Afrika Statistika

ISSN 2316-090X

Afrika Statistika
Vol. 3 (2), 2018, pages 1645–1666.
DOI: http://dx.doi.org/10.16929/as/1645.126

A negative binomial mixture integer-valued GARCH model

Mamadou Lamine DIOP*, **Aliou DIOP and Abdou Kâ DIONGUE**

LERSTAD, Université Gaston Berger, BP 234 Saint-Louis, Sénégal.

Received on March 14, 2018; Accepted on May 6, 2018

Abstract. This paper generalizes the negative binomial integer-valued GARCH model (NBINGARCH) to a negative binomial mixture integer-valued GARCH (NBMINGARCH) for modeling time series of counts with presence of overdispersion. This class of models consists of a mixture of K stationary or non-stationary negative binomial integer-valued GARCH components. The advantage of these models over the NBINGARCH models includes the ability to handle multimodality and nonstationary components. Compared to the MINGARCH models, this class of models is more flexible to describe the greater degrees of overdispersion. The necessary and sufficient first and second order stationarity conditions are investigated. The estimation of parameters is done through an EM algorithm and the model is selected by some information criterions. Some simulation results and real data application are provided.

Key words: Time series of counts, Mixture model, Negative binomial, Overdispersion.
AMS 2010 Mathematics Subject Classification : 62F10, 62M10.

---

*Mamadou Lamine Diop: diopml@yahoo.fr
Aliou Diop: aliou.diop@ugb.edu.sn
Abdou Kâ Diongue : abdou.diongue@ugb.edu.sn

Résumé. (French) Ce papier propose une généralisation des modèles NBINGARCH aux modèles de mélange NB-MINGARCH pour la modélisation des séries temporelles de comptage. Ces modèles possèdent des avantages par rapport aux modèles NBINGARCH en raison de leur capacité à prendre en compte la multimodalité souvent observée dans la distribution conditionnelle. Comparée aux modèles MINGARCH, cette classe de modèles a une plus grande flexibilité à décrire les séries présentant un degré de surdispersion élevé. Les conditions de la stationnarité des moments d'ordre 1 et 2 sont établies. L'estimation des paramètres de ces modèles est faite en utilisant la procédure EM. La sélection des modèles est effectuée par une règle de décision basée sur des critères d'information. Quelques résultats de simulations sont présentés. Ces modèles sont aussi appliqués sur un exemple de données réelles.

1. Introduction

The *DARMA models*, developed by Jacobs and Lewis (1978), are the first class of models for modeling time series count data. But, they have some disadvantages because they do not allow to model the series taking an infinite number of values. The INAR models are the first whose structure is similar to that of ARMA models. They are studied by Al-Osh and Alzaid (1987) and McKenzie (1985) by using the thinning operation of Steutel and Van Harn (1979). Due to its limitations, several authors have proposed either extensions or new models. For example, an extension of the INAR(1) process which is useful for modeling discrete-time dependent counting processes is introduced by Al-Osh and Alzaid (1990), while Kachour and Yao (2009) have proposed a new class of autoregressive models for integer-valued time series using the rounding operator. Other models have also emerged, with particularity as the ability to capture some specific phenomena often displayed by counts data, such as the overdispersion or the excess of zeros. In recent years, several models have been proposed to describe the dependence structure of the observations. We refer to Ferland *et al.* (2006), Neumann (2011), Zhu (2012a), Zhu (2012b)) and Davis and Liu (2012). For other references, see, Fokianos and Neumann (2013), Moysiadis and Fokianos (2014) and Fried *et al.* (2015).

However, as pointed out in some works, these models are obtained by assuming that the series is unimodal whereas many time series may exhibit multimodality either in the marginal or the conditional distribution. The main tools generally used to take account this phenomenon are the mixture models (see Jalali and Pemberton (1995) or Mclachlan and Peel (2000)).Authors such as Zhu *et al.* (2010) have generalized the INARCH model to the mixture model (MINARCH). The model is obtained with the Poisson distribution and has an advantage over the INARCH model because of its ability to handle multimodality and non-stationary components. It is in the same line ideas that Diop *et al.* (2016) generalized the MINARCH model to a mixture integer-valued GARCH (MINGARCH). This model includes the ability to take into account the moving average (*MA*) components of the series. In statistical research, the Poisson distribution is generally the first probability distribution considered to analyze count data. But, for the analysis of count time series, the variability of the series is often larger than the mean. In literature, the negative binomial is considered to be the prototype for the integer-valued time series when this phenomenon is observed (see Hoef and Boveng (2007), Liden and Mantyniemi (2011), Molla and Muniswamy (2012)).

In this contribution, we introduce a new class of models that is the negative binomial mixture GARCH by using the negative binomial distribution. These models are a generalization of the NBINGARCH model (introduced by Zhu (2011)) to a mixture model. The advantages of this mixture model over the NBINGARCH model include the ability to handle multimodality and non-stationary components. This class of models is also very flexible to describe the greater degrees of overdispersion.

The paper is organized as follows. In Section 2, we describe the NB-MINGARCH model and the stationarity conditions. In Section 3, we discuss the estimation procedure with some simulation experiments. A real data application is presented in Section 4. In Section 5, we conclude by giving some remarks.

2. The negative binomial mixture integer-valued GARCH

To define this class of models, assume that $\{Y_t,\ t \in \mathbb{Z}\}$ is a time series of counts. The NB-MINGARCH $(K; r_1, \ldots, r_K; p_1, \ldots, p_K; q_1, \ldots, q_K)$ model is defined by

$$\begin{cases} Y_t = \sum\limits_{k=1}^{K} \mathbf{1}_{(\eta_t = k)} Y_{kt}, \\ Y_{kt}|\mathcal{F}_{t-1} : NB(r_k, p_{kt}), \\ \frac{1-p_{kt}}{p_{kt}} = \lambda_{kt} = \alpha_{k0} + \sum\limits_{i=1}^{p_k} \alpha_{ki} Y_{t-i} + \sum\limits_{j=1}^{q_k} \beta_{kj} \lambda_{k(t-j)}, \end{cases} \quad (1)$$

where $r_k \in \mathbb{N}^*,\ \alpha_{k0} > 0,\ \alpha_{ki} \geq 0,\ \beta_{kj} \geq 0,\ (i = 1, \cdots, p_k,\ j = 1, \cdots, q_k,\ k = 1, \cdots, K)$, $NB(r,p)$ is the negative binomial distribution with parameters r and p, $\mathbf{1}_{(.)}$ denotes the indicator function, p_k and q_k are respectively the order of AR and MA for the k^{th} component, $\mathcal{F}_{t-1}$ is the information set up to time $t-1$ and η_t is a sequence of independent and identically distributed random variables. It is assumed that:

- Y_{t-j} and η_t are independent for all $t > 0$ and $j > 0$,
- given $\mathcal{F}_{t-1}$, Y_{kt} and η_t are conditionally independent,
- the process $\{\eta_t,\ t \in \mathbb{N}\}$ is such that $\mathbb{P}(\eta_t = k) = \alpha_k,\ k = 1, \cdots, K$ with $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_K$ and $\sum\limits_{k=1}^{K} \alpha_k = 1$.

The conditional probability mass function of Y_{kt} has the form

$$\mathbb{P}\left(Y_{kt} = y|\mathcal{F}_{t-1}\right) = \binom{y + r_k - 1}{r - 1} p_{kt}^{r_k} (1 - p_{kt})^y, \text{ for } y = 0, 1, \ldots$$

where

$$p_{kt} = \frac{1}{1+\lambda_{kt}}, \quad 1 - p_{kt} = \frac{\lambda_{kt}}{1+\lambda_{kt}}.$$

The conditional mean and variance of Y_t are given by

$$\mathbb{E}\Big(Y_t|\mathcal{F}_{t-1}\Big) = \sum_{k=1}^{K}\alpha_k r_k \lambda_{kt},$$

and

$$\mathrm{Var}\Big(Y_t|\mathcal{F}_{t-1}\Big) = \mathbb{E}\Big(Y_t|\mathcal{F}_{t-1}\Big) + \sum_{k=1}^{K}(r_k + r_k^2)\alpha_k\lambda_{kt}^2 - \Big(\sum_{k=1}^{K}\alpha_k r_k \lambda_{kt}\Big)^2.$$

The variance is given by

$$\begin{aligned}
\mathrm{Var}\Big(Y_t\Big) &= \mathbb{E}\Big(\mathrm{Var}(Y_t|\mathcal{F}_{t-1})\Big) + \mathrm{Var}\Big(\mathbb{E}(Y_t|\mathcal{F}_{t-1})\Big) \\
&= \mathbb{E}\Big(Y_t\Big) + \sum_{k=1}^{K}(r_k + r_k^2)\alpha_k\mathbb{E}\Big(\lambda_{kt}^2\Big) - \Big(\mathbb{E}\Big(\sum_{k=1}^{K}\alpha_k r_k\lambda_{kt}\Big)\Big)^2 \\
&\geq \mathbb{E}\Big(Y_t\Big) + \mathbb{E}\Big(\sum_{k=1}^{K}\alpha_k r_k^2\lambda_{kt}^2 - \Big(\sum_{k=1}^{K}\alpha_k r_k\lambda_{kt}\Big)^2\Big) + \sum_{k=1}^{K}\alpha_k r_k\mathbb{E}\Big(\lambda_{kt}^2\Big).
\end{aligned}$$

Using the Jensen's inequality, it is easy to see that:

$$\sum_{k=1}^{K}\alpha_k r_k^2\lambda_{kt}^2 - \Big(\sum_{k=1}^{K}\alpha_k r_k\lambda_{kt}\Big)^2 > 0.$$

Hence, the variance is larger than the mean, which indicates that the NB-MINGARCH model is able to describe the time series count with overdispersion. In addition, with the negative binomial distribution, the larger the probability p_{kt} is, the more the variance is superior to the mean. This property favors using a negative binomial over a Poisson distribution. In the formula for the variance, we can also remark that the largest degree of overdispersion can be described with appropriate choices for the parameters r_k.

Remark 1. If $K = 1$, then the NB-MINGARCH model becomes the negative binomial integer-valued GARCH model (NBINGARCH model) proposed by Zhu (2011).

Let now introduce the polynomials $D_k(B) = 1 - \beta_{k1}B - \cdots - \beta_{kq}B^q$, where B is the backshift operator. We consider the hypothesis H_1 and H_2 given by

$$\begin{aligned} H_1 &: \textit{For } k = 1, \ldots, K, \textit{ the roots of } D_k(z) = 0 \textit{ lie outside the unit circle.} \\ H_2 &: \textit{For any fixed } t \textit{ and } k, \quad \lambda_{kt} < \infty \textit{ a.s.} \end{aligned}$$

The condition H_1 is equivalent to $\sum_{j=1}^{q} \beta_{kj} < 1$, for any fixed k.
In the following, define: $p = \max(p_1, \cdots, p_K)$, $q = \max(q_1, \cdots, q_K)$, and $L = \max(p, q)$.
We assume that $\alpha_{ki} = 0$, for all $i > p_k$ and $\beta_{kj} = 0$, for all $j > q_k$.

The first stationarity conditions for the NB-MINGARCH model (1) is given in the following theorem.

Theorem 1. *Assume that the conditions H_1 and H_2 are satisfied. A necessary and sufficient condition for model (1) to be stationary in the mean is that the roots of the equation:*

$$1 - \sum_{k=1}^{K} \alpha_k r_k \left(\frac{\sum_{i=1}^{p_k} \alpha_{ki} Z^{-i}}{1 - \sum_{i=1}^{q_k} \beta_{kj} Z^{-j}} \right) = 0 \tag{2}$$

lie inside the unit circle.

Proof: Let $\mu_t = \mathbb{E}(Y_t) = \sum_{k=1}^{K} \alpha_k r_k \mathbb{E}(\lambda_{kt})$, for all $t \in \mathbb{Z}$. Recall that $\lambda_{kt} = \alpha_{k0} + \sum_{i=1}^{p_k} \alpha_{ki} Y_{t-i} + \sum_{j=1}^{q_k} \beta_{kj} \lambda_{k(t-j)}$. For all $m \geq 1$, by recurrence, we have

$$\begin{aligned} \lambda_{kt} &= \alpha_{k_0} + \sum_{i=1}^{L} \alpha_{ki} Y_{t-i} + \sum_{l=1}^{m} \sum_{j_1, \cdots, j_l = 1}^{L} \alpha_{k0} \beta_{kj_1} \cdots \beta_{kj_l} \\ &+ \sum_{l=1}^{m} \sum_{j_1, \cdots, j_{l+1} = 1}^{L} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} Y_{t-j_1-\cdots-j_l-j_{l+1}} \\ &+ \sum_{j_1, \cdots, j_{m+1} = 1}^{L} \beta_{kj_1} \cdots \beta_{kj_{m+1}} \lambda_{k(t-j_1-\cdots-j_{m+1})}. \end{aligned}$$

For $k = 1, \cdots, K$, let $C_{k0} = \sum_{l=0}^{\infty} \sum_{j_1 \cdots j_l = 1}^{L} \alpha_{k0} \beta_{kj_1} \cdots \beta_{kj_l}$ and define

$$\lambda'_{kt} = C_{k0} + \sum_{i=1}^{L} \alpha_{ki} Y_{t-i} + \sum_{l=1}^{\infty} \sum_{j_1, \cdots, j_{l+1} = 1}^{L} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} Y_{t-j_1-j_2-\cdots-j_{l+1}}, \text{ for all } t > 0.$$

Since $\sum_{j=1}^{L} \beta_{kj} < 1$, it is easy to see that $\lambda'_{kt} < \infty$ a.s. for any fixed t and k.

We will show below that $\lambda_{kt} = \lambda'_{kt}$ almost surely as $m \to \infty$ for any fixed t and k. In what follows, C will denote any positive constants whose value is unimportant and may vary from line to line. Let t and k be fixed now. It follows that for any $m \geq 1$

$$
\begin{aligned}
|\lambda_{kt} - \lambda'_{kt}| \leq & \sum_{l=m+1}^{\infty} \sum_{j_1 \cdots j_l = 1}^{L} \alpha_{k0} \beta_{kj_1} \cdots \beta_{kj_l} + \sum_{l=m+1}^{\infty} \sum_{j_1, \cdots, j_{l+1}=1}^{L} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} Y_{t-j_1-j_2-\cdots-j_{l+1}} \\
& + \sum_{j_1, \cdots, j_{m+1}=1}^{L} \beta_{kj_1} \cdots \beta_{kj_{m+1}} \lambda_{k(t-j_1-\cdots-j_{m+1})}.
\end{aligned}
$$

First

$$
\mathbb{E}\left\{ \sum_{j_1 \cdots j_l=1}^{L} \alpha_{k0} \beta_{kj_1} \cdots \beta_{kj_l} + \sum_{j_1, \cdots, j_{l+1}=1}^{L} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} Y_{t-j_1-j_2-\cdots-j_{l+1}} \right\} \leq C \left(\sum_{j=1}^{L} \beta_{kj} \right)^{l}
$$

and

$$
\mathbb{E}\left\{ \sum_{j_1, \cdots, j_{m+1}=1}^{L} \beta_{kj_1} \cdots \beta_{kj_{m+1}} \lambda_{k(t-j_1-\cdots-j_{m+1})} \right\} \leq C \left(\sum_{j=1}^{L} \beta_{kj} \right)^{m+1}.
$$

The expectation of the right-hand side of the above is bounded by

$$
M = C \left(1 - \sum_{j=1}^{L} \beta_{kj} \right)^{-1} \left(\sum_{j=1}^{L} \beta_{kj} \right)^{m+1}.
$$

Let $A_m = \left\{ |\lambda_{kt} - \lambda'_{kt}| > \frac{1}{m} \right\}$. Then

$$
\mathbb{P}(A_m) \leq mC \left(1 - \sum_{j=1}^{L} \beta_{kj} \right)^{-1} \left(\sum_{j=1}^{L} \beta_{kj} \right)^{m+1}.
$$

According to the convergence of the series associated to $U_m = mC \left(1 - \sum_{j=1}^{L} \beta_{kj} \right)^{-1} \left(\sum_{j=1}^{L} \beta_{kj} \right)^{m+1}$, we can see that $\sum_{m=1}^{\infty} \mathbb{P}(A_m) < \infty$. Then, using Borel-Cantelli lemma and the fact that $A_m \subset A_{m+1}$, we can show that $\lambda_{kt} = \lambda'_{kt}$ a.s.
Therefore,

$$
\mu_t = \left\{ \sum_{k=1}^{K} \alpha_k r_k C_{k0} + \sum_{l=0}^{\infty} \sum_{j_1 \cdots j_{l+1}=1}^{L} \sum_{k=1}^{K} \alpha_k r_k \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} \mu_{t-j_1-j_2-\cdots-j_{l+1}} \right\}. \quad (3)
$$

The equation can be rewritten as:

$$\left\{1 - \sum_{l=0}^{\infty} \sum_{j_1 \cdots j_{l+1}=1}^{L} \sum_{k=1}^{K} \alpha_k r_k \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} B^{(j_1+j_2+\cdots+j_{l+1})}\right\} \mu_t = \sum_{k=1}^{K} \alpha_k r_k C_{k0},$$

where B is the backward shift operator. The necessary and sufficient condition for the existence of stationary solution is that all roots of the equation:

$$1 - \sum_{l=0}^{\infty} \sum_{j_1 \cdots j_{l+1}=1}^{L} \sum_{k=1}^{K} \alpha_k r_k \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} Z^{-(j_1+j_2+\cdots+j_{l+1})} = 0$$

lie inside the unit circle (see Goldberg (1958)).
Since $\sum_{j=1}^{q_k} \beta_{kj} < 1$ and $\|Z\| > 1$, this equation is equivalent to

$$1 - \sum_{k=1}^{K} \alpha_k r_k \left(\sum_{i=1}^{p_k} \alpha_{ki} Z^{-i}\right) \sum_{l=0}^{\infty} \left(\sum_{j=1}^{q_k} \beta_{kj} Z^{-j}\right)^l = 0.$$

Hence, the equation (2) follows. □

As an illustration, we give two special cases of Theorem 1 in Corollary 1 and Corollary 2. First, consider the NBINGARCH model corresponding to the NB-MINGARCH model with $K = 1$.

Corollary 1. *Assume that $p \geq q$. A necessary and sufficient condition for the NBINGARCH(p, q) model to be stationary is that the roots of the equation*

$$1 - r_1 \sum_{i=1}^{p} \alpha_{1i} Z^{-i} - \sum_{j=1}^{q} \beta_{1j} Z^{-j} = 0$$

lie inside the unit circle (as in Zhu (2011)).

Now, consider in the following corollary the NB-MINGARCH model with $p_k = q_k = 1$ for all $k = 1, \cdots, K$.

Corollary 2. *A necessary and sufficient condition for the NB-MINGARCH$(K; r_1, \cdots, r_K; 1, \cdots, 1; 1, \cdots, 1)$ model to be stationary in the mean is that the roots of the equation*

$$1 + C_1 Z^{-1} + C_2 Z^{-2} + \cdots + C_K Z^{-K} = 0$$

lie inside the unit circle, where

$$C_1 = -\sum_{k=1}^{K}\left(\delta_k + \alpha_k r_k \gamma_k\right)$$

and

$$C_j = (-1)^j \left[\sum_{k_1>k_2>\cdots>k_j}^{K} \delta_{k_1}\delta_{k_2}\cdots\delta_{k_j} + \sum_{k=1}^{K}\alpha_k r_k \gamma_k \left(\sum_{\substack{k_1>k_2>\cdots>k_{j-1}\\ k_1\neq k,\ k_2\neq k,\cdots,k_{j-1}\neq k}}^{K} \delta_{k_1}\delta_{k_2}\cdots\delta_{k_{j-1}}\right)\right],$$

for $j = 2,\cdots,K$*, with* $\gamma_k = \alpha_{k1}$ *and* $\delta_k = \beta_{k1}$*.*

Proof: The equation (2) becomes

$$1 - \sum_{k=1}^{K}\frac{\alpha_k r_k \gamma_k Z^{-1}}{1-\delta_k Z^{-1}} = 0.$$

The previous equation is equivalent to:

$$\prod_{k=1}^{K}\left(1-\delta_k Z^{-1}\right) - \sum_{k=1}^{K}\alpha_k r_k \gamma_k Z^{-1}\prod_{\substack{k'=1\\k'\neq k}}^{K}(1-\delta_{k'}Z^{-1}) = 1 + C_1 Z^{-1} + C_2 Z^{-2} + \cdots + C_2 Z^{-K} = 0. \square$$

If the process $\{Y_t\}_{t\in\mathbb{Z}}$ is first-order stationary, then $\mathbb{E}(Y_t)$ is a constant independent of t. From the equation (3), we have

$$\mathbb{E}(Y_t) = \mu = \sum_{k=1}^{K}\alpha_k r_k C_{k0} + \mu\sum_{l=0}^{\infty}\sum_{j_1\cdots j_{l+1}=1}^{L}\sum_{k=1}^{K}\alpha_k r_k \alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}.$$

Hence

$$\mu = \frac{\sum_{k=1}^{K}\left(\frac{\alpha_k r_k \alpha_{k0}}{1-\sum_{j=1}^{q_k}\beta_{kj}}\right)}{1-\sum_{k=1}^{K}\left(\frac{\sum_{i=1}^{p_k}\alpha_k r_k \alpha_{ki}}{1-\sum_{j=1}^{q_k}\beta_{kj}}\right)}.$$

In the following proposition, we give a simple condition for the existence of first-order stationarity.

Proposition 1. *A necessary condition for the existence of first-order stationarity is:*

$$\sum_{j=1}^{q_k}\beta_{kj} < 1, \qquad \text{for } k = 1,\ldots,K \quad \text{and} \quad \sum_{k=1}^{K}\left(\frac{\sum_{i=1}^{p_k}\alpha_k r_k \alpha_{ki}}{1-\sum_{j=1}^{q_k}\beta_{kj}}\right) < 1.$$

Remark 2. As a special case, a necessary condition for the NB-MINGARCH$(2; r_1, r_2; 1, 1; 1, 1)$ model to be stationary in the mean is:

$$\beta_{11} < 1,\ \beta_{21} < 1 \qquad \text{and} \qquad \frac{\alpha_1 r_1 \alpha_{11}}{1-\beta_{11}} + \frac{\alpha_2 r_2 \alpha_{21}}{1-\beta_{21}} < 1.$$

Now, we suppose that the process $\{Y_t\}_{t\in\mathbb{Z}}$ is first-order stationary. The following theorem gives a second-order stationarity condition for the NB-MINGARCH model.

Theorem 2. *Under the hypothesis of Theorem 1, assume that the process $(Y_t)_{t\in\mathbb{Z}}$ is first-order stationary. Then, a necessary and sufficient condition for the process to be second-order stationary is that all roots of $1 - c_1 Z^{-1} - c_2 Z^{-2} - \cdots - c_L Z^{-L} = 0$ lie inside the unit circle, where*

$$\begin{aligned}
c_u &= \sum_{k=1}^{K} \alpha_k (r_k + r_k^2) \left(\Delta_{k,u} - \sum_{v=1}^{L-1} \Lambda_{kv} b_{vu} \omega_{u0} \right),\ u = 1, \cdots, L-1, \\
c_L &= \sum_{k=1}^{K} \alpha_k (r_k + r_k^2) \Delta_{k,L}, \\
\Delta_{k,i} &= \Delta_{k,i}^{(1)} + \Delta_{k,i}^{(2)}, \\
\Delta_{k,i}^{(1)} &= \sum_{l=0}^{\infty} \sum_{\substack{j_{l+2}=i \\ j_{l+2}=j_1+\cdots+j_{l+1}}}^{L} \alpha_{kj_{l+1}} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_l} \\
\Delta_{k,i}^{(2)} &= \sum_{\substack{l=0 \\ l'=0}}^{\infty} \sum_{\substack{j_1+\cdots+j_{l+2}=i \\ j_1+\cdots+j_{l+2}=j'_1+\cdots+j'_{l'+1}}}^{L} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_{l+1}} \alpha_{kj'_{l'+1}} \beta_{kj'_1} \cdots \beta_{kj'_{l'}} \\
\Lambda_{kv} &= \Lambda_{kv}^{(1)} + \Lambda_{kv}^{(2)} \\
\Lambda_{kv}^{(1)} &= \sum_{l=0}^{\infty} \sum_{|j_{l+2}-j_1-\cdots-j_{l+1}|=v}^{L} \alpha_{kj_{l+1}} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_l} \\
\Lambda_{kv}^{(2)} &= \sum_{\substack{l=0 \\ l'=0}}^{\infty} \sum_{|j_1+\cdots+j_{l+2}-j'_1-\cdots-j'_{l'+1}|=v}^{L} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_{l+1}} \alpha_{kj'_{l'+1}} \beta_{kj'_1} \cdots \beta_{kj'_{l'}},
\end{aligned}$$

$\Gamma = (\omega_{ij})_{i,j=1}^{L-1}$, $\Gamma^{-1} = (b_{ij})_{i,j=1}^{L-1}$, *two matrices such that* $\omega_{ii} = \sum_{l=0}^{\infty} \sum_{k=1}^{K} \delta_{iikl} - 1$, $\omega_{iu} = \sum_{l=0}^{\infty} \sum_{k=1}^{K} \delta_{iukl}$ *for* $u \neq i$ *and* $\delta_{iukl} = \alpha_k r_k \sum_{|i-j_1-\cdots-j_{l+1}|=u} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l}$.

Proof: Let $\gamma_{i,t} = \mathbb{E}(Y_t Y_{t-i})$ for $i = 0, 1, \cdots, L$. We have

$$\gamma_{i,t} = \sum_{i=1}^{K} \alpha_k r_k \mathbb{E}(\lambda_{kt} Y_{t-i}).$$

Using the same arguments as in the proof of Theorem 1, we can show that almost surely

$$\begin{aligned}\gamma_{i,t} &= \sum_{k=1}^{K}\alpha_{k0}\alpha_k r_k\mathbb{E}(Y_{t-i}) + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\cdots,j_l=1}^{L}\alpha_{k0}\alpha_k r_k\beta_{kj_1}\cdots\beta_{kj_l}\mathbb{E}(Y_{t-i})\\ &\quad+\sum_{l=0}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\cdots,j_{l+1}=1}^{L}\alpha_k r_k\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}\mathbb{E}(Y_{t-j_1-\cdots-j_{l+1}}Y_{t-i})\\ &= I + II\end{aligned}$$

where

$$\begin{aligned}II &= \sum_{l=0}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\cdots,j_{l+1}=1}^{L}\alpha_k r_k\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}\gamma_{|i-j_1-\cdots-j_{l+1}|,t}\\ &= \sum_{l=0}^{\infty}\sum_{k=1}^{K}\sum_{j_1+\cdots+j_{l+1}=i}^{L}\alpha_k r_k\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}\gamma_{0,t-i}\\ &\quad+\sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1+\cdots+j_{l+1}\neq i}^{L}\alpha_k r_k\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}\gamma_{|i-j_1-\cdots-j_{l+1}|,t}\\ &= \sum_{l=0}^{\infty}\sum_{k=1}^{K}\delta_{i0kl}\gamma_{0,t-i} + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{u=1}^{L-1}\delta_{iukl}\gamma_{u,t}\end{aligned}$$

with

$$\delta_{iukl} = \sum_{|i-j_1-\cdots-j_{l+1}|=u}\alpha_k r_k\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}.$$

Moreover, using the same notation, we get

$$\begin{aligned}I &= \left\{\sum_{k=1}^{K}\alpha_{k0}\alpha_k r_k + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\cdots,j_l=1}^{L}\alpha_{k0}\alpha_k r_k\beta_{kj_1}\cdots\beta_{kj_l}\right\}\mu\\ &= \left\{\sum_{l=0}^{\infty}\sum_{k=1}^{K}\sum_{j_1,\cdots,j_l=1}^{L}\alpha_{k0}\alpha_k r_k\beta_{kj_1}\cdots\beta_{kj_l}\right\}\mu =: K_1.\end{aligned}$$

Finally, for $i = 1,\cdots,L$,

$$\gamma_{i,t} = K_1 + \sum_{l=0}^{\infty}\sum_{k=1}^{K}\delta_{i0kl}\gamma_{0,t-i} + \sum_{l=1}^{\infty}\sum_{k=1}^{K}\sum_{u=1}^{L-1}\delta_{iukl}\gamma_{u,t}.$$

It is equivalent to

$$K_1 + \omega_{i0}\gamma_{0,t-i} + \sum_{u=1}^{L-1} \omega_{iu}\gamma_{u,t} = 0$$

where

$$\omega_{i0} = \sum_{l=0}^{\infty}\sum_{k=1}^{K}\delta_{i0kl}, \quad \omega_{ii} = \sum_{l=0}^{\infty}\sum_{k=1}^{K}\delta_{iikl} - 1 \text{ and } \omega_{iu} = \sum_{l=0}^{\infty}\sum_{k=1}^{K}\delta_{iukl} \text{ for } u \neq i.$$

Let $\Gamma = (\omega_{ij})_{i,j=1}^{L-1}$ and $\Gamma^{-1} = (b_{ij})_{i,j=1}^{L-1}$.
Then

$$B(\gamma_{1,t}, \cdots, \gamma_{L-1,t})' = -\left(K_1 + \omega_{10}\gamma_{0,t-1}, \cdots, K_1 + \omega_{(L-1)0}\gamma_{0,t-(L-1)}\right)'$$

which is equivalent to

$$(\gamma_{1,t}, \cdots, \gamma_{L-1,t})' = -B^{-1}\left(K_1 + \omega_{10}\gamma_{0,t-1}, \cdots, K_1 + \omega_{(L-1)0}\gamma_{0,t-(L-1)}\right)'.$$

We can show that

$$\gamma_{i,t} = -K_1 \sum_{u=1}^{L-1} b_{iu} - \sum_{u=1}^{L-1} b_{iu}\omega_{u0}\gamma_{0,t-u}.$$

The second moment is given by

$$\gamma_{0,t} = \mathbb{E}(Y_t) + \sum_{k=1}^{K}\alpha_k(r_k + r_k^2)\mathbb{E}(\lambda_{kt}^2).$$

For $k = 1, \cdots, K$, we have

$$\begin{aligned} \lambda_{kt}^2 &= \Big(\alpha_{k0} + \sum_{i=1}^{L}\alpha_{ki}Y_{t-i} + \sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)}\Big)\lambda_{kt} \\ &= \alpha_{k0}\lambda_{kt} + \sum_{i=1}^{L}\alpha_{ki}Y_{t-i}\lambda_{kt} + \sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)}\lambda_{kt}. \end{aligned}$$

The condition H_1 implies that the process $\{\lambda_{kt},\ t \in \mathbb{Z}\}$ is first-order stationary, for $k = 1, \cdots, K$. Hence,

$$\mathbb{E}(\lambda_{kt}) = \frac{\alpha_{k0} + \sum_{i=1}^{L}\alpha_{ki}\mu}{1 - \sum_{j=1}^{L}\beta_{kj}}, \quad \text{for} \quad k = 1, \cdots, K.$$

We deduce that

$$\begin{aligned}\mathbb{E}\Big(\sum_{i=1}^{L}\alpha_{ki}Y_{t-i}\lambda_{kt}\Big) &= \mathbb{E}\Big(C_{k0}\sum_{i=1}^{L}\alpha_{ki}Y_{t-i}+\sum_{i=1}^{L}\alpha_{ki}Y_{t-i}\sum_{l=0}^{\infty}\sum_{j_1,\cdots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}Y_{t-j_1-j_2-\cdots-j_{l+1}}\Big)\\ &= \mathbb{E}\Big(C_{k0}\sum_{i=1}^{L}\alpha_{ki}Y_{t-i}+\sum_{l=0}^{\infty}\sum_{j_1,\cdots,j_{l+2}=1}^{L}\alpha_{kj_{l+1}}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_l}Y_{t-j_1-j_2-\cdots-j_{l+1}}Y_{t-j_{l+2}}\Big)\\ &= C_{k0}\mu\sum_{i=1}^{L}\alpha_{ki}+\sum_{i=1}^{L}\Delta_{k,i}^{(1)}\gamma_{0,t-i}+\sum_{v=1}^{L-1}\Lambda_{kv}^{(1)}\gamma_{v,t}\end{aligned}$$

where

$$\Delta_{k,i}^{(1)} = \sum_{l=0}^{\infty}\sum_{\substack{j_{l+2}=i\\ j_{l+2}=j_1+\cdots+j_{l+1}}}^{L}\alpha_{kj_{l+1}}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_l}$$

and

$$\Lambda_{kv}^{(1)} = \sum_{l=0}^{\infty}\sum_{|j_{l+2}-j_1-\cdots-j_{l+1}|=v}^{L}\alpha_{kj_{l+1}}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_l}.$$

Moreover

$$\begin{aligned}\sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)} &= \sum_{j=1}^{L}\beta_{kj}\left\{C_{k0}+\sum_{l=0}^{\infty}\sum_{j_1,\cdots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}Y_{t-j-j_1-j_2-\cdots-j_{l+1}}\right\}\\ &= C_{k0}\sum_{j=1}^{L}\beta_{kj}+\sum_{l=0}^{\infty}\sum_{j_1,\cdots,j_{l+2}=1}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}Y_{t-j_1-j_2-\cdots-j_{l+2}}.\end{aligned}$$

Hence

$$\begin{aligned}\sum_{j=1}^{L}\beta_{kj}\lambda_{k(t-j)}\lambda_{kt} &= \left\{C_{k0}\sum_{j=1}^{L}\beta_{kj}+\sum_{l=0}^{\infty}\sum_{j_1,\cdots,j_{l+2}=1}^{L}\alpha_{kj_{l+2}}\beta_{kj_1}\cdots\beta_{kj_{l+1}}Y_{t-j_1-j_2-\cdots-j_{l+2}}\right\}\\ &\quad\times\left\{C_{k0}+\sum_{l=0}^{\infty}\sum_{j_1,\cdots,j_{l+1}=1}^{L}\alpha_{kj_{l+1}}\beta_{kj_1}\cdots\beta_{kj_l}Y_{t-j_1-j_2-\cdots-j_{l+1}}\right\}\end{aligned}$$

$$= C_{k0}^2 \sum_{j=1}^{L} \beta_{kj} + C_{k0} \sum_{j=1}^{L} \beta_{kj} \sum_{l=0}^{\infty} \sum_{j_1,\cdots,j_{l+1}=1}^{L} \alpha_{kj_{l+1}} \beta_{kj_1} \cdots \beta_{kj_l} Y_{t-j_1-j_2-\cdots-j_{l+1}}$$
$$+ \ C_{k0} \sum_{l=0}^{\infty} \sum_{j_1,\cdots,j_{l+2}=1}^{L} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_{l+1}} Y_{t-j_1-j_2-\cdots-j_{l+2}}$$
$$+ \sum_{\substack{l=0 \\ l'=0}}^{\infty} \sum_{\substack{j_1,\cdots,j_{l+2}=1 \\ j'_1,\cdots,j'_{l'+1}=1}}^{L} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_{l+1}} \alpha_{kj'_{l'+1}} \beta_{kj'_1} \cdots \beta_{kj'_{l'}} Y_{t-j_1-j_2-\cdots-j_{l+2}} Y_{t-j'_1-j'_2-\cdots-j'_{l'+1}}.$$

The term $\mathbb{E}\Big(\sum_{j=1}^{L} \beta_{kj} \lambda_{k(t-j)} \lambda_{kt} \Big)$ is given by

$$\mathbb{E}\Big(\sum_{j=1}^{L} \beta_{kj} \lambda_{k(t-j)} \lambda_{kt} \Big) = C_{k0}^2 \sum_{j=1}^{L} \beta_{kj} + 2C_{k0}\mu \sum_{l=0}^{\infty} \sum_{j_1,\cdots,j_{l+2}=1}^{L} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_{l+1}}$$
$$+ \sum_{i=1}^{L} \Delta_{k,i}^{(2)} \gamma_{0,t-i} + \sum_{v=1}^{L-1} \Lambda_{kv}^{(2)} \gamma_{v,t}$$

where

$$\Delta_{k,i}^{(2)} = \sum_{\substack{l=0 \\ l'=0}}^{\infty} \sum_{\substack{j_1+\cdots+j_{l+2}=i \\ j_1+\cdots+j_{l+2}=j'_1+\cdots+j'_{l'+1}}}^{L} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_{l+1}} \alpha_{kj'_{l'+1}} \beta_{kj'_1} \cdots \beta_{kj'_{l'}}$$
$$\Lambda_{kv}^{(2)} = \sum_{\substack{l=0 \\ l'=0}}^{\infty} \sum_{|j_1+\cdots+j_{l+2}-j'_1-\cdots-j'_{l'+1}|=v}^{L} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_{l+1}} \alpha_{kj'_{l'+1}} \beta_{kj'_1} \cdots \beta_{kj'_{l'}}.$$

Let us denote $\Delta_{k,i} = \Delta_{k,i}^{(1)} + \Delta_{k,i}^{(2)}$ and $\Lambda_{kv} = \Lambda_{kv}^{(1)} + \Lambda_{kv}^{(2)}$.

For $k = 1, \cdots, K$, the expectation of λ_{kt}^2 is given by

$$\mathbb{E}(\lambda_{k,t}^2) = C_k + \sum_{i=1}^{L} \Delta_{k,i} \gamma_{0,t-i} + \sum_{v=1}^{L-1} \Lambda_{kv} \gamma_{v,t}$$

with

$$C_k = \alpha_{k0} \mathbb{E}(\lambda_{kt}) + C_{k0} \sum_{i=1}^{L} \alpha_{ki} \mu + C_{k0}^2 \sum_{j=1}^{L} \beta_{kj} + 2C_{k0} \sum_{l=0}^{\infty} \sum_{j_1,\cdots,j_{l+2}=1}^{L} \alpha_{kj_{l+2}} \beta_{kj_1} \cdots \beta_{kj_{l+1}} \mu.$$

Then

$$\begin{aligned}\gamma_{0,t} &= \mu+\sum_{k=1}^{K}\alpha_k(r_k+r_k^2)\left(C_k+\sum_{i=1}^{L}\Delta_{k,i}\gamma_{0,t-i}+\sum_{v=1}^{L-1}\Lambda_{kv}\gamma_{v,t}\right)\\ &= \mu+\sum_{k=1}^{K}\alpha_k(r_k+r_k^2)\left[C_k+\sum_{u=1}^{L}\Delta_{k,u}\gamma_{0,t-u}+\sum_{v=1}^{L-1}\Lambda_{kv}\left(-K_1\sum_{u=1}^{L-1}b_{vu}-\sum_{u=1}^{L-1}b_{vu}\omega_{u0}\gamma_{0,t-u}\right)\right]\\ &= c_0+\sum_{k=1}^{K}\alpha_k(r_k+r_k^2)\left[\sum_{u=1}^{L}\Delta_{k,u}\gamma_{0,t-u}-\sum_{u=1}^{L-1}\left(\sum_{v=1}^{L-1}\Lambda_{kv}b_{vu}\omega_{u0}\right)\gamma_{0,t-u}\right]\end{aligned}$$

where

$$c_0 = \mu+\sum_{k=1}^{K}\alpha_k(r_k+r_k^2)\left[C_k-K_1\sum_{v=1}^{L-1}\Lambda_{kv}\sum_{u=1}^{L-1}b_{vu}\right].$$

Hence

$$\gamma_{0,t} = c_0+\sum_{k=1}^{K}\alpha_k(r_k+r_k^2)\left[\sum_{u=1}^{L-1}\left(\Delta_{k,u}-\sum_{v=1}^{L-1}\Lambda_{kv}b_{vu}\omega_{u0}\right)\gamma_{0,t-u}+\Delta_{k,L}\gamma_{0,t-L}\right] \tag{4}$$

Let

$$c_u = \sum_{k=1}^{K}\alpha_k(r_k+r_k^2)\left(\Delta_{k,u}-\sum_{v=1}^{L-1}\Lambda_{kv}b_{vu}\omega_{u0}\right),\ u=1,\cdots,L-1 \text{ and } c_L=\sum_{k=1}^{K}\alpha_k(r_k+r_k^2)\Delta_{k,L}.$$

Then the equation (4) is equivalent to

$$\gamma_{0,t} = c_0+\sum_{u=1}^{L}c_u\gamma_{0,t-u}. \tag{5}$$

A necessary and sufficient condition for the process to be second-order stationary is that all roots of $1-c_1Z^{-1}-c_2Z^{-2}-\cdots-c_LZ^{-L}=0$ lie inside the unit circle. □

For illustration, we consider in the following corollary the NBINARCH model corresponding to the NB-MINGARCH model with $K=1$ and $\beta_{1j}=0$ for all $j=1,\ldots,q_1$.

Corollary 3. *A necessary and sufficient condition for the process to be second-order stationary is that all roots of* $1-c_1Z^{-1}-c_2Z^{-2}-\cdots-c_LZ^{-p}=0$ *lie inside the unit* *circle, where*

$$\begin{aligned} c_u &= (r_1+r_1^2)\left(\alpha_{1u}-\sum_{v=1}^{p-1}\sum_{|i-j|=v}\alpha_{1i}\alpha_{1j}b_{vu}\omega_{u0}\right),\ u=1,\cdots,p-1 \textit{ and} \\ c_p &= (r_1+r_1^2)\alpha_{1p} \\ \textit{and } \Gamma &= (\omega_{ij})_{i,j=1}^{p-1},\ \Gamma^{-1}=(b_{ij})_{i,j=1}^{p-1},\ \textit{two matrices such that} \\ \omega_{uu} &= r_1\sum_{|i-u|=u}\alpha_{1i}-1 \textit{ and } \omega_{lu}=r_1\sum_{|i-u|=u}\alpha_{1i},\ u\neq l. \end{aligned}$$

If the process $\{Y_t\}_{t\in\mathbb{Z}}$ following a NB-MINGARCH$(K;r_1,\cdots,r_K;p_1,\cdots,p_K;q_1,\cdots,q_K)$ model is second-order stationary, then from (5), we have

$$\gamma_{0,t}=\frac{c_0}{1-\sum_{u=1}^{L}c_u}.$$

Hence, necessary second order stationary condition for a special case is given by the following proposition.

Proposition 2. *For the NB-MINGARCH$(K;r_1,\cdots,r_K;1,\cdots,1;1,\cdots,1)$ model, the second order stationary condition is:*

$$\beta_{k1}<1,\ k=1,\cdots,K \ \ and \ \sum_{k=1}^{K}\alpha_k(r_k+r_k^2)\alpha_{k1}^2<1.$$

Remark 3. As a special case, a necessary condition for the NB-MINGARCH$(1;r;1;1)$ for the second-order stationarity is: $\beta_{k1}<1,\ k=1,\cdots,K \ \ and \ \ (r+r^2)\alpha_{11}^2<1$.

In the following proposition, we give a necessary and sufficient condition of the second-order stationarity for the NB-MINGARCH$(K;r_1,\cdots,r_K;2,\cdots,2;1,\cdots,1)$ model.

Proposition 3. *Let $(Y_t)_{t\in\mathbb{Z}}$ be a NB-MINGARCH$(K;r_1,\cdots,r_K;2,\cdots,2;1,\cdots,1)$ model. Assume that the conditions H_1 and H_2 hold. If the process $(Y_t)_{t\in\mathbb{Z}}$ is first-order stationary then a necessary and sufficient condition for the process to be* *second-order stationary is:* $c_1 + c_2 < 1$ *where*

$$c_1 = \sum_{k=1}^{K} \alpha_k (r_k + r_k^2) \alpha_{k1}^2 + \frac{\left(\sum\limits_{k=1}^{K} \alpha_k (r_k + r_k^2) \Lambda_{k1}\right)\left(\sum\limits_{k=1}^{K} \alpha_k r_k \alpha_{k1}\right)}{1 - \sum\limits_{k=1}^{K} \alpha_k r_k \left(\alpha_{k1}\beta_{k1} + \alpha_{k2}\right)}$$

and

$$c_2 = \sum_{k=1}^{K} \alpha_k (r_k + r_k^2) \left(\alpha_{k2}^2 + 2\alpha_{k1}\alpha_{k2}\beta_{k1} + \alpha_{k1}^2 \beta_{k1}^2\right)$$

with

$$\Lambda_{k1} = 2\alpha_{k1}\alpha_{k2} + 2\alpha_{k1}^2\beta_{k1} + 2\alpha_{k2}^2\beta_{k1} + 2\alpha_{k1}^2\beta_{k1}^3 + 3\alpha_{k1}\alpha_{k2}\beta_{k1}^2 + 3\alpha_{k1}\alpha_{k2}\beta_{k1}^4 + \alpha_{k1}^2\beta_{k1}^5 + \alpha_{k1}\alpha_{k2}\beta_{k1}^6, \ k = 1, \cdots, K.$$

3. Estimation procedure and simulation results

In this section, we use the expectation-maximization (EM) algorithm to estimate the parameters. Suppose that the n-dimensional vectors of observations $Y = (Y_1, \cdots, Y_n)$ is generated from the NB-MINGARCH$(K; p_1, \ldots, p_K; q_1, \ldots, q_K; r_1, \ldots, r_K)$ model and the number of failures r_k is known, for $k = 1, \ldots, K$. Let $Z = (Z_1, \cdots, Z_n)$ be the random matrix variable where $Z_t = (Z_{1,t}, \cdots, Z_{K,t})^T$ is a vector whose components are defined by:

$$Z_{i,t} = \begin{cases} 1 & \text{if } Y_t \text{ comes from the } i-\text{th component; } 1 \leq i \leq K, \\ 0 & \text{otherwise.} \end{cases}$$

The vectors Z_t are not observed and its distribution is

$$\mathbb{P}(Z_t = (1, 0, \cdots, 0)^T) = \alpha_1, \cdots, \mathbb{P}(Z_t = (0, 0, \cdots, 0, 1)^T) = \alpha_K.$$

Let $\alpha = (\alpha_1, \cdots, \alpha_{K-1})^T$, $\alpha_{(k)} = (\alpha_{k0}, \alpha_{k1}, \cdots, \alpha_{kp_k})^T$, $\beta_{(k)} = (\beta_{k1}, \cdots, \beta_{kq_k})^T$ $\theta_{(k)} = (\alpha_{(k)}^T, \beta_{(k)}^T)$ and $\theta = (\alpha, \theta_{(1)}, \cdots, \theta_{(K)})^T \in \Theta$ (The parameters space).

Given Z_t, the distribution of the complete data (Y_t, Z_t) is then given by

$$\prod_{k=1}^{K} \left(\alpha_k \binom{Y_t + r_k - 1}{r_k - 1} p_{kt}^{r_k} (1 - p_{kt})^{Y_t} \right)^{Z_{kt}}$$

and the log-likelihood function at time t is given by

$$l_t = \sum_{k=1}^{K} Z_{kt} \left\{ \log(\alpha_k) - (r_k + Y_t) \log(1 + \lambda_{kt}) + Y_t \log(\lambda_{kt}) + \log(\binom{Y_t + r_k - 1}{r_k - 1}) \right\}.$$

The conditional log-likelihood is then given by

$$l^*(\theta) = \sum_{t=L+1}^{n} \sum_{k=1}^{K} Z_{kt} \left\{ \log(\alpha_k) - (r_k + Y_t)\log(1+\lambda_{kt}) + Y_t \log(\lambda_{kt}) + \log \binom{Y_t + r_k - 1}{r-1} \right\}.$$

The estimates of the parameters are obtained by iterating the two steps (E-step and M-step) of the EM algorithm until convergence. These steps are described in Diop *et al.* (2016).

Now, we investigate the performance of the EM estimation method by using Monte Carlo simulations. We use 100 independent realizations of the model (1) with sizes $n = 100$, $n = 200$ and $n = 500$. The model used is a NB-MINGARCH$(2; 1, 1; 1, 1; 2, 3)$ with parameter values

$$\begin{pmatrix} \alpha_1 & \alpha_{10} & \alpha_{11} & \beta_{11} \\ \alpha_2 & \alpha_{20} & \alpha_{21} & \beta_{21} \end{pmatrix} = \begin{pmatrix} 0.65 & 0.20 & 0.20 & 0.50 \\ 0.35 & 0.30 & 0.30 & 0.40 \end{pmatrix}.$$

The random initialization method is employed in this paper. Although it is probably the most commonly way employed for initiating the EM algorithm, this method has a limitation in that it is known to converge surely to a local maximum but not always to a global maximum. In practice, an extension consists of repeating it several times from different initial values and selecting the solution that maximizes the likelihood among these values. The estimation of the number of failures r_k is carried out by adopting the approach of Benjamin *et al.* (2003) and Davis and Wu (2009); that is, by maximizing the likelihood with respect to the parameter for different r_k values. The performances of the estimators are evaluated by the mean square error (RMSE) and the mean absolute error (MAE).

The results in Table 1 show that the proposed estimation method gives small mean square error and reasonably small mean absolute error. The performance of the estimate improves when the sample size increases. But these results are less satisfactory for the parameters α_{k0} and r_k. In fact, the obtained values of MSE and MAE for α_{k0} and r_k are a little bit high and that the convergence of estimators to the true value is slow. Similar results are observed in Diop *et al.* (2016) and Zhu *et al.* (2010) in regards to parameter α_{k0}. So we will have to seek to develop a method that will help to estimate only the two parameters (α_{k0}, r_k) so as to improve the quality of the estimation of the NB-MINGARCH model .

4. Real data application

In this section, we shall investigate the computer aided dispatch (CAD) calls data, which are more representative of the volume and extent of crimes that do not have victims, such as drug dealing, prostitution, and gambling than offense or arrest records. We restrict our attention to the time series representing a count of CAD drug calls reported in the 22nd police car beat in Pittsburgh, during one month. There are 144 available observations that represent the number of CAD

Table 1. *Results of the simulation study.*

Sample size	k		α_k	α_{k0}	α_{k1}	β_{k1}	r_k
100	1	True values	0.6500	0.2000	0.2000	0.5000	2.0000
		Mean estimated	0.5467	0.3286	0.2005	0.4854	2.3251
		MSE	0.0329	0.1200	0.0124	0.0535	0.1873
		MAE	0.1250	0.2271	0.0878	0.1859	0.3363
	2	True values	0.3500	0.3000	0.3000	0.4000	3.0000
		Mean estimated	0.4533	0.4391	0.2785	0.3712	3.3097
		MSE	0.0329	0.1867	0.0154	0.0532	0.2140
		MAE	0.1250	0.2856	0.0937	0.1816	0.3705
200	1	True values	0.6500	0.2000	0.2000	0.5000	2.0000
		Mean estimated	0.6276	0.3302	0.1963	0.4852	2.3031
		MSE	0.0145	0.0956	0.0045	0.0234	0.1545
		MAE	0.0873	0.1929	0.0534	0.1101	0.3031
	2	True values	0.3500	0.3000	0.3000	0.4000	3.0000
		Mean estimated	0.3724	0.4259	0.2988	0.3747	3.3002
		MSE	0.0145	0.1724	0.0129	0.0246	0.1975
		MAE	0.0873	0.2541	0.0676	0.1080	0.3624
500	1	True values	0.6500	0.2000	0.2000	0.5000	2.0000
		Mean estimated	0.6597	0.2195	0.2031	0.5044	2.2867
		MSE	0.0107	0.0216	0.0024	0.0074	0.1430
		MAE	0.0649	0.1042	0.0359	0.0667	0.2867
	2	True values	0.3500	0.3000	0.3000	0.4000	3.0000
		Mean estimated	0.3403	0.3277	0.3105	0.3841	3.2966
		MSE	0.0107	0.0936	0.0100	0.0159	0.1562
		MAE	0.0649	0.1826	0.0600	0.0945	0.3039

drug calls of 12 years (from 1990 through 2001). The data are available online at the forecasting principles site (http://www.forecastingprinciples.com), in the section about crime data. The mean and the variance are estimated as 6.3056 and 22.0249, respectively. Hence the data seem to be overdispersed. The histogram (in Fig. 1) of the data and the bimodality index of Der and Everitt (2002) used in Zhu *et al.* (2010) show that the series seem to be bimodal. Zhu *et al.* (2010) have shown that the two-components mixture MINARCH model is more appropriate for this dataset than the INARCH model. However, the results of Diop *et al.* (2016) indicate that the MINGARCH (with two components) model should be preferred to the MINARCH for this dataset.

We will investigate this count data by fitting a NB-MINGARCH model to the series. The problem of model selection requires three aspects. First, we must select the number of components K. Second, we must estimate the number of failures r_k. Three, the model identification problem needs to be addressed (i.e the *AR* polynomial order, p_k, and the *MA* polynomial order, q_k). The selection of K and r_k, $k = 1, \ldots, K$ is more important because it will affect the interpretation of the model and the estimation of the orders is dependent on the selected number of components. This estimation is carried out by using some information criteria. In this paper, three criteria are considered: the Akaike information criterion (AIC), the Bayesian information criterion (BIC) and the mixture regression criterion (MRC) proposed by Naik *et al.* (2007). We consider a NB-MINGARCH model with $K = 1, 2$. To simplify, for $K = 1, 2$, consider the model with: $p_1 = \cdots = p_K = p$, $q_1 = \cdots = q_K = q$ and $r_1 = \cdots = r_K = r$, where $1 \leq p \leq 3$, $0 \leq q \leq 3$ and $r = 1, \cdots, 15$. The number of failures and the order of the components are selected to be that minimizing the criteria. Table 2 indicates the results obtained. For each criterion, the minimum is represented by the underlined value.

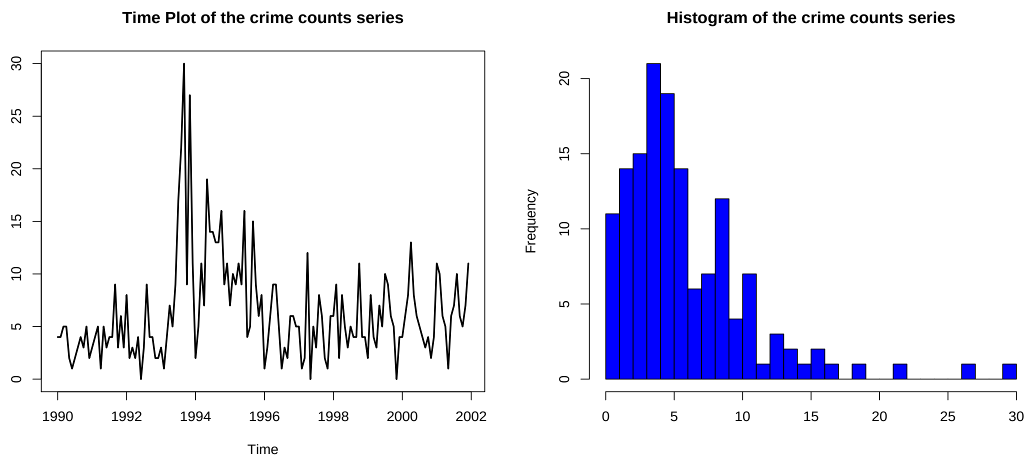

Fig. 1. *Crime counts series: the time plot and the histogram.*

Table 2 shows that the BIC and the MRC retain the single-component mixture model, but this does not agree with the histogram and the bimodality index. It is not surprising since these criterion penalize more than the AIC. The AIC suggests the two-component mixture model respectively with $(p, q) = (1, 3)$, which confirm the bimodality observed in the histogram. These results are in concordance with those obtained by Diop *et al.* (2016) and Zhu *et al.* (2010). In addition, the values of

Table 2. *AIC, BIC and MRC values for the crime counts series, $K = 1, 2$.*

	Order	AIC $p=1$	AIC $p=2$	AIC $p=3$	BIC $p=1$	BIC $p=2$	BIC $p=3$	MRC $p=1$	MRC $p=2$	MRC $p=3$
	$q=0$	762.25	758.57	760.34	771.14	770.40	775.09	549.43	546.63	545.87
$K=1$	$q=1$	759.00	759.71	762.34	770.84	774.45	780.03	550.60	548.57	547.84
$\widehat{r}=5$	$q=2$	760.91	762.02	764.02	775.68	779.75	784.66	547.98	550.16	548.45
	$q=3$	757.45	763.17	763.27	775.13	783.81	786.86	545.91	549.75	546.83
			(807.35)			(822.41)			(545.87)	
	$q=0$	765.65	755.25	754.01	786.39	775.94	780.55	716.02	730.81	729.87
$K=2$	$q=1$	753.83	751.60	749.64	774.57	778.21	782.08	739.64	727.22	722.46
$\widehat{r}=12$	$q=2$	751.81	754.81	753.25	778.54	787.33	797.58	733.61	724.13	728.71
	$q=3$	749.13	752.25	757.06	781.57	790.58	801.29	729.96	724.38	728.70
			(751.82)			(780.92)			(537.59)	

The values in the brackets $(\cdot)$ represent the minimums given by the MINGARCH models for each criterion (see also Diop *et al.* (2016)).

the AIC and the BIC obtained for the NB-MINGARCH model are globally better than those of the MINGARCH model. Thus, we can conclude that the NB-MINGARCH model is more appropriate for this dataset than the MINGARCH model.

5. Concluding remarks

In this work, a new class of time series models which generalize the NBIN-GARCH model to a negative binomial mixture integer-valued GARCH model (NB-MINGARCH) is proposed. It is potentially useful in modeling integer-valued time series with greater degree of overdispersion, multimodality and non-stationary components. The estimation of the parameters is done by using the EM algorithm. The model selection can be done with three foregoing information criteria. However, many problems for this class of models remain open. The properties of the ergodicity can be established for these models. It would also be interesting to study the properties of the estimators for the parameters.

Acknowledgments

We wish to thank the referee as well as the associate editor for helpful comments and suggestions, which have permitted to improve the presentation of this paper and the editors of highly professional handling of the publication process.

References

Al-Osh, M.A. and Alzaid, A.A. (1987). First order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.*, 8, 261-275.

Al-Osh, M.A. and Alzaid, A.A. (1990). Integer-valued pth-order autoregressive structure (INAR(p)) process. . *J. Appl. Probab.*, 27(2), 314-324.

Benjamin, M.A., Rigby, R.A. and Stasinopoulos, D.M. (2003). Generalized autoregressive moving average models. *Journal of the American Statistical association*, 98, 214-223.

Davis, A. and Wu, R. (2009). A negative binomial model for time series of counts. *Biometrika*, 96, 735-49.

Davis, A., Liu, H. (2012). Theory and inference for a class of observation-driven models with application to time series of counts. *Preprint, arXiv:1204.3915*.

Der, G. and Everitt, B.S. (2002). *A Handbook of Statistical Analysis using SAS, second ed.* Chapman & Hall/CRC, London. ISBN 1-5848-8245-X

Diop, M.L., Diop, A. and Diongue, A.K. (2016). A mixture integer-valued GARCH model. *Revstat Statistical Journal*, 14(3), 245-271.

Ferland, R., Latour, A. and Oraichi, D. (2006). Integer-valued GARCH process. *J. Time Ser. Anal.*, 27, 923-42.

Fokianos, K., Neumann, M. (2013). A goodness-of-fit test for Poisson count processes. *Electronic Journal of Statistics,* 7, 793-819.

Fried, R., Agueusop I., Bornkamp, B., Fokianos, K., Fruth, J. and Ickstadt, K. (2015). Retrospective Bayesian outlier detection in INGARCH series. *Statistics & Computing*, 25, 365-374.

Goldberg, S. (1958). *Introduction to Difference Equations*. New York: Wiley. ISBN 978-0-387-27602-1.

Hoef, J.M. and Boveng, P.L. (2007). Quasi-poisson vs. negative binomial regression: how should we model overdispersed count data ? *Ecology*, 88(11), 2766-2772.

Jacobs, P.A. and Lewis, P.A.W. (1978a). Discrete time series generated by mixture I: correlational and runs properties. *J. Roy. Statist. Soc.*, Ser B 40(1), 94-105.

Jain, G.C. and Consul, P.C. (1971). A generalized negative binomial distribution. *SIAM Journal on Applied Mathematics*, 21(4), 501-513.

Jalali, A. and Pemberton, J. (1995). Mixture models for time series. *J. Appi. Prob.*, 32, 123-138.

Kachour, M. and Yao, J.F. (2009). First-Order Rounded Integer-Valued Autoregressive (RINAR(1)) Process. *J. Time Ser. Anal.*, 30, 417-448.

Liden, A. and Mantyniemi, S. (2011). Using the negative distribution to model overdispersion in ecological count data. *Ecological Society of America*, 92(7), 1414-1421.

Mclachlan, G. and Peel, D. (2000). *Finite Mixture Models*. New York: Wiley. ISBN: 978-0-471-00626-8.

McKenzie, E. (1985). Simple models for discrete variate time series. *Water Resources Bulletin*, 22, 645-650.

Molla, D.T. and Muniswamy, B. (2012). Power of Tests for Negative Binomial Regression Coefficients in Count Data. *Int. J. Math.*, 3(8), 3136-3142.

Moysiadis T. and Fokianos K. (2014). On binary and categorical time series models with feedback. *Journal of Multivariate Analysis*, 131, 209-228.

Naik, P.A., Shi, P. and Tsai, C.L. (2007). Extending the Akaike information criterion to mixture regression models. *Journal of the American Statistical Association*, 120, 244-254.

Neumann M. (2011). Absolute regularity and ergodicity of Poisson count processes. *Bernoulli*, 1268- 1284.

Steutel, F.W. and Van Harn, K. (1979). Discrete analogues of self decomposability and stability. *Ann. Probab.*, 7, 893-899.

Zhu, F., Li, Q. and Wang, D. (2010). A mixture integer-valued ARCH model. *J. Stat. Plann. Infer.*, 140, 2025-2036.

Zhu, F. (2011). A negative binomial integer-valued GARCH model. *J. Time Ser. Anal.*, 32, 54-67.

Zhu, F. (2012). Modeling overdispersed or underdispersed count data with generalized poisson integer-valued GARCH models. *J. Math. Anal. Appl.*, 389, 58-71.

Zhu, F. (2012). Zero-inflated Poisson and negative binomial integer-valued GARCH models. *Journal of Statistical Planning and Inference*, 142, 826-839.