(\frac{1}{1-(\beta+\beta^2)(s-1)}\right)^\kappa$$

and therefore $\{X_t\} \sim \text{NB}\left(\kappa, \frac{1}{1+\beta(1+\beta)}\right)$.

Proof. Let $z = \beta * \varepsilon_{t-1} | \varepsilon_{t-1} \sim \text{NB}\left(\varepsilon_{t-1}, \frac{1}{1+\beta}\right)$, then it is possible to write $G_z(s) = \left(\frac{1}{1-\beta(s-1)}\right)^{\varepsilon_{t-1}}$. Thus, it can be written that

$$\begin{aligned} G_X(s) &= E[s^{\varepsilon_t + \beta * \varepsilon_{t-1}}] = G_\varepsilon(s) E[E[s^{\beta * \varepsilon_{t-1}} | \varepsilon_{t-1}]] = G_\varepsilon(s) E\left[\left(\frac{1}{1-\beta(s-1)}\right)^{\varepsilon_{t-1}}\right] \\ &= G_\varepsilon(s) G_\varepsilon\left(\frac{1}{1-\beta(s-1)}\right) = \frac{1}{(1-\beta(s-1))^\kappa} \left[\frac{1}{1-\beta\left(\frac{1}{1-\beta(s-1)}-1\right)}\right]^\kappa \\ &= \left(\frac{1}{1-(\beta+\beta^2)(s-1)}\right)^\kappa, \end{aligned}$$

hence, $X_t \sim \text{NB}(\kappa, (\beta + \beta^2))$.

Using a different parametrization of the NB distribution (see [12] for details), considering $\kappa = k$, $\beta + \beta^2 = p$, and $p = \frac{1}{1+p} = \frac{1}{1+\beta(1+\beta)}$, $q = 1 - p = \frac{\beta(1+\beta)}{1+\beta(1+\beta)}$, then it is possible to write that $X_t \sim \text{NB}\left(\kappa, \frac{1}{1+\beta(1+\beta)}\right)$.

By using the properties of the NB thinning operator given in Lemma 3 of [21], namely $E[\beta * X] = \beta E[X]$ and $E(\beta * X)^2 = \beta^2 E(X^2) + \beta(1+\beta)E(X)$, and stationarity of the process, it can be easily proven that the first- and second-order moments of INMA-NB(1) model defined in (3) are given by

$$\begin{aligned}
E[X_t] &= \kappa\beta(1+\beta); \\
\text{Var}[X_t] &= \kappa\beta(1+\beta)(1+\beta(1+\beta)); \\
\gamma_{X_t}(k) = \text{Cov}(X_{t-k}, X_t) &= \begin{cases} \kappa\beta^2(1+\beta), & k=1, \\ 0, & k>1. \end{cases}
\end{aligned} \tag{4}$$

Note that the index of dispersion of the process $\{X_t\}$ defined in (3) is given by $\frac{\sigma^2}{\mu} = 1 + \beta + \beta^2 > 1$ and thus this model is appropriate for overdispersed data. Motivated by this fact, the INMA-NB(1) model is extended to the bivariate case in the next section.

3 Bivariate INMA Models Based on the NB Thinning Operator

It is well known that the generalization of discrete distributions to a multivariate context is neither straightforward nor unique leading to multivariate distributions with different properties. Here, the Bivariate Negative Binomial distribution is constructed based on the compound correlated bivariate Poisson distribution proposed by [7] and designated by Bivariate Negative Binomial-type II (BNB_{II}) distribution (see [7, 14] for details). Let Y_1, Y_2, Y_0 and Λ denote random variables such that $Y_i|\Lambda = \lambda \sim \text{Po}(\theta_i\lambda)$, $i = 1, 2, 0$ are mutually independent. If Λ is Gamma then the joint (unconditional) distribution of $X_1 = Y_1 + Y_0$ and $X_2 = Y_2 + Y_0$ has pgf given by

$$G_{\mathbf{X}}(s_1, s_2) = \left(\frac{q}{1 - p_1 s_1 - p_2 s_2 - p_3 s_1 s_2} \right)^{\nu} \tag{5}$$

where $p_i = \begin{cases} \theta_i/(\theta_1 + \theta_2 + \theta_0 + \beta), & i = 1, 2 \\ \theta_0/(\theta_1 + \theta_2 + \theta_0 + \beta), & i = 3 \end{cases}$ and $q = 1 - (p_1 + p_2 + p_3) = \beta/(\theta_1 + \theta_2 + \theta_0 + \beta)$. Denote $\mathbf{X} = (X_1, X_2) \sim \text{BNB}_{\text{II}}(p_1, p_2, p_3, \nu)$.

Now, let $\{\mathbf{X}_t\} = \{(X_{1,t}, X_{2,t})\}, t \in \mathbb{Z}$, be a non-negative integer-valued bivariate random variable. Then $\{\mathbf{X}_t\}$ is a first-order Bivariate INteger-valued Moving Average model, referred as BINMA-BNB_{II}(1, 1), if satisfies the following equations

$$\begin{cases} X_{1,t} = \varepsilon_{1,t} + \beta_1 * \varepsilon_{1,t-1}, \\ X_{2,t} = \varepsilon_{2,t} + \beta_2 * \varepsilon_{2,t-1}, \end{cases} \tag{6}$$

where $\beta_j \in]0, 1[, j = 1, 2$; $*$ is the NB thinning operator given in (2) and the innovation process $\{\varepsilon_t\} = \{(\varepsilon_{1,t}, \varepsilon_{2,t})\}, t \in \mathbb{Z}$, is an i.i.d. sequence of bivariate random variables that follows a BNB_{II} distribution with parameters (p_1, p_2, p_3, ν)

where $p_1 = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \phi + \tau}$, $p_2 = \frac{\lambda_2}{\lambda_1 + \lambda_2 + \phi + \tau}$ and $p_3 = \frac{\phi}{\lambda_1 + \lambda_2 + \phi + \tau}$, with $\nu, \tau, \lambda_1, \lambda_2 > 0$ and $\phi \in [0, \min(\lambda_1, \lambda_2)[$.

As referred by [7, 14], the marginal distribution of $\varepsilon_{j,t}$ is univariate NB, i.e., $\varepsilon_{j,t} \sim \text{NB}\left(v, \frac{\tau}{(\lambda_j + \phi) + \tau}\right)$, $j = 1, 2$, with expected value and variance given by (for $j = 1, 2$) $E[\varepsilon_{j,t}] = (\lambda_j + \phi) \frac{v}{\tau}$ and $\text{Var}[\varepsilon_{j,t}] = (\lambda_j + \phi)^2 \frac{v}{\tau^2} + (\lambda_j + \phi) \frac{v}{\tau}$, respectively, and covariance defined as $\text{Cov}(\varepsilon_{1,t}, \varepsilon_{2,t}) = (\lambda_1 + \phi)(\lambda_2 + \phi) \frac{v}{\tau^2} + \phi \frac{v}{\tau}$.

It can be proved (for details see [25]) that the first- and second-order moments of the BINMA-BNB $_{\text{II}}(1, 1)$ model are given by (for $j = 1, 2$)

$$\begin{aligned} E[X_{j,t}] &= (1 + \beta_j)(\lambda_j + \phi) \frac{v}{\tau}; \\ \text{Var}[X_{j,t}] &= (1 + \beta_j^2)(\lambda_j + \phi)^2 \frac{v}{\tau^2} + (2\beta_j^2 + \beta_j + 1)(\lambda_j + \phi) \frac{v}{\tau}; \\ \gamma_{X_j}(k) &= \text{Cov}(X_{j,t-k}, X_{j,t}) = \begin{cases} \beta_j(\lambda_j + \phi) \frac{v}{\tau} \left((\lambda_j + \phi) \frac{1}{\tau} + 1 \right), & k = 1, \\ 0, & k > 1; \end{cases} \\ \gamma_{X_1, X_2}(0) &= \text{Cov}(X_{1,t}, X_{2,t}) = (1 + \beta_1 \beta_2) \left((\lambda_1 + \phi)(\lambda_2 + \phi) \frac{v}{\tau^2} + \phi \frac{v}{\tau} \right); \\ \gamma_{X_1, X_2}(1) &= \text{Cov}(X_{1,t}, X_{2,t-1}) = \beta_1 \left((\lambda_1 + \phi)(\lambda_2 + \phi) \frac{v}{\tau^2} + \phi \frac{v}{\tau} \right); \\ \gamma_{X_2, X_1}(1) &= \text{Cov}(X_{1,t-1}, X_{2,t}) = \beta_2 \left((\lambda_1 + \phi)(\lambda_2 + \phi) \frac{v}{\tau^2} + \phi \frac{v}{\tau} \right); \\ \gamma_{X_1, X_2}(k) &= \gamma_{X_2, X_1}(k) = 0, \text{ if } k > 1. \end{aligned} \quad (7)$$

Additionally, the joint pgf of $\mathbf{X}_t = \{(X_{1,t}, X_{2,t})\}$ is given by

$$G_{\mathbf{X}}(\mathbf{s}) = G_{\boldsymbol{\varepsilon}}(\mathbf{s}) G_{\boldsymbol{\varepsilon}} \left(\frac{\tau}{\tau - (\lambda_1 + \phi)(s_1 - 1)}, \frac{\tau}{\tau - (\lambda_2 + \phi)(s_2 - 1)} \right),$$

where $G_{\boldsymbol{\varepsilon}}(\mathbf{s}) = G_{(\varepsilon_{1,t}, \varepsilon_{2,t})}(s_1, s_2)$. Then, $G_{\mathbf{X}}(\mathbf{s})$ can be written by

$$\begin{aligned} G_{\mathbf{X}}(\mathbf{s}) &= \left[1 - \left(\frac{\lambda_1 + \phi}{\tau} \right) (s_1 - 1) - \left(\frac{\lambda_2 + \phi}{\tau} \right) (s_2 - 1) - \frac{\phi}{\tau} (s_1 - 1)(s_2 - 1) \right]^{-v} \times \\ &\quad \times \left[1 - \left(\frac{\lambda_1 + \phi}{\tau} \right) \left(\frac{\tau}{\tau - (\lambda_1 + \phi)(s_1 - 1)} - 1 \right) - \left(\frac{\lambda_2 + \phi}{\tau} \right) \times \right. \\ &\quad \times \left. \left(\frac{\tau}{\tau - (\lambda_2 + \phi)(s_2 - 1)} - 1 \right) - \frac{\phi}{\tau} \left(\frac{\tau}{\tau - (\lambda_1 + \phi)(s_1 - 1)} - 1 \right) \times \right. \\ &\quad \times \left. \left. \left(\frac{\tau}{\tau - (\lambda_2 + \phi)(s_2 - 1)} - 1 \right) \right]^{-v}. \end{aligned}$$

Furthermore, it can be proven (by contradiction) that the joint distribution of the BINMA-BNB $_{\text{II}}(1, 1)$ model it is not BNB $_{\text{II}}$ (details in [25]).

The index of dispersion of the BINMA-BNB $_{\text{II}}(1, 1)$ model is given by $\frac{\sigma_j^2}{\mu_j} = 1 + \beta_j + \beta_j^2 > 1$, with $\beta_j = \frac{\lambda_j + \phi}{\tau}$. Therefore, as for the univariate case, this model is appropriate when overdispersion occurs.

4 Parameter Estimation and Monte Carlo Simulation Results

In this section, the estimation of the parameters is discussed and the small sample properties are illustrated through the results of a small Monte Carlo experiment.

4.1 Generalized Method of Moments

It is well known that likelihood-based procedures are unfeasible in MA models, with exception of the Poisson INMA(1) model based on binomial thinning operation, where the conditional distribution $f(X_t|X_{t-1})$ of both the Poisson INAR(1) and the Poisson INMA(1) models are of the same form (see [18]). As referred by [24], in the INMA context, the full or conditional likelihood density function is arduous and renders considerable computational difficulties. Therefore, the Generalized Method of Moment (GMM) methodology based on first- and second-order moments of the process is considered. The GMM estimator was firstly introduced by [9] into the econometric literature and, since then, has been widely applied in several fields.

Suppose we have an observed sample $\mathbf{X}_n = \{X_t : t = 1, \dots, n\}$ from which we want to estimate an unknown $q \times 1$ parameter vector $\boldsymbol{\theta}$ with true value $\boldsymbol{\theta}_0$ and consider a vector $\mathbf{T}_n = \mathbf{T}_n(\mathbf{X}_n)$ of $k \geq q$ summary statistics with expectation $\boldsymbol{\alpha}(\boldsymbol{\theta}) = E[\mathbf{T}_n]$ (where $\boldsymbol{\alpha}(\boldsymbol{\theta})$ are the theoretical counterparts) under the model. The so called moment condition is defined by $E[\mathbf{h}_n(\boldsymbol{\theta}; \mathbf{X}_n)] = \mathbf{0}$, where $\mathbf{h}_n(\boldsymbol{\theta}; \mathbf{X}_n) = \mathbf{T}_n - \boldsymbol{\alpha}(\boldsymbol{\theta})$ is a continuous $k \times 1$ vector function of $\boldsymbol{\theta}$, and $E[\mathbf{h}_n(\boldsymbol{\theta}; \mathbf{X}_n)]$ exists and is finite for all t and $\boldsymbol{\theta}$. In practice, the moment condition is replaced by its sample analogous $\frac{1}{n} \sum_{t=1}^n \mathbf{h}_n(\boldsymbol{\theta}; X_t) = \mathbf{0}$, and an estimator $\hat{\boldsymbol{\theta}}$ can be obtained as the solution of the last equation.

Note that when $k = q$, we obtain the Method of Moments (MM) estimator and we say that $\boldsymbol{\theta}$ is just-identified. The Generalized Method of Moments estimator is obtained when $k > q$ and then we say that $\boldsymbol{\theta}$ is over-identified. Then, the GMM estimator of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_n = \arg \min_{\boldsymbol{\theta}} \{ \mathbf{h}_n(\boldsymbol{\theta}; \mathbf{X}_n)' \mathbf{W}_n \mathbf{h}_n(\boldsymbol{\theta}; \mathbf{X}_n) \},$$

where $[\cdot]'$ denotes transpose and \mathbf{W}_n is a $k \times k$ is any symmetric and positive definite weight matrix that may depend on the data but that converges in probability to a positive definite matrix \mathbf{W} .

The GMM estimator thus defined is asymptotically consistent (see [8, 9, 11] for additional details). The covariance matrix of the GMM estimator $\hat{\boldsymbol{\theta}}_n$ converges to $\mathbf{C} = \left(\mathbf{M}(\boldsymbol{\theta}_0) \right)^{-1} \tilde{\boldsymbol{\Sigma}} \left(\mathbf{M}'(\boldsymbol{\theta}_0) \right)^{-1}$, where $\tilde{\boldsymbol{\Sigma}} = \mathbf{H}'(\boldsymbol{\theta}_0) \mathbf{W}_n \mathbf{S} \mathbf{W}_n' \mathbf{H}(\boldsymbol{\theta}_0)$, $\mathbf{M}(\boldsymbol{\theta}) = \mathbf{H}'(\boldsymbol{\theta}) \mathbf{W} \mathbf{H}(\boldsymbol{\theta})$, $\mathbf{H}(\boldsymbol{\theta})$ is the limiting matrix of the partial derivatives of \mathbf{h}_n and \mathbf{S} is the limiting covariance matrix of \mathbf{h}_n . The smallest attainable asymptotic variance is obtained when $\mathbf{W}_n = \mathbf{S}^{-1}$.

Remark 1. Note that, in general, is not possible to find an analytical solution for the minimization of the quadratic form $\mathbf{h}_n(\boldsymbol{\theta}, \mathbf{X}_n)' \mathbf{W}_n \mathbf{h}_n(\boldsymbol{\theta}, \mathbf{X}_n)$ and we have to resort to numerical procedures. In order to obtain an efficient GMM estimator we can reformulate the GMM criterium as

$$Q_n(\boldsymbol{\theta}, \mathbf{X}_n) = \mathbf{h}_n(\boldsymbol{\theta}, \mathbf{X}_n)' \mathbf{W}_n(\boldsymbol{\theta}) \mathbf{h}_n(\boldsymbol{\theta}, \mathbf{X}_n), \quad (8)$$

where the weight matrix, which depends on the parameters, is obtained via plug-in or empirical estimation as the covariance matrix of \mathbf{h}_n and minimize this quadratic form with respect to $\boldsymbol{\theta}$. This procedure is called the *continuously updated GMM estimator*.

4.2 Monte Carlo Results

To illustrate the estimation procedure and to analyse the small sample properties of the (*continuously updated*) GMM estimators for the parameters of the proposed models, the results of a simulation study are presented. Thus, 5000 independent replicates of time series of length $n = 200, 500$ and 1000 are generated from each of the models. The mean estimate and the standard error of the estimates are obtained from the 5000 replications.

The minimization of (8) is performed by the R function *optim*, which accomplished a general-purpose optimization based on Nelder-Mead, quasi-Newton and conjugate-gradient algorithms and includes an option for box-constrained optimization [20].

INMA-NB(1) model:

For the univariate model proposed in Section 2, the independent replicates generated from the INMA-NB(1) process (3) with two parameters, $\boldsymbol{\theta} = (\kappa, \beta)$. There are three summary statistics related with the first- and second-order moments: mean, variance and autocovariance at lag 1, defined in (4).

The initial values for GMM estimation are obtained from the method of moments, namely $\hat{\beta} = \frac{\hat{\gamma}_x(1)}{\bar{x}}$, and $\hat{\kappa} = \frac{\hat{s}^2}{\hat{\beta}(1+\hat{\beta})(1+\hat{\beta}+\hat{\beta}^2)}$ where \bar{x} is the sample mean, \hat{s}^2 is the sample variance and $\hat{\gamma}_x(1)$ is the sample autocovariance in lag 1. The minimization of (8) is performed numerically and subject to $\beta \in]0, 1[$, $\kappa > 0$, for parameters with values $\kappa \in \{3.0, 5.0\}$ and $\beta \in \{0.2, 0.4, 0.6, 0.8\}$.

The mean estimates and standard errors from the 5000 replicates are given in Table 4.2. The sample mean and standard errors decrease as the sample size increases, indicating that the distribution of the estimators is consistent and symmetric. The results show that, in general, $\hat{\kappa}$ is overestimated while $\hat{\beta}$ is underestimated.

Table 1 Sample mean and standard errors (in brackets) of the estimates for the INMA-NB(1) models.

n	$\theta = (\kappa, \beta)$	$\hat{\kappa}$	$\hat{\beta}$	$\theta = (\kappa, \beta)$	$\hat{\kappa}$	$\hat{\beta}$
200		4.957 (5.185)	0.185 (0.086)		7.338 (5.883)	0.190 (0.083)
500	(3, 0.2)	3.566 (1.847)	0.193 (0.058)	(5, 0.2)	5.769 (2.565)	0.195 (0.054)
1000		3.248 (0.878)	0.196 (0.041)		5.382 (1.428)	0.196 (0.039)
200		3.561 (1.716)	0.382 (0.103)		5.812 (2.487)	0.386 (0.100)
500	(3, 0.4)	3.195 (0.720)	0.392 (0.066)	(5, 0.4)	5.297 (1.162)	0.393 (0.063)
1000		3.091 (0.483)	0.396 (0.047)		5.151 (0.763)	0.396 (0.045)
200		3.305 (0.952)	0.581 (0.111)		5.469 (1.479)	0.582 (0.105)
500	(3, 0.6)	3.122 (0.525)	0.592 (0.072)	(5, 0.6)	5.181 (0.842)	0.593 (0.069)
1000		3.058 (0.359)	0.596 (0.052)		5.085 (0.570)	0.597 (0.049)
200		3.331 (0.699)	0.756 (0.106)		5.519 (1.123)	0.760 (0.101)
500	(3, 0.8)	3.124 (0.409)	0.783 (0.072)	(5, 0.8)	5.183 (0.648)	0.786 (0.068)
1000		3.054 (0.290)	0.793 (0.053)		5.071 (0.467)	0.795 (0.051)

BINMA-BNB $\Pi(1, 1)$ model:

For the BINMA-BNB $\Pi(1, 1)$ model there are five parameters of interest $\theta = (\lambda_1, \lambda_2, \phi, \nu, \tau)$ and nine summary statistics concerning with the first- and second-order moments: mean, variance, autocovariance at lag 1 for each series and the cross-covariance at lag 1, 0 and -1, refer to equation 7. Since the method of moments yields (highly) non-linear equations and often non admissible estimates, the starting values for the minimization of (8) are given arbitrarily by $\theta^* = (\lambda_1, \lambda_2, \phi, \nu, \tau) = (1, 1, 0.5, 1, 2)$, satisfying the constraints in the parameters $\lambda_j > 0$, for $j = 1, 2$, $\phi \in [0, \min(\lambda_1, \lambda_2)[$ and $\nu, \tau > 0$. Two models were generated with the set of parameters given by $\theta = (1.0, 1.0, 0.3, 1.0, 2.0)$ and $\theta = (0.6, 0.9, 0.4, 1.0, 2.0)$, respectively. The mean estimates and corresponding standard errors are given in Table 2. For both set of parameters, the estimates are generally biased with standard errors that decrease as the sample size increases. Note that the asymptotic theory suggests that it is preferable to include as many summary statistics as possible to maximize que information extracted from the data, and therefore the GMM estimator is always over-identified, since the number of parameters is less than the number of sum-

mary statistics. However, this redundant information may lead to deterioration in the GMM performance, in particular, with the highly correlated sample moments resulting in a badly conditioned weighting matrix [2]. Thus, further analysis is needed concerning the estimation of the BINMA-BNB $\Pi(1, 1)$ model.

Table 2 Mean and standard errors (in brackets) of the estimates for the BINMA-BNB $\Pi(1, 1)$ model.

$\theta = (\lambda_1, \lambda_2, \phi, \nu, \tau)$	n	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\phi}$	$\hat{\nu}$	$\hat{\tau}$
(1.0, 1.0, 0.3, 1.0, 2.0) $\beta_1 = \beta_2 = 0.65$	200	0.930 (0.216)	0.943 (0.241)	0.313 (0.227)	1.220 (0.332)	2.236 (0.280)
	500	1.008 (0.144)	1.010 (0.159)	0.329 (0.184)	1.088 (0.168)	2.195 (0.159)
	1000	1.043 (0.103)	1.044 (0.113)	0.326 (0.136)	1.047 (0.112)	2.177 (0.114)
(0.6, 0.9, 0.4, 1.0, 2.0) $\beta_1 = 0.5; \beta_2 = 0.65$	200	0.751 (0.193)	1.055 (0.257)	0.370 (0.210)	1.065 (0.342)	2.354 (0.298)
	500	0.786 (0.134)	1.105 (0.166)	0.417 (0.167)	0.942 (0.158)	2.303 (0.137)
	1000	0.803 (0.102)	1.127 (0.120)	0.426 (0.132)	0.908 (0.109)	2.282 (0.103)

5 Real Data Illustration

In this section, a dataset is used to illustrate the univariate model and methods developed previously. To assess the fit the model, a range of diagnostic and validation tools based on parametric resampling are used.

Consider a dataset concerning the number of Sex Offences reported in the 21st police car beat in Pittsburgh (Pennsylvania, USA), during one month. The data consist of $n = 144$ observations, from January of 1990 to December of 2001 (see Figure 1). The dataset is available from the Forecasting Principles site <http://www.forecastingprinciples.com/index.php/crimedata>. From a preliminary analysis of the sample mean and variance of the data, Table 3, it is possible to conclude that the data seem to be overdispersed. Furthermore, the values of the sample autocorrelation function (ACF) in Figure 1, which are nearly zero after the lag 1, suggest that a first-order model is appropriate to this data set. These data has been studied previously by [21] who fitted an autoregressive model. However, considering that it is expected that the sex offenders stay a limited time in the system and that these individuals can come and go several times during their life times in the system, a moving average model, INMA-NB(1) may provide a good fit.

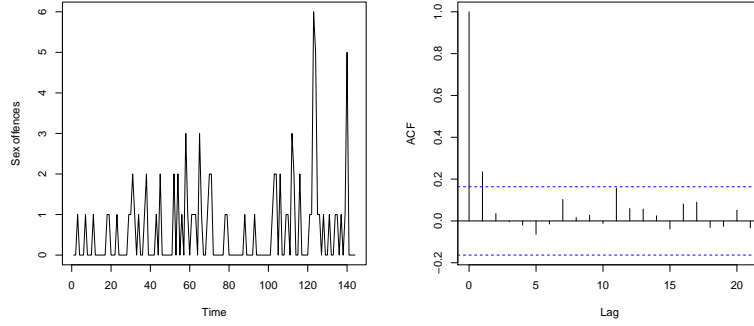


Fig. 1 Time series (left panel) and autocorrelation function (right panel) of the monthly number of Sex Offences, from 1990 to 2001, registered in Pittsburgh.

Table 3 Sample measures for the sex offences data set.

	Mean	Variance	ACF(1)
Sample	0.590	1.027	0.235
INMA-NB(1)	0.588	1.011	0.282
NGINAR(1) [21]	0.587	0.932	0.165

The obtained GMM estimates are $\hat{\theta} = (\hat{\kappa}, \hat{\beta}) = (0.485, 0.816)$ with standard errors given by $(0.162, 0.329)$, respectively (starting value $\theta^* = (\kappa, \beta) = (1.13, 0.41)$ given by the method of moments).

The parametric bootstrap (see [13] for details) is used to assess the adequacy of the model to represent specific features of interest of the data, in this case the autocorrelation function. The fitted model is used to generate 5000 (univariate) time series samples, all with the same number of observations as the original data set, which are then used to construct an empirical distribution for the ACF. Figure 2 represents the acceptance envelope computed from the 2.5% and 97.5% quantiles of the empirical distribution for the ACF. It is clear that the model represents adequately the autocorrelation. Furthermore, Table 3 indicates that the fitted INMA-NB(1) model is competitive model when compared with the NGINAR(1) model fitted to this data set by [21], specially in what concerns to the autocorrelation function.

6 Final Remarks

Modelling overdispersed time series of counts is a relevant issue and several autoregressive type models have been proposed in the literature. However, often a

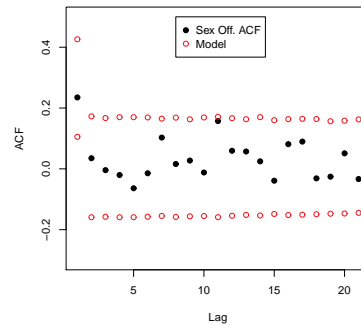


Fig. 2 Acceptance envelope for the autocorrelation function for the sex offences data set.

moving average model may be more adequate for the data under analysis and, so far, the literature is lacking MA models appropriate for overdispersed time series of counts. This work contributes to closing this gap with new first-order integer-valued moving average univariate and bivariate models based on the NB thinning operation. The univariate process with NB arrivals presents a NB marginal distribution. However, in the case of bivariate model with bivariate NB innovations the joint distribution is not the bivariate NB. Nevertheless, this bivariate model can still account for overdispersion. For estimation purposes, the GMM is considered and the finite-sample behaviour is analysed through a simulation study. The proposed univariate model is fruitfully applied to a real dataset. Further studies of high-order INMA-NB models will be reported elsewhere.

Acknowledgements This research was supported by the Portuguese national funding agency for science, research and technology (FCT), within the Center for Research and Development in Mathematics and Applications (CIDMA), project UID/MAT/04106/2019.

References

1. Al-Osh, M. A., Alzaid, A. A.: Integer-valued moving average (INMA) process. *Statist. Papers* 29, 281–300 (1988) [10.1007/BF02924535](https://doi.org/10.1007/BF02924535)
2. Andersen, T., Sørensen, B.: GMM Estimation of a Stochastic Volatility Model: A Monte Carlo Study. *J. Bus. Econ. Stat.* 14, 328–352 (1996) [doi:10.2307/1392446](https://doi.org/10.2307/1392446)
3. Brännäs, K., Hall, A.: Estimation in integer-valued moving average models. *Appl. Stoch. Models Bus. Ind.* 17, 277–291 (2001) [doi: 10.1002/asmb.445](https://doi.org/10.1002/asmb.445)
4. Brännäs, K., Hellström, J., Nordström, J.: A new approach to modelling and forecasting monthly guest nights in hotels. *Int. J. Forecast.* 18, 19–30 (2002) [doi: 10.1016/S0169-2070\(01\)00104-2](https://doi.org/10.1016/S0169-2070(01)00104-2)
5. Cossette H., Marceau E., Maume-Deschamps V.: Discrete-time risk models based on time series for count random variables. *ASTIN Bulletin: The Journal of the IAA* 40, 123–150

- (2010) doi: 10.2143/AST.40.1.2049221
6. Cossette H., Marceau E., Tourelle F.: Risk models based on time series for count random variables. *Ins.: Mathematics Econ.* 48, 19–28 (2011) doi: 10.1016/j.insmatheco.2010.08.007
 7. Edwards, C. B., Gurland, J.: A class of distributions applicable to accidents. *J. Am. Stat. Assoc.* 56, 503–517 (1961) doi: 10.1080/01621459.1961.10480641
 8. Greene, W. H. *Econometric Analysis*. Prentice Hall, Upper Saddle River, N.J. (2000)
 9. Hansen, L. P.: Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054 (1982) doi: 10.2307/1912775
 10. Hu X., Zhang L., Sun W.: Risk model based on the first-order integer-valued moving average process with compound Poisson distributed innovations. *Scand. Actuar. J.* 2018, 412–425 (2018) doi: 10.1080/03461238.2017.1371067
 11. Jesus, J., Chandler, R. E.: Estimating functions and the generalized method of moments. *Interface Focus* 1, 871–885 (2011) doi: 10.1098/rsfs.2011.0057
 12. Johnson, N. L., Kemp, A. W., Kotz, S.: *Univariate discrete distributions*. John Wiley & Sons, New Jersey (2005)
 13. Jung, R. C., McCabe, B. P. M., Tremayne, A. R.: Model validation and diagnostics. In: Davis, R. A., Holan, S. H., Lund, R., Ravishanker, N. (eds.) *Handbook of Discrete-Valued Time Series*, pp. 189–218. Chapman & Hall/CRC Press (2015)
 14. Kocherlakota, S., Kocherlakota, K.: *Bivariate discrete distributions*. Markel Dekker, New York (1992)
 15. Mamode Khan, N., Sunecher, Y., Jowaheer, V.: Analyzing the Full BINMA Time Series Process Using a Robust GQL Approach. *J. Time Ser. Econom.* 9 (2016) doi: 10.1515/jtse-2015-0019
 16. Mamode Khan, N., Sunecher, Y., Jowaheer, V.: Inferential methods for an unconstrained non-stationary BINMA time series process with Poisson innovations, *J. Stat. Theory Pract.* 11, 76–106 (2017) doi: 10.1080/15598608.2016.1258600
 17. Ma, D., Wang, D., Cheng, J.: Bidimensional discrete-time risk models based on bivariate claim count time series. *J. Inequal. Appl.* 2015: 105, (2015) doi: 10.1186/s13660-015-0618-3
 18. McKenzie, E.: Some ARMA models for dependent sequences of Poisson counts. *Adv. Appl. Probab.* 20, 822–835 (1988) doi: 10.2307/1427362
 19. Quoreshi, A. M. M. S.: Bivariate time series modeling of financial count data. *Commun. Stat. - Theor. M.* 35, 1343–1358 (2006) doi: 10.1080/03610920600692649
 20. R Development Core Team: *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing (2008) ([url =http://www.R-project.org](http://www.R-project.org))
 21. Ristić, M. M., Bakouch, H. S., Nastić, A. S.: A new geometric first-order integer-valued autoregressive (NGINAR (1)) process. *J. Stat. Plan. Inference* 139, 2218–2226 (2009) doi: 10.1016/j.jspi.2008.10.007
 22. Scotto, M. G., Weiß, C. H., Gouveia, S. Thinning-based models in the analysis of integer-valued time series: a review. *Stat. Model.* 15, 590–618 (2015) doi: 10.1177/1471082X15584701
 23. Steutel, F. W., Van Harn, K.: Discrete analogues of self-decomposability and stability. *Ann. Probab.* 7, 893–899 (1979) doi: 10.1214/aop/1176994950
 24. Sunecher, Y., Mamode Khan, N., Jowaheer, V.: Estimating the parameters of a BINMA Poisson Model for a non-stationary bivariate time series. *Commun. Stat. Simul. Comput.* 46, 6803–6827 (2016) doi: 10.1080/03610918.2016.1212068
 25. Torres, C.: Contributions to the analysis of multivariate count data. PhD Thesis, Universidade do Porto, Portugal (2016)
 26. Weiß, C. H.: Serial dependence and regression of Poisson INARMA models. *J. Stat. Plan. Inference* 138, 2975–2990 (2008) doi: 10.1016/j.jspi.2007.11.009
 27. Weiß, C. H.: *An Introduction to Discrete-Valued Time Series*. John Wiley & Sons, Inc., Chichester (2018)
 28. Yu, K., Zou, H., Shi, D.: Integer-valued moving average models with structural changes. *Math. probl. eng.* 2014, Article ID 231592, 5 pages (2014) doi: 10.1155/2014/231592
 29. Yu, K., Zou, H.: The Combined Poisson INMA(q) Models for Time Series of Counts. *J. Appl. Math.* 2015, Article ID 457842, 7 pages (2015) doi: 10.1155/2015/457842

30. Zhang L., Hu X., Duan B.: Optimal reinsurance under adjustment coefficient measure in a discrete risk model based on Poisson MA(1) process. *Scand. Actuar. J.* 2015, 455–467 (2015)
doi: 10.1080/03461238.2013.849615
31. Zou, H., Yu, K.: First order threshold integer-valued moving average processes, *Dynam. Cont. Dis. Ser. B* 21, 197–205 (2014)