FACTA UNIVERSITATIS (NIŠ)
Ser. Math. Inform. Vol. 39, No 1 (2024), 99–122
https://doi.org/10.22190/FUMI230203007P
Original Scientific Paper

A MIXTURE INTEGER-VALUED AUTOREGRESSIVE MODEL WITH A STRUCTURAL BREAK

Predrag M. Popović[1], Miroslav M. Ristić[2] and Milena S. Stojanović[2]

[1] **University of Niš, Faculty of Civil Engineering and Architecture**
Department of Mathematics, Physic and Informatics
Aleksandra Medvedeva 14, 18000 Niš, Serbia
[2] **University of Niš, Faculty of Sciences and Mathematics**
Department of Mathematics, Višegradska 33, 18000 Niš, Serbia

Abstract. In this manuscript we introduce a mixture integer-valued autoregressive model with a structural break. The introduced model is a mixture of an INAR(1) model with the binomial thinning operator and an INAR(1) model with the negative binomial thinning operator. Some properties of the introduced model are derived. The unknown parameters of the model are estimated by some methods and the performances of the obtained estimators are checked by simulations. At the end of the paper, two possible applications of the model are provided and discussed.
Keywords: Binomial thinning, Integer-valued autoregressive model, Mixture of INAR models, Structural break, Negative binomial thinning.

1. Introduction

In this manuscript we introduce a mixture integer-valued autoregressive model with a structural break motivated by the following real examples. We observe the number of infected people from a virus. At first, the activity of the virus is low and up to the moment τ, the number of infected people from the virus behaves like an INAR(1) model with the binomial thinning operator [13, 1]. After the moment τ, when a large number of reproductions are created, the activity of

Received February 03, 2023, revised: May 24, 2023, accepted: May 24, 2023
Communicated by Aleksandar Nastić
Corresponding Author: Predrag M. Popović. E-mail addresses: popovicpredrag@yahoo.com (P. Popović), miristic72@gmail.com (M. Ristić), milena_aleksic93@yahoo.com (M. Stojanović)
2010 *Mathematics Subject Classification.* Primary 62M10

the virus increases significantly and a greater number of people can be infected. Now, the number of infected people can be described as an INAR(1) model with the negative binomial thinning operator [14]. Also, we can observe the number of criminal offenses in some places. In the beginning, criminal groups are poorly organized and therefore, the number of criminal offenses can be described by an INAR(1) model with the binomial thinning operator. After some time criminal groups are better organized and their activities become more frequently. Therefore, after a moment τ, the number of crimes can be represented by an INAR(1) model with the negative binomial thinning operator.

Regarding these real examples, in this manuscript we introduce an integer-valued autoregressive model with one structural break τ based on two thinning operators, the binomial thinning operator " $\circ$ " and the negative binomial thinning operator " $*$ ". Exactly, we consider a time series model $\{X_t\}$, $t \in \mathcal{Z} \equiv \{0, \pm 1, \pm 2, \dots\}$, given by

$$X_t = \begin{cases} \alpha \circ X_{t-1} + \varepsilon_t, & t \leqslant \tau, \\ \beta * X_{t-1} + \varepsilon_t, & t > \tau, \end{cases} \tag{1.1}$$

where α and β belong to $(0, 1)$, τ is an integer, all the counting series incorporated in $\alpha\circ$ and $\beta*$ are mutually independent sequences of independent and identically distributed (i.i.d.) random variables with Bernoulli(α), and Geometric($\beta/(1+\beta)$) distributions, respectively. Here, Bernoulli(α) indicates to a Bernoulli distributed random variable with success α and Geometric($\beta/(1+\beta)$) indicates to a geometric distributed random variable with mean β. We also suppose that $\{\varepsilon_t\}$ is a sequence of independent random variables, all the counting series are independent of X_t and ε_s for all t and s, and X_t and ε_s are independent for all $t < s$. The distribution of the random variable X_t can be different over the times $t \leqslant \tau$ and $t > \tau$, so the model given by (1.1) can be a non-stationary model. It is possible that the distribution of the random variable X_t is the same over all times $t \leqslant \tau$ and $t > \tau$ and in this case we have a process stationary in distribution. This type of process is also stationary in mean (a first-order stationary process) and it can be also a second-order stationary process only when the thinning parameters α and β are identical. In all other cases it is a non-stationary model, but stationary separately on each regime, i.e. it is stationary before and after a structural break.

The introduced model is an integer-valued autoregressive model with one structural break. Some other forms of these models have been widely investigated in the past. Also, special attentions have been given to development of the methods for detections of the breaks or the changes in these models. A comprehensive list of important references in detections of breaks and changes in time series can be found in [6] and [2]. Kashikar et al. [11] introduced an INAR model of the first- and higher-order with structural breaks. In their model, different binomial thinning operators have been used in each regime to generate values of the considered time series model. The authors have supposed that the innovations are independent Poisson distributed random variables with identical parameters inside each regime and with different parameters between different regimes. Hudecová [9] considered autoregressive binary time series with changes and has proposed method for detection of changes. Yu et al. [15] introduced an integer-valued moving average model with structural breaks and researched its properties. Hudecová et al. [10] developed procedures based on the probability generating function for detections of the changes in the integer-valued autoregressive model of the first order. Chen and Lee [4] introduced a zero-inflated generalized Poisson autoregressive models with structural breaks. Recently, Kim and Lee [12] introduced a residual-based CUSUM test for the PINAR(1) model which can be used as an alternative to classical CUSUM tests, while Cui and Wu [5] considered how to detect parameter changes in observation-driven models for count time series.

The manuscript is organized as follows. In Section 2, we introduce a first-order integer-valued autoregressive model as a mixture of two INAR(1) models based on two well-known thinning operators, the binomial and the negative binomial thinning operator. Some properties of the model with geometric marginals including conditional properties and correlation structure are derived. Section 3 covers some estimation issues. Here, we consider two methods of estimations, the conditional maximum likelihood method and the method of the conditional least squares. The performances of the obtained estimates are checked by simulations for different true values of the parameters, different sample sizes and different positions for structural break. In Section 4, we discuss possible applications of the introduced model on two real data sets about some criminal acts. The manuscript ends with some concluding remarks and discussion about further developments related to the introduced model and its generalization.

2. Construction and properties

In this section we derive some properties of the model introduced in the introduction. First, we start with the definition of the model in general case.

Definition 1. Suppose that α and β are real numbers from $(0,1)$, τ is an integer, all the counting series incorporated in $\alpha\circ$ and $\beta *$ are mutually independent sequences of i.i.d. random variables with Bernoulli(α) and Geometric($\beta/(1+\beta)$) distributions. Also, suppose that $\{\varepsilon_t\}$ is a sequence of independent random variables, all the counting series are performed independently of X_t and ε_s for all t and s, and the random variables X_t and ε_s are independent for all $t < s$. A process $\{X_t\}$ is said to be a mixture integer-valued autoregressive model with a structural break if it satisfies equation (1.1) for all $t \in \mathcal{Z}$.

Now, we consider a case of the mixture integer-valued autoregressive model with a structural break under the assumption that the random variable X_t has the geometric distribution as follows. For $t \leqslant \tau$ we suppose that X_t has Geometric($\mu_1/(1+\mu_1)$) distribution and for $t > \tau$ we suppose that X_t has Geometric($\mu_2/(1+\mu_2)$) distribution. Both parameters μ_1 and μ_2 are positive real numbers. Thus, our model has 5 parameters: 2 thinning parameters α and β; 2 mean parameters μ_1 and μ_2, and 1 structural break parameter τ. The number of the unknown parameters can be reduced. First, it is possible that the random variable X_t has the same distribution over all times, so in this case we have that the parameters μ_1 and μ_2 are identical. Also, the thinning parameters α and β can be identical too. Thus, the number of unknown parameters can be reduced to 3 unknown parameters. As mentioned above, in the case when we have identical thinning parameters and identical mean parameters, our model is reduced to a second-order stationary process with same marginal distributions.

Our model is completely determined if the distribution of the random variable ε_t is known. Regarding this, the following theorem gives its distribution.

Theorem 1. *Let $\{X_t\}$, $t \in \mathcal{Z}$, be the mixture integer-valued autoregressive model with a structural break given by (1) and let us suppose that X_t has Geom$(\mu_1/(1+\mu_1))$ distribution for $t \leq \tau$, and Geom$(\mu_2/(1+\mu_2))$ distribution for $t > \tau$. If $\alpha \in (0,1)$, $\beta \in \Big(0, \min\{\frac{\mu_2}{1+\mu_2}, \frac{\mu_2}{1+\mu_1}\}\Big]$, $\mu_1 > 0$, $\mu_2 > 0$, and $\tau \in \mathcal{Z}$, then the random variable ε_t is distributed as follows:*

$$\varepsilon_t \stackrel{d}{=} \begin{cases} 0, & w.p. \quad \alpha, \\ Geom\big(\frac{\mu_1}{1+\mu_1}\big), & w.p. \quad 1-\alpha, \end{cases} \qquad if\ t \leq \tau, \tag{2.1}$$

$$\varepsilon_t \stackrel{d}{=} \begin{cases} Geom\big(\frac{\beta}{1+\beta}\big), & w.p. \quad \frac{\beta\mu_1}{\mu_2-\beta}, \\ Geom\big(\frac{\mu_2}{1+\mu_2}\big), & w.p. \quad 1-\frac{\beta\mu_1}{\mu_2-\beta}, \end{cases} \qquad if\ t = \tau+1, \tag{2.2}$$

$$\varepsilon_t \stackrel{d}{=} \begin{cases} Geom\big(\frac{\beta}{1+\beta}\big), & w.p. \quad \frac{\beta\mu_2}{\mu_2-\beta}, \\ Geom\big(\frac{\mu_2}{1+\mu_2}\big), & w.p. \quad 1-\frac{\beta\mu_2}{\mu_2-\beta}, \end{cases} \qquad if\ t \geqslant \tau+2. \tag{2.3}$$

Proof. Let Φ_{X_t}, $\Phi_{X_{t-1}}$ and Φ_{ε_t} be the probability generating functions of the random variables X_t, X_{t-1} and ε_t, respectively. Let us first consider the case $t \leqslant \tau$. In this case, we have that both random variables X_t and X_{t-1} have Geom$\big(\frac{\mu_1}{1+\mu_1}\big)$ distributions. Thus, from (1.1) we have that

$$\Phi_{X_t}(s) = \Phi_{X_{t-1}}(1-\alpha+\alpha s)\Phi_{\varepsilon_t}(s).$$

We obtain that

$$\Phi_{\varepsilon_t}(s) = \frac{\Phi_{X_t}(s)}{\Phi_{X_{t-1}}(1-\alpha+\alpha s)} = \frac{1+\alpha\mu_1-\alpha\mu_1 s}{1+\mu_1-\mu_1 s} = \alpha + (1-\alpha)\cdot\frac{1}{1+\mu_1-\mu_1 s}.$$

The function $\Phi_{\varepsilon_t}(s)$ is well defined for $\alpha \in (0,1)$. Thus, we obtain that the random variable ε_t is a mixture of 0 with probability α and the random variable which has geometric distribution Geom$\big(\frac{\mu_1}{1+\mu_1}\big)$ with probability $1-\alpha$.

If $t = \tau+1$, then X_t has a geometric distribution Geom$(\mu_2/(1+\mu_2))$ and X_{t-1} has a geometric distribution Geom$(\mu_1/(1+\mu_1))$. Thus,

$$\Phi_{X_t}(s) = \Phi_{X_{t-1}}\left(\frac{1}{1+\beta-\beta s}\right)\Phi_{\varepsilon_t}(s)$$

and we get

$$\Phi_{\varepsilon_t}(s) = \frac{\Phi_{X_t}(s)}{\Phi_{X_{t-1}}(\frac{1}{1+\beta-\beta s})} = \frac{1+\beta(1+\mu_1)-\beta(1+\mu_1)s}{(1+\beta-\beta s)(1+\mu_2-\mu_2 s)}$$
$$= \frac{\beta\mu_1}{\mu_2-\beta}\cdot\frac{1}{1+\beta-\beta s} + \left(1-\frac{\beta\mu_1}{\mu_2-\beta}\right)\cdot\frac{1}{1+\mu_2-\mu_2 s}.$$

The function $\Phi_{\varepsilon_t}(s)$ is well defined for $0 < \beta \leq \min\{1, \frac{\mu_2}{1+\mu_1}\}$. Thus, we obtain that the random variable ε_t is a mixture of two random variables which have geometric distribution $\text{Geom}\big(\frac{\beta}{1+\beta}\big)$ with probability $\frac{\beta\mu_1}{\mu_2-\beta}$ and geometric distribution $\text{Geom}\big(\frac{\mu_2}{1+\mu_2}\big)$ with probability $1-\frac{\beta\mu_1}{\mu_2-\beta}$.

The third case, $t \geqslant \tau_1 + 2$, can be considered in similar manner which implies that

$$\Phi_{\varepsilon_t}(s) = \frac{\beta\mu_2}{\mu_2-\beta}\cdot\frac{1}{1+\beta-\beta s} + \left(1-\frac{\beta\mu_2}{\mu_2-\beta}\right)\cdot\frac{1}{1+\mu_2-\mu_2 s},$$

where $0 < \beta \leq \frac{\mu_2}{1+\mu_2}$. Having in mind the above interval for parameter β, it follows that $\beta \in \Big(0, \min\{\frac{\mu_2}{1+\mu_2}, \frac{\mu_2}{1+\mu_1}\}\Big]$, which competes the proof of theorem. □

The introduced mixture integer-valued autoregressive model with a structural break is obviously a first-order Markov process, so the transition probabilities can be derived by considering the conditional probabilities $\pi(x_t|x_{t-1}) \equiv P(X_t = x_t|X_{t-1} = x_{t-1})$, where x_t and x_{t-1} are non-negative integers. These conditional probabilities are given by the following theorem and they will be used later for the derivation of the conditional log-likelihood function and for the conditional maximum likelihood estimation.

Theorem 2. *Let $\{X_t\}$, $t \in \mathcal{Z}$, be the model given by Definition 1 which satisfies the assumptions of Theorem 1. Let us define the functions $m(t) = \min(x_t, x_{t-1})$, $b(y,i,\theta) = \binom{y}{i}\theta^i(1-\theta)^{y-i}$, $g(y,\theta) = \frac{\theta^y}{(1+\theta)^{y+1}}$ and $h(y,i,\theta) = \binom{y+i-1}{i}\frac{\theta^i}{(1+\theta)^{y+i}}$. Let I_A be the indicator function of the event A. Then*

$$\pi(x_t|x_{t-1}) =$$
$$= \begin{cases} \sum_{i=0}^{m(t)} b(x_{t-1},i,\alpha)\left[\alpha I_{\{x_t=i\}} + (1-\alpha)g(x_t-i,\mu_1)\right], & t\leqslant\tau \\ \sum_{i=0}^{x_t} h(x_{t-1},i,\beta)\left[\frac{\beta\mu_1}{\mu_2-\beta}g(x_t-i,\beta) + \left(1-\frac{\beta\mu_1}{\mu_2-\beta}\right)g(x_t-i,\mu_2)\right], & t=\tau+1 \\ \sum_{i=0}^{x_t} h(x_{t-1},i,\beta)\left[\frac{\beta\mu_2}{\mu_2-\beta}g(x_t-i,\beta) + \left(1-\frac{\beta\mu_2}{\mu_2-\beta}\right)g(x_t-i,\mu_2)\right], & t\geqslant\tau+2. \end{cases}$$

Proof. We will prove theorem only for the case $t \leqslant \tau$. All other cases can be proved similarly. Let us first suppose that $x_{t-1} > 0$. Since $t \leqslant \tau$, we have that

$$\pi(x_t|x_{t-1}) = P(\alpha \circ X_{t-1} + \varepsilon_t = x_t|X_{t-1} = x_{t-1}).$$

Now, the random variable $\alpha \circ X_{t-1}$ for given $X_{t-1} = x_{t-1}$ has the binomial distribution with parameters x_{t-1} and α. Let us denote this random variable as $Bin(x_{t-1}, \alpha)$. Then

$$\pi(x_t|x_{t-1}) = P(Bin(x_{t-1}, \alpha) + \varepsilon_t = x_t) = \sum_{i=0}^{m(t)} b(x_{t-1}, i, \alpha) P(\varepsilon_t = x_t - i).$$

According to distribution (2.1), we have that

$$P(\varepsilon_t = x_t - i) = \alpha I_{\{x_t = i\}} + (1-\alpha) g(x_t - i, \mu_1).$$

Replacing this in the above equation we obtain the expression for the conditional probability when x_{t-1} is positive integer. When $x_{t-1} = 0$ we have that

$$\pi(x_t|0) = P(\varepsilon_t = x_t) = \alpha I_{\{x_t = 0\}} + (1-\alpha) g(x_t, \mu_1).$$

□

As a next property we consider the covariance and the correlation structure of the introduced model. These properties will be used later for the estimation of the unknown parameters. This structure is given by the following theorem.

Theorem 3. *Let $\{X_t\}$, $t \in \mathcal{Z}$, be the model given by Definition 1 which satisfies the assumptions of Theorem 1. Then*

(*a*) *the covariance function of the random variables X_t and X_{t+k}, $k \geqslant 0$, is positive and it is given as*

$$\gamma_t(k) \equiv Cov(X_t, X_{t+k}) = \begin{cases} \alpha^k \mu_1(1+\mu_1), & \text{if } t+k \leqslant \tau, \\ \alpha^{\tau-t}\beta^{t+k-\tau}\mu_1(1+\mu_1), & \text{if } t \leqslant \tau < t+k, \\ \beta^k \mu_2(1+\mu_2), & \text{if } \tau < t, \end{cases}$$

(*b*) *the correlation function of the random variables X_t and X_{t+k}, $k \geq 0$, is positive, always less than 1 and given by*

$$\rho_t(k) \equiv Corr(X_t, X_{t+k}) = \begin{cases} \alpha^k, & \text{if } t+k \leqslant \tau, \\ \alpha^{\tau-t}\beta^{t+k-\tau}\sqrt{\frac{\mu_1(1+\mu_1)}{\mu_2(1+\mu_2)}}, & \text{if } t \leqslant \tau < t+k, \\ \beta^k, & \text{if } \tau < t. \end{cases}$$

Proof. (a) Let us first consider the covariance function between the random variables X_t and X_{t+k}. If $t+k \leq \tau$, then using the independency between X_t and the counting series incorporated in $\alpha \circ X_{t+k-1}$, and the independency between the random variables X_t and ε_{t+k}, we have that

$$\begin{aligned}\gamma_t(k) &= Cov(X_t, \alpha \circ X_{t+k-1} + \varepsilon_{t+k}) \\ &= \alpha Cov(X_t, X_{t+k-1}) \\ &= \alpha^k Var(X_t).\end{aligned}$$

Since $t \leqslant \tau$, then $Var(X_t) = \mu_1(1+\mu_1)$ and $\gamma_t(k) = \alpha^k \cdot \mu_1(1+\mu_1)$. Obviously, the covariance function is positive in this case.

If $t \leqslant \tau < t+k$, then using the independency of the random variables considered in the first case, we have that

$$\begin{aligned}\gamma_t(k) &= Cov(X_t, \beta * X_{t+k-1} + \varepsilon_{t+k}) \\ &= \beta Cov(X_t, X_{t+k-1}) \\ &= \beta^{t+k-\tau} Cov(X_t, X_\tau).\end{aligned}$$

Since $Cov(X_t, X_\tau) = Cov(X_t, X_{t+(\tau-t)})$ and $\tau - t \geqslant 0$, we can apply the result of the first case which implies that $\gamma_t(k) = \beta^{t+k-\tau}\alpha^{\tau-t}Var(X_t)$. Finally, since $t \leqslant \tau$, we have that $\gamma_t(k) = \beta^{t+k-\tau}\alpha^{\tau-t}\mu_1(1+\mu_1)$. Obviously, the covariance function is positive in this case. The third case can be considered in similar way which implies the proof of the first part of theorem.

(b) The correlation function of the random variables X_t and X_{t+k} can be represented via the corresponding covariance function as

$$\rho_t(k) = \frac{\gamma_t(k)}{\sqrt{Var(X_t) \cdot Var(X_{t+k})}}.$$

If $t + k \leq \tau$, then we have that the random variables X_t and X_{t+k} have the same distributions which together with $\gamma_t(k) = \alpha^k Var(X_t)$ imply that $\rho_t(k) = \alpha^k$.

If $t \leqslant \tau < t+k$, then we have that the random variables X_t and X_{t+k} have geometric distributions with means μ_1 and μ_2, respectively. Then, using this and the result of the first part of theorem, we have that

$$\rho_t(k) = \frac{\alpha^{\tau-t}\beta^{t+k-\tau}\mu_1(1+\mu_1)}{\sqrt{\mu_1(1+\mu_1)\cdot\mu_2(1+\mu_2)}} = \alpha^{\tau-t}\beta^{t+k-\tau}\sqrt{\frac{\mu_1(1+\mu_1)}{\mu_2(1+\mu_2)}}.$$

The last case can be considered in similar way which implies the proof of the second part of theorem related to expression of the correlation function.

Obviously, the correlation function is positive in all cases. Let us now prove that the correlation function $\rho_t(k)$ is always less than 1. Since the parameters α and β belong to $(0,1)$, this conclusion is obviously in cases: $t + k \leq \tau$ and $t > \tau$. Let us consider the case $t \leqslant \tau < t+k$. Since $\beta \in \Big(0, \min\{\frac{\mu_2}{1+\mu_2}, \frac{\mu_2}{1+\mu_1}\}\Big)$, we have that $\beta < \mu_2/(1+\mu_1) < (1+\mu_2)/\mu_1$. From this we obtain that $\beta < \sqrt{\frac{\mu_2(1+\mu_2)}{\mu_1(1+\mu_1)}}$, which implies that $\rho_t(k)$ is less than 1. $\square$

Remark 2.1. From the results of the previous theorem we can conclude that the correlation function of the random variables X_t and X_{t+k}, $k \geq 0$ can be written in the form $\rho_t(k) = ac^k$, where a is positive and $c \in \{\beta, \alpha\}$, which implies that $Corr(X_t, X_{t+k})$ converges to 0 when $k \longrightarrow \infty$.

In the following theorem we present some conditional properties of the introduced model which can be derived without using the complicated conditional probabilities.

Theorem 4. *Let $\{X_t\}$, $t \in \mathcal{Z}$, be the model given by Definition 1 which satisfies the assumptions of Theorem 1. Then:*

(a) The conditional expectation of the random variable X_{t+k}, $k \geqslant 0$, for given X_t is a linear function of X_t given by

$$E(X_{t+k}|X_t) = \begin{cases} \alpha^k(X_t - \mu_1) + \mu_1, & \text{if } t+k \leqslant \tau, \\ \alpha^{\tau-t}\beta^{t+k-\tau}(X_t - \mu_1) + \mu_2, & \text{if } t \leqslant \tau < t+k, \\ \beta^k(X_t - \mu_2) + \mu_2, & \text{if } \tau < t. \end{cases}$$

(b) The conditional variance of the random variable X_{t+k} for given X_t is of the form

$$V(X_{t+k}|X_t) = a_{t,k}X_t + b_{t,k}, \tag{2.4}$$

where $a_{t,k}$ and $b_{t,k}$ are given respectively as

$$a_{t,k} = \begin{cases} \alpha^k(1-\alpha^k), & t+k \leqslant \tau, \\ \frac{\alpha^{\tau-2t}\beta^{t+k-2\tau}}{1-\beta}(\alpha^t\beta^\tau + \alpha^t\beta^{\tau+1} - \alpha^\tau\beta^{t+k} - 2\alpha^t\beta^{t+k+1} & \\ +\alpha^\tau\beta^{t+k+1}), & t \leqslant \tau < t+k, \\ \dfrac{\beta^k(1+\beta)(1-\beta^k)}{1-\beta}, & \tau < t, \end{cases}$$

and

$$b_{t,k} = \begin{cases} \mu_1(1-\alpha^k)(1+\mu_1+\mu_1\alpha^k), & t+k \leqslant \tau, \\ \dfrac{\alpha^{-2t}\beta^{-2\tau}}{1-\beta}[\alpha^{2t}(1-\beta)\beta^{2\tau}\mu_2(1+\mu_2) - \alpha^{2\tau}(1-\beta)\beta^{2t+2k}\mu_1^2 & \\ +\alpha^{t+\tau}\beta^{t+k}(2\beta^{t+k+1} - \beta^\tau - \beta^{\tau+1})], & t \leqslant \tau < t+k, \\ \dfrac{(1-\beta^k)\mu_2[1-\beta-2\beta^{k+1}+(1-\beta)(1+\beta^k)\mu_2]}{1-\beta}, & \tau < t. \end{cases}$$

Proof. (a) Let us first consider the conditional expectation of the random variable X_{t+k}, $k \geqslant 0$, for given X_t. Let $t + k \leq \tau$. Then using the property of the binomial thinning operator $E(\alpha\circ X_{t+k-1}|X_{t+k-1})=\alpha X_{t+k-1}$ and the first-order Markovian property, we have that $E(X_{t+k}|X_t)=\alpha E(X_{t+k-1}|X_t)+E(\varepsilon_{t+k})$. Applying the last equation $k-1$ more times, we have that $E(X_{t+k}|X_t)=\alpha^kX_t+\sum_{j=0}^{k-1}\alpha^jE(\varepsilon_{t+k-j})$. Since $t+k\leqslant\tau$, the random variables ε_{t+k}, ε_{t+k-1}, $\dots$, ε_{t+1}, are identically distributed random variables. Also, from (2.1) we have that $E(\varepsilon_{t+k-j})=(1-\alpha)\mu_1$ for $j\in\{0,1,\dots,k-1\}$. Thus,

$$E(X_{t+k}|X_t)=\alpha^kX_t+\mu_1(1-\alpha^k). \tag{2.5}$$

Let $t\leqslant\tau<t+k$. First, using the property of the negative binomial thinning operator $E(\beta{*}X_{t+k-1}|X_{t+k-1})=\beta X_{t+k-1}$ and applying $t+k-\tau$ times the obtained equation, we have that

$$\begin{aligned} E(X_{t+k}|X_t) &= \beta E(X_{t+k-1}|X_t)+E(\varepsilon_{t+k}) \\ &= \beta^{t+k-\tau}E(X_\tau|X_t)+\sum_{j=0}^{t+k-\tau-2}\beta^jE(\varepsilon_{t+k-j})+\beta^{t+k-\tau-1}E(\varepsilon_{\tau+1}). \end{aligned}$$

The random variables ε_{t+k}, ε_{t+k-1}, $\dots$, $\varepsilon_{\tau+2}$, have the distribution (2.3) which implies that the expectation is $E(\varepsilon_{t+k-j})=\mu_2(1-\beta)$ for $j\in\{0,1,\dots,t+k-\tau-2\}$. According to (2.2), we have that $E(\varepsilon_{\tau+1})=\mu_2-\beta\mu_1$. Replacing these results in expression of $E(X_{t+k}|X_t)$, we have that

$$E(X_{t+k}|X_t)=\beta^{t+k-\tau}E(X_\tau|X_t)-\beta^{t+k-\tau}\mu_1+\mu_2.$$

Since $E(X_\tau|X_t)=E(X_{t+(\tau-t)}|X_t)$ and $\tau-t\geqslant 0$, using the result (2.5) for $k=\tau-t$, we obtain that

$$E(X_{t+k}|X_t)=\beta^{t+k-\tau}\alpha^{\tau-t}(X_t-\mu_1)+\mu_2.$$

The third case can be considered in similar way which implies the proof of the first part of theorem.

(b) The easiest way to calculate the conditional variance $Var(X_{t+k}|X_t)$ is to use the conditional probability generating function $g_{t,k}(s)=E(s^{X_{t+k}}|X_t)$ and the property

$$Var(X_{t+k}|X_t)=g''_{t,k}(1)+g'_{t,k}(1)-[g'_{t,k}(1)]^2. \tag{2.6}$$

We have three cases: $t+k\leq\tau$, $\tau<t$ and $t\leq\tau<t+k$. We will consider each case separately.

At first we suppose that $t+k\leq\tau$. Conditional probability generating function of the random variable X_{t+k}, $k\geqslant 0$, for given X_t is

$$\begin{aligned} g_{t,k}(s) &= E(s^{\alpha\circ X_{t+k-1}+\varepsilon_{t+k}}\mid X_t) \\ &= E[(1-\alpha+\alpha s)^{X_{t+k-1}}|X_t]E(s^{\varepsilon_{t+k}}) \\ &= g_{t,k-1}(1-\alpha+\alpha s)\varphi(s), \end{aligned} \tag{2.7}$$

where $\varphi(s) = \dfrac{1+\alpha\mu_1-\alpha\mu_1 s}{1+\mu_1-\mu_1 s}$. In fact, we have a recurrent formula $g_{t,k}(s)$ with an initial condition $g_{t,0}(s) = s^{X_t}$. Thus, we obtain that

$$g_{t,k}(s) = \frac{1+\mu_1\alpha^k-\mu_1\alpha^k s}{1+\mu_1-\mu_1 s}(1-\alpha+\alpha s)^{X_t}. \tag{2.8}$$

According to (2.6), we have

$$Var(X_{t+k}|X_t) = \alpha^k(1-\alpha^k)X_t + \mu_1(1-\alpha^k)(1+\mu_1+\mu_1\alpha^k). \tag{2.9}$$

Let us consider now the third case $\tau \le t$. The second case will be considered at the end of the proof. Conditional probability generating function of the random variable X_{t+k}, $k \geqslant 0$, for given X_t is given by

$$\begin{aligned} g_{t,k}(s) &= E(s^{\beta * X_{t+k-1}+\varepsilon_{t+k}} \mid X_t) \\ &= E\Big[\Big(\frac{1}{1+\beta-\beta s}\Big)^{X_{t+k-1}} \mid X_t\Big]E(s^{\varepsilon_{t+k}}) \\ &= g_{t,k-1}\Big(\frac{1}{1+\beta-\beta s}\Big)\psi(s), \end{aligned} \tag{2.10}$$

where $\psi(s) = \frac{1+\beta(1+\mu_2)-\beta(1+\mu_2)s}{(1+\beta-\beta s)(1+\mu_2-\mu_2 s)}$. The function $g_{t,k}(s)$ can be written in the following form

$$g_{t,k}(s) = (K_k(s))^{X_t}B_k(s), \tag{2.11}$$

where

$$K_k(s) = \frac{1-\beta^k-\beta(1-\beta^{k-1})s}{1-\beta^{k+1}-\beta(1-\beta^k)s}$$

and

$$B_k(s) = \frac{1-\beta^{k+1}+(1-\beta)\beta^k\mu_2-\beta[1-\beta^k+(1-\beta)\beta^{k-1}\mu_2]s}{[1-\beta^{k+1}-\beta(1-\beta^k)s][1+\mu_2-\mu_2 s]}.$$

Then, by applying (2.6), we have that

$$\begin{aligned} Var(X_{t+k} \mid X_t) &= \frac{\beta^k(1+\beta)(1-\beta^k)}{1-\beta}X_t \\ &+ \frac{(1-\beta^k)\mu_2[1-\beta-2\beta^{k+1}+(1-\beta)(1+\beta^k)\mu_2]}{1-\beta}. \end{aligned}$$

At the end, we suppose that $t \le \tau < t+k$. Since $t+k = \tau+1+(t+k-\tau-1)$, we have the following

$$
\begin{aligned}
&E(s^{X_{t+k}} \mid X_t) = \\
&= E[(K_{t+k-\tau-1}(s))^{X_{\tau+1}} \mid X_t] B_{t+k-\tau-1}(s) \\
&= E[(K_{t+k-\tau-1}(s))^{\beta * X_\tau + \varepsilon_{\tau+1}} \mid X_t] B_{t+k-\tau-1}(s) \\
&= E\Big[\Big(\frac{1}{1+\beta-\beta K_{t+k-\tau-1}(s)}\Big)^{X_\tau} \mid X_t\Big] \\
&\times \frac{1+\beta(1+\mu_1)-\beta(1+\mu_1)K_{t+k-\tau-1}(s)}{(1+\beta-\beta K_{t+k-\tau-1}(s))(1+\mu_2-\mu_2 K_{t+k-\tau-1}(s))} B_{t+k-\tau-1}(s).
\end{aligned}
$$

Let $L(s) = (1+\beta-\beta K_{t+k-\tau-1}(s))^{-1}$. Then, the next applies

$$
\begin{aligned}
&E(s^{X_{t+k}} \mid X_t) = \\
&= E((L(s))^{X_\tau} \mid X_t) \frac{1+\beta(1+\mu_1)-\beta(1+\mu_1)K_{t+k-\tau-1}(s)}{(1+\beta-\beta K_{t+k-\tau-1}(s))(1+\mu_2-\mu_2 K_{t+k-\tau-1}(s))} B_{t+k-\tau-1}(s) \\
&= (1-\alpha^{\tau-t}+\alpha^{\tau-t}L(s))^{X_t} \frac{1+\mu_1\alpha^{\tau-t}-\mu_1\alpha^{\tau-t}L(s)}{1+\mu_1-\mu_1 L(s)} \\
&\times \frac{1+\beta(1+\mu_1)-\beta(1+\mu_1)K_{t+k-\tau-1}(s)}{(1+\beta-\beta K_{t+k-\tau-1}(s))(1+\mu_2-\mu_2 K_{t+k-\tau-1}(s))} B_{t+k-\tau-1}(s).
\end{aligned}
$$

Finally, from (2.6), we have that

$$
\begin{aligned}
&Var(X_{t+k} \mid X_t) = \\
&= \frac{\alpha^{\tau-2t}\beta^{t+k-2\tau}}{1-\beta}(\alpha^t\beta^\tau + \alpha^t\beta^{\tau+1} - \alpha^\tau\beta^{t+k} - 2\alpha^t\beta^{t+k+1} + \alpha^\tau\beta^{t+k+1})X_t \\
&+ \frac{\alpha^{-2t}\beta^{-2\tau}}{1-\beta}[\alpha^{2t}(1-\beta)\beta^{2\tau}\mu_2(1+\mu_2) - \alpha^{2\tau}(1-\beta)\beta^{2t+2k}\mu_1^2 \\
&+ \alpha^{t+\tau}\beta^{t+k}(2\beta^{t+k+1} - \beta^\tau - \beta^{\tau+1})]
\end{aligned}
$$

and the theorem is proved by this. □

3. Estimation of the unknown parameters

In this section we consider estimation of the unknown parameters of our model. We suppose that we have a realization $(X_1, X_2, \dots, X_N)$ of size N of the mixture integer-valued autoregressive model with a structural break given by (1.1). We want to estimate the position of the structural break τ and to estimate the parameters α, β, μ_1 and μ_2. We consider two estimation methods: the conditional maximum likelihood method and the conditional least squares method. In the following two subsections we analyze both methods.

3.1. Conditional maximum likelihood estimation

The maximum likelihood method has been widely used for the estimation of the structural breaks and change points depending on the considered problem. The maximum likelihood method derives the estimators of the structural breaks or change points by maximizing the log likelihood function with respect to the structural breaks or given change points. For example, Hinkley and Hinkley [7] considered the maximum likelihood method for the estimation of one change point for the sequence of independent random variables with Bernoulli distributions, Horváth [8] used the maximum likelihood method to test changes in the parameters of independent normal distributed random variables, Avery and Anderson [3] used it for the estimation of two change points in DNA sequences etc.

Since our observations are dependent we consider the conditional maximum likelihood method which estimates the unknown structural break τ and the parameters α, β, μ_1 and μ_2 by maximizing the conditional log-likelihood function $L = \log P(X_i = x_i, 2 \leqslant i \leqslant N | X_1 = x_1)$. Using the fact that our model is a first-order Markov process and the results of Theorem 2, we obtain that the conditional log-likelihood function is given by

$$\begin{aligned} &L(\tau,\alpha,\beta,\mu_1,\mu_2) = \\ &= \sum_{t=2}^{\tau} \log \left\{ \sum_{i=0}^{m(t)} b(x_{t-1},i,\alpha) \left[\alpha I_{\{x_t=i\}} + (1-\alpha) g(x_t - i,\mu_1)\right] \right\} \\ &+ \log \left\{ \sum_{i=0}^{x_{\tau+1}} h(x_\tau,i,\beta) \left[\frac{\beta\mu_1}{\mu_2-\beta} g(x_{\tau+1}-i,\beta) + \left(1 - \frac{\beta\mu_1}{\mu_2-\beta}\right) g(x_{\tau+1}-i,\mu_2) \right] \right\} \\ &+ \sum_{t=\tau+2}^{N} \log \left\{ \sum_{i=0}^{x_t} h(x_{t-1},i,\beta) \left[\frac{\beta\mu_2}{\mu_2-\beta} g(x_t-i,\beta) + \left(1-\frac{\beta\mu_2}{\mu_2-\beta}\right) g(x_t-i,\mu_2) \right] \right\}, \end{aligned}$$

where $m(t) = \min(x_t, x_{t-1})$.

We can consider two approaches based on the conditional maximum likelihood estimation method. Which approach will be used depends on the sample size. Thus, if the sample size is not too large, we can estimate the unknown parameters as follows. For each fixed and known τ we estimate the parameters α, β, μ_1 and μ_2 by maximizing the conditional log-likelihood function L. Let us denote these estimates by $\widehat{\alpha}_\tau$, $\widehat{\beta}_\tau$, $\widehat{\mu}_{1,\tau}$ and $\widehat{\mu}_{2,\tau}$. Then, the estimate of the structural break $\widehat{\tau}$ is obtained as the value τ which maximizes the function $L(\tau, \widehat{\alpha}_\tau, \widehat{\beta}_\tau, \widehat{\mu}_{1,\tau}, \widehat{\mu}_{2,\tau})$. This approach works very well and gives good estimates for any sample size, but can be slow when the sample size is greater than 2000. In this case, we can estimate all five parameters by maximizing the conditional log-likelihood function L with respect to all these parameters. The simulations show that the first approach gives better estimates, especially it estimates better the structural break.

3.2. Conditional least squares estimation

According to the results presented in the first part of Theorem 4, the conditional least squares estimators are obtained by minimizing the function Q given as

$$\begin{aligned} Q &= \sum_{t=1}^{\tau-1}[X_{t+1} - \alpha X_t - (1-\alpha)\mu_1]^2 + (X_{\tau+1} - \beta X_\tau - \mu_2 + \beta\mu_1)^2 \\ &\quad + \sum_{t=\tau+1}^{N-1}[X_{t+1} - \beta X_t - (1-\beta)\mu_2]^2 \end{aligned}$$

with respect to the unknown parameters τ, α, β, μ_1 and μ_2. If the parameter τ is known, then the estimators of the parameters α, β, μ_1 and μ_2 can be obtained as the solutions of the system of equations

$$\begin{aligned} \frac{\partial Q}{\partial \alpha} &= \sum_{t=1}^{\tau-1}[X_{t+1} - \alpha X_t - (1-\alpha)\mu_1](-X_t + \mu_1) = 0, \\ \frac{\partial Q}{\partial \beta} &= (X_{\tau+1} - \beta X_\tau - \mu_2 + \beta\mu_1)(-X_\tau + \mu_1) + \\ &\quad + \sum_{t=\tau+1}^{N-1}[X_{t+1} - \beta X_t - (1-\beta)\mu_2](-X_t + \mu_2) = 0, \\ \frac{\partial Q}{\partial \mu_1} &= -(1-\alpha)\sum_{t=1}^{\tau-1}[X_{t+1} - \alpha X_t - (1-\alpha)\mu_1] + \beta(X_{\tau+1} - \beta X_\tau - \mu_2 + \beta\mu_1) = 0, \\ \frac{\partial Q}{\partial \mu_2} &= (X_{\tau+1} - \beta X_\tau - \mu_2 + \beta\mu_1) + (1-\beta)\sum_{t=\tau+1}^{N-1}[X_{t+1} - \beta X_t - (1-\beta)\mu_2] = 0, \end{aligned}$$

with respect to the parameters α, β, μ_1 and μ_2. Similarly as in the first approach discussed in the conditional maximum likelihood estimation, we can derive the conditional least squares estimates as follows. For each fixed and known true value of the parameter τ we estimate the remaining parameters α, β, μ_1 and μ_2 by minimizing the function Q. Let us denote again these estimates by $\widehat{\alpha}_\tau$, $\widehat{\beta}_\tau$, $\widehat{\mu}_{1,\tau}$ and $\widehat{\mu}_{2,\tau}$. Then, the estimate of the structural break $\widehat{\tau}$ is obtained as the value τ which minimizes the function $Q(\tau, \widehat{\alpha}_\tau, \widehat{\beta}_\tau, \widehat{\mu}_{1,\tau}, \widehat{\mu}_{2,\tau})$, i.e.

$$\hat{\tau} = \arg\min_{\tau} Q(\tau, \widehat{\alpha}_\tau, \widehat{\beta}_\tau, \widehat{\mu}_{1,\tau}, \widehat{\mu}_{2,\tau}),$$

where we consider the realized values of the function Q for all $\tau \in [1, N-1]$.

3.3. Simulations

To check the performances of the estimators obtained by the conditional maximum likelihood method and the conditional least squares method we simulated 1000 samples for the following cases of true values of the parameters. We considered two cases: (a) the true values are $\alpha = 0.2$, $\beta = 0.3$, $\mu_1 = 1$ and $\mu_2 = 2$; and (b) the true values are $\alpha = 0.4$, $\beta = 0.8$, $\mu_1 = 4$ and $\mu_2 = 10$. Thus, in the first case we supposed that the correlations between the observations are small and the mean parameters are moderate and similar. In the second case we supposed that the correlation and the mean in the case of the negative binomial thinning operator are significantly larger.

For both cases we consider samples of sizes $n = 100$, $n = 200$ and $n = 500$. For each different sample size we consider three different cases: $\tau = n/4$, $\tau = n/2$ and $\tau = 3n/4$.

The maximization and the minimization have been performed by using the `optim` function from R statistical software and using the method Nelder-Mead. Some parts of the code have been written in `Rcpp` to improve the speed of the estimation. For each case we considered the mean, median, lower (Q_1) and upper (Q_3) quartiles, and standard error (SE) of the obtained estimates. The flow of the estimation procedure can be divided into two steps. In the first step, for all $\tau \in [1, T-1]$ we estimate parameters α, β, μ_1 and μ_2. Consequently, we obtained values of functions L and Q discussed in the previous subsections. In the second step, from these calculated values, we pick the value of τ which maximise the value of L for CML and minimize the value of Q for the CLS method. All the results are presented in Tables 3.1–3.3.

For big enough and small enough sample size, we conclude that there is regularity. As the value of the structural parameter τ rises, as the value SE of $\widehat{\alpha}$ and $\widehat{\mu}_1$ decreases. For the value SE of $\widehat{\beta}$ and $\widehat{\mu}_2$, we can conclude the opposite. While the value of the structural parameter τ rises, the value SE of $\widehat{\beta}$ and $\widehat{\mu}_2$ also rises. This correctness is justified by the number of information we receive. When the value of the structural parameter τ increase, the number of information for $\widehat{\alpha}$ and $\widehat{\mu}_1$ grows and the number of information for $\widehat{\beta}$ and $\widehat{\mu}_2$ decreases. When we have more information the error is smaller and conversely.

We can notice that the estimates are quite close to the true values when the sample size is 500, while there are some deviations for samples of size 100. Actually, the CML method shows remarkable results even for samples of sizes 100 and 200, while the CLS method are not quite accurate in these cases, especially for parameters μ_1 and μ_2. The CML method provides consistent estimates for both set of parameters, while CLS gives slightly better results when the value of parameters are smaller, i.e. for the parameter set a). Especially interesting part is the estimation of the structural break (parameter τ). While the CML method is quite accurate for all tested cases, the CLS method shows some deviations from true value even for the samples of size 500.

After all we can conclude that the estimates converges to their true values with the increase of the sample size, while the convergence is faster with CML than with CLS method.

$n = 100,\ \tau = 25,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	40.3010	31	22	58	0.8389	$\widehat{\tau}$	56.178	53	32	83	0.8509
$\widehat{\tau}-\tau$	15.301	6	−3	33	0.8389	$\widehat{\tau}-\tau$	31.178	28	7	58	0.8509
$\widehat{\alpha}$	0.2113	0.1795	0.0240	0.3111	0.0065	$\widehat{\alpha}$	0.2539	0.2099	0.0663	0.3602	0.0074
$\widehat{\beta}$	0.2838	0.2836	0.1722	0.3761	0.0055	$\widehat{\beta}$	0.2392	0.2025	0.0300	0.3434	0.0073
$\widehat{\mu}_1$	0.9436	0.8422	0.5160	1.3039	0.0205	$\widehat{\mu}_1$	1.4065	1.2500	0.8447	1.6695	0.0335
$\widehat{\mu}_2$	2.1913	2.0686	1.7100	2.5004	0.0379	$\widehat{\mu}_2$	2.9528	2.3678	1.8130	3.1826	0.0679
$n = 100,\ \tau = 50,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	52.697	53	40.75	69.25	0.7504	$\widehat{\tau}$	64.957	65	51	82	0.6916
$\widehat{\tau}-\tau$	2.697	3	-9.25	19.25	0.7504	$\widehat{\tau}-\tau$	14.957	15	1	32	0.6916
$\widehat{\alpha}$	0.2059	0.1853	0.0728	0.2918	0.0058	$\widehat{\alpha}$	0.2113	0.1791	0.0674	0.3014	0.0060
$\widehat{\beta}$	0.2823	0.2667	0.1599	0.3919	0.0057	$\widehat{\beta}$	0.2386	0.1999	0.0305	0.3582	0.0073
$\widehat{\mu}_1$	0.9375	0.8892	0.6514	1.1087	0.0187	$\widehat{\mu}_1$	1.1521	1.0440	0.8236	1.2945	0.0225
$\widehat{\mu}_2$	2.2583	2.1321	1.6926	2.6426	0.0338	$\widehat{\mu}_2$	3.0559	2.4966	1.8853	3.4039	0.0664
$n = 100,\ \tau = 75,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	63.263	75	44	83	0.8781	$\widehat{\tau}$	73.353	80	70	89.25	0.7596
$\widehat{\tau}-\tau$	-11.737	0	-31	8	0.8781	$\widehat{\tau}-\tau$	-1.647	5	-5	14.25	0.7596
$\widehat{\alpha}$	0.2223	0.1937	0.1058	0.2901	0.0061	$\widehat{\alpha}$	0.1994	0.1661	0.0764	0.2721	0.0057
$\widehat{\beta}$	0.2820	0.2449	0.1072	0.4095	0.0071	$\widehat{\beta}$	0.2331	0.1606	0.0001	0.3600	0.0083
$\widehat{\mu}_1$	0.9271	0.8868	0.7090	1.0701	0.0163	$\widehat{\mu}_1$	1.1652	0.9633	0.8051	1.1488	0.0482
$\widehat{\mu}_2$	2.3613	2.1068	1.4056	2.9152	0.0464	$\widehat{\mu}_2$	3.1387	2.5917	1.8062	3.7004	0.0713
$n = 200,\ \tau = 50,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	68.376	54	45	75	1.4224	$\widehat{\tau}$	104.191	86	56	159	1.723
$\widehat{\tau}-\tau$	18.376	4	-5	25	1.4224	$\widehat{\tau}-\tau$	54.191	36	6	109	1.723
$\widehat{\alpha}$	0.1895	0.1748	0.0755	0.2741	0.0050	$\widehat{\alpha}$	0.2222	0.2080	0.1015	0.3206	0.0053
$\widehat{\beta}$	0.2980	0.2990	0.2356	0.3548	0.0039	$\widehat{\beta}$	0.2703	0.2590	0.1577	0.3518	0.0061
$\widehat{\mu}_1$	0.9489	0.8727	0.6749	1.1541	0.0162	$\widehat{\mu}_1$	1.2340	1.1950	0.8670	1.5363	0.0170
$\widehat{\mu}_2$	2.1757	2.0817	1.8718	2.3235	0.0259	$\widehat{\mu}_2$	2.8910	2.2823	1.9894	2.9201	0.0651
$n = 200,\ \tau = 100,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	101.519	102	91	117	1.1599	$\widehat{\tau}$	127.34	117	101.75	160	1.2865
$\widehat{\tau}-\tau$	1.519	2	-9	17	1.1599	$\widehat{\tau}-\tau$	27.34	17	1.75	60	1.2865
$\widehat{\alpha}$	0.1943	0.1854	0.1256	0.2527	0.0038	$\widehat{\alpha}$	0.2035	0.1960	0.1005	0.2842	0.0044
$\widehat{\beta}$	0.2964	0.2927	0.2223	0.3644	0.0040	$\widehat{\beta}$	0.2701	0.2594	0.1431	0.3541	0.0061
$\widehat{\mu}_1$	0.9363	0.9254	0.7842	1.0716	0.0097	$\widehat{\mu}_1$	1.1220	1.0431	0.8798	1.2406	0.0177
$\widehat{\mu}_2$	2.1895	2.1050	1.8272	2.4042	0.0276	$\widehat{\mu}_2$	2.9048	2.3379	1.9872	2.9028	0.0637
$n = 200,\ \tau = 150,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	139.624	151	137.75	162	1.3726	$\widehat{\tau}$	156.556	159	150	179	1.1203
$\widehat{\tau}-\tau$	-10.376	1	-12.25	12	1.3726	$\widehat{\tau}-\tau$	6.556	9	0	29	1.1203
$\widehat{\alpha}$	0.1982	0.1943	0.1395	0.2436	0.0035	$\widehat{\alpha}$	0.1928	0.1893	0.1221	0.2542	0.0034
$\widehat{\beta}$	0.3099	0.2914	0.1975	0.4010	0.0056	$\widehat{\beta}$	0.2770	0.2483	0.0962	0.4009	0.0074
$\widehat{\mu}_1$	0.9559	0.9457	0.8239	1.0627	0.0128	$\widehat{\mu}_1$	1.0598	0.9968	0.8846	1.1105	0.0170
$\widehat{\mu}_2$	2.3912	2.2345	1.7516	2.7357	0.0370	$\widehat{\mu}_2$	3.1465	2.5441	2.0297	3.4113	0.0668

Table 3.1: Conditional maximum likelihood and conditional least squares estimates for different cases of true values of the parameters

$n = 500,\ \tau = 125,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$

CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	135.523	128	119	142	1.7701	$\widehat{\tau}$	191.46	139	127	190.25	3.6558
$\widehat{\tau}-\tau$	10.523	3	-6	17	1.7701	$\widehat{\tau}-\tau$	66.46	14	2	65.25	3.6558
$\widehat{\alpha}$	0.1933	0.1882	0.1343	0.2454	0.0031	$\widehat{\alpha}$	0.1994	0.1896	0.1129	0.2687	0.0039
$\widehat{\beta}$	0.2969	0.2977	0.2655	0.3290	0.0017	$\widehat{\beta}$	0.2859	0.2859	0.2327	0.3328	0.0039
$\widehat{\mu}_1$	0.9482	0.9449	0.8151	1.0640	0.0075	$\widehat{\mu}_1$	1.0919	1.0107	0.8798	1.2274	0.0104
$\widehat{\mu}_2$	2.0448	2.0370	1.9024	2.1460	0.0118	$\widehat{\mu}_2$	2.4803	2.0982	1.9556	2.2773	0.0531

$n = 500,\ \tau = 250,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$

CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	255.23	253	246	265.25	1.4007	$\widehat{\tau}$	289.542	262	251	298.25	2.1692
$\widehat{\tau}-\tau$	5.23	3	-4	15.25	1.4007	$\widehat{\tau}-\tau$	39.542	12	1	48.25	2.1692
$\widehat{\alpha}$	0.1952	0.1935	0.1606	0.2279	0.0019	$\widehat{\alpha}$	0.1991	0.1945	0.1481	0.2446	0.0024
$\widehat{\beta}$	0.2954	0.2923	0.2539	0.3355	0.0020	$\widehat{\beta}$	0.2893	0.2838	0.2297	0.3405	0.0038
$\widehat{\mu}_1$	0.9742	0.9713	0.8955	1.0601	0.0042	$\widehat{\mu}_1$	1.0360	1.0127	0.9278	1.1172	0.0051
$\widehat{\mu}_2$	2.0509	2.0371	1.8962	2.1879	0.0079	$\widehat{\mu}_2$	2.3906	2.1054	1.9496	2.3229	0.0429

$n = 500,\ \tau = 375,\ \alpha = 0.2,\ \beta = 0.3,\ \mu_1 = 1,\ \mu_2 = 2$

CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	372.582	377	369	391	1.8381	$\widehat{\tau}$	394.592	386	375	419	1.8383
$\widehat{\tau}-\tau$	-2.418	2	-6	16	1.8381	$\widehat{\tau}-\tau$	19.592	11	0	44	1.8383
$\widehat{\alpha}$	0.1983	0.1997	0.1720	0.2291	0.0015	$\widehat{\alpha}$	0.2017	0.1963	0.1568	0.2406	0.0025
$\widehat{\beta}$	0.2971	0.2922	0.2341	0.3555	0.0032	$\widehat{\beta}$	0.2803	0.2754	0.1946	0.3551	0.0048
$\widehat{\mu}_1$	0.9715	0.9734	0.9106	1.0404	0.0041	$\widehat{\mu}_1$	1.0258	1.0023	0.9322	1.0779	0.0082
$\widehat{\mu}_2$	2.1009	2.0907	1.8623	2.3062	0.0158	$\widehat{\mu}_2$	2.5460	2.2123	1.9534	2.5673	0.0481

$n = 100,\ \tau = 25,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$

CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	27.315	25	22	31	0.374	$\widehat{\tau}$	52.778	47	25	86	0.9949
$\widehat{\tau}-\tau$	2.315	0	-3	6	0.374	$\widehat{\tau}-\tau$	27.778	22	0	61	0.9949
$\widehat{\alpha}$	0.3823	0.3939	0.3196	0.4604	0.0042	$\widehat{\alpha}$	0.6072	0.7011	0.3497	0.8357	0.0090
$\widehat{\beta}$	0.7953	0.8013	0.7595	0.8356	0.0018	$\widehat{\beta}$	0.5914	0.6536	0.4690	0.8017	0.0077
$\widehat{\mu}_1$	3.8424	3.6386	2.6067	4.9137	0.0555	$\widehat{\mu}_1$	4.8332	2.9441	1.5312	8.1632	0.1367
$\widehat{\mu}_2$	10.1301	9.6837	7.7273	11.9703	0.1113	$\widehat{\mu}_2$	6.5530	5.7130	4.1488	8.2864	0.1325

$n = 100,\ \tau = 50,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$

CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	49.606	50	46	56	0.4526	$\widehat{\tau}$	49.159	47	23	76	0.9257
$\widehat{\tau}-\tau$	-0.394	0	-4	6	0.4526	$\widehat{\tau}-\tau$	-0.841	-3	-27	26	0.9257
$\widehat{\alpha}$	0.3901	0.3899	0.3441	0.4414	0.0028	$\widehat{\alpha}$	0.4820	0.4096	0.2909	0.7275	0.0086
$\widehat{\beta}$	0.7776	0.7962	0.7471	0.8369	0.0036	$\widehat{\beta}$	0.6073	0.6687	0.5105	0.7896	0.0074
$\widehat{\mu}_1$	3.8695	3.8411	3.1233	4.6293	0.0396	$\widehat{\mu}_1$	4.1442	3.0642	1.9004	5.4229	0.1905
$\widehat{\mu}_2$	10.1451	9.3875	7.1875	12.2876	0.1524	$\widehat{\mu}_2$	6.8419	6.0001	4.1815	8.2232	0.1612

$n = 100,\ \tau = 75,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$

CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	69.431	75	67	78	0.5738	$\widehat{\tau}$	49.801	49	24	75	0.9141
$\widehat{\tau}-\tau$	-5.569	0	-8	3	0.5738	$\widehat{\tau}-\tau$	-25.199	-26	-51	0	0.9141
$\widehat{\alpha}$	0.3986	0.4027	0.3614	0.4390	0.0027	$\widehat{\alpha}$	0.3847	0.3467	0.2244	0.5031	0.0076
$\widehat{\beta}$	0.7389	0.7898	0.7034	0.8460	0.0058	$\widehat{\beta}$	0.5889	0.6245	0.4577	0.7546	0.0068
$\widehat{\mu}_1$	3.7731	3.7484	3.2043	4.3642	0.0361	$\widehat{\mu}_1$	3.6433	3.3320	2.2435	4.5701	0.0699
$\widehat{\mu}_2$	10.0257	9.0245	5.4754	13.107	0.1956	$\widehat{\mu}_2$	6.2752	5.7130	4.1102	7.5401	0.1171

Table 3.2: Conditional maximum likelihood and conditional least squares estimates for different cases of true values of the parameters

$n = 200,\ \tau = 50,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	52.079	50	47	56	0.4835	$\widehat{\tau}$	99.802	77.5	39	179	2.1926
$\widehat{\tau} - \tau$	2.079	0	-3	6	0.4835	$\widehat{\tau} - \tau$	49.802	27.5	-11	129	2.1926
$\widehat{\alpha}$	0.3949	0.3962	0.3464	0.4409	0.0027	$\widehat{\alpha}$	0.6169	0.7021	0.3983	0.8215	0.0077
$\widehat{\beta}$	0.7955	0.7977	0.7721	0.8223	0.0012	$\widehat{\beta}$	0.6260	0.7097	0.5035	0.7978	0.0076
$\widehat{\mu}_1$	3.9902	3.8837	3.1168	4.7795	0.0385	$\widehat{\mu}_1$	4.8815	2.7730	1.8929	8.3710	0.1216
$\widehat{\mu}_2$	9.9912	9.7218	8.2559	11.3364	0.0797	$\widehat{\mu}_2$	7.5863	7.0549	4.9478	9.0451	0.1589
$n = 200,\ \tau = 100,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	100.343	100	96	105	0.4853	$\widehat{\tau}$	88.594	72	35	140	1.9663
$\widehat{\tau} - \tau$	0.343	0	-4	5	0.4853	$\widehat{\tau} - \tau$	-11.406	-28	-65	40	1.9663
$\widehat{\alpha}$	0.3977	0.3990	0.3668	0.4298	0.0016	$\widehat{\alpha}$	0.5060	0.4450	0.3367	0.7127	0.0074
$\widehat{\beta}$	0.7944	0.7996	0.7650	0.8298	0.0016	$\widehat{\beta}$	0.6658	0.7230	0.6100	0.7899	0.0062
$\widehat{\mu}_1$	3.9851	3.9772	3.5154	4.4516	0.0235	$\widehat{\mu}_1$	3.6213	2.5677	1.9437	4.6694	0.0807
$\widehat{\mu}_2$	10.0778	9.65	7.9898	11.8108	0.1068	$\widehat{\mu}_2$	6.9571	6.756	4.6707	8.2727	0.1127
$n = 200,\ \tau = 150,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	147.603	150	144.75	155	0.6018	$\widehat{\tau}$	105.155	108.5	51	157	1.8747
$\widehat{\tau} - \tau$	-2.397	0	-5.25	5	0.601	$\widehat{\tau} - \tau$	-44.845	-41.5	-99	7	1.8747
$\widehat{\alpha}$	0.3990	0.3994	0.3739	0.4246	0.0013	$\widehat{\alpha}$	0.4139	0.3896	0.2983	0.4885	0.0060
$\widehat{\beta}$	0.7810	0.7943	0.7449	0.8349	0.0029	$\widehat{\beta}$	0.6495	0.6914	0.5943	0.7713	0.0059
$\widehat{\mu}_1$	3.9524	3.9663	3.5614	4.3297	0.0193	$\widehat{\mu}_1$	3.5984	3.5848	2.3423	4.3866	0.0517
$\widehat{\mu}_2$	10.2504	9.6642	7.1598	12.4546	0.1494	$\widehat{\mu}_2$	6.4835	6.0226	4.4474	7.4225	0.1027
$n = 500,\ \tau = 125,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	126.589	125	123	130	0.4317	$\widehat{\tau}$	214.71	123	74	440.25	5.6351
$\widehat{\tau} - \tau$	1.589	0	-2	5	0.4317	$\widehat{\tau} - \tau$	89.71	-2	-51	315.25	5.6351
$\widehat{\alpha}$	0.3976	0.3995	0.3726	0.4241	0.0013	$\widehat{\alpha}$	0.6082	0.5840	0.444	0.7999	0.0063
$\widehat{\beta}$	0.7979	0.7995	0.7820	0.8150	0.0008	$\widehat{\beta}$	0.6796	0.7602	0.6771	0.8054	0.0067
$\widehat{\mu}_1$	4.0661	3.9800	3.5957	4.5168	0.0223	$\widehat{\mu}_1$	4.4798	2.3832	2.0444	7.9776	0.1049
$\widehat{\mu}_2$	9.9584	9.8093	8.8011	11.0295	0.0499	$\widehat{\mu}_2$	8.1966	8.2114	5.9552	9.8036	0.1216
$n = 500,\ \tau = 250,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	251.822	250	248	256	0.4341	$\widehat{\tau}$	183.334	124.5	67.75	282.25	4.7419
$\widehat{\tau} - \tau$	1.822	0	2	6	0.4341	$\widehat{\tau} - \tau$	-66.666	-125.5	-182.25	32.25	4.7419
$\widehat{\alpha}$	0.4008	0.4018	0.3831	0.4187	0.0008	$\widehat{\alpha}$	0.5115	0.4664	0.3784	0.6269	0.0057
$\widehat{\beta}$	0.7977	0.7993	0.7777	0.8186	0.0009	$\widehat{\beta}$	0.7231	0.7659	0.7187	0.7973	0.0051
$\widehat{\mu}_1$	3.9649	3.9336	3.6907	4.2383	0.0128	$\widehat{\mu}_1$	3.3701	2.2528	1.9854	4.0988	0.0659
$\widehat{\mu}_2$	10.0991	9.9783	8.7784	11.3061	0.0597	$\widehat{\mu}_2$	7.3444	7.2657	5.7130	8.3134	0.1048
$n = 500,\ \tau = 375,\ \alpha = 0.4,\ \beta = 0.8,\ \mu_1 = 4,\ \mu_2 = 10$											
CML	Mean	Median	Q_1	Q_3	SE	CLS	Mean	Median	Q_1	Q_3	SE
$\widehat{\tau}$	374.349	375	370	381	0.5138	$\widehat{\tau}$	268.048	297	137	382	4.3605
$\widehat{\tau} - \tau$	-0.651	0	-5	6	0.5138	$\widehat{\tau} - \tau$	-106.952	-78	-238	7	4.3605
$\widehat{\alpha}$	0.4001	0.4011	0.3851	0.4153	7e-04	$\widehat{\alpha}$	0.4315	0.4174	0.3657	0.4779	0.0038
$\widehat{\beta}$	0.7938	0.7973	0.766	0.8229	0.0014	$\widehat{\beta}$	0.7346	0.7499	0.697	0.7938	0.0036
$\widehat{\mu}_1$	4.0037	4.0093	3.7538	4.2318	0.0111	$\widehat{\mu}_1$	3.4848	3.6406	2.6151	4.2147	0.0365
$\widehat{\mu}_2$	10.1226	9.709	8.1434	11.5507	0.1103	$\widehat{\mu}_2$	6.6011	6.0863	4.5934	7.1362	0.1267

Table 3.3: Conditional maximum likelihood and conditional least squares estimates for different cases of true values of the parameters

4. Illustrative examples

In this section we discuss possible applications of our introduced model. According to the motivations mentioned in the introductory section, we consider two real data sets about some criminal counts observed in Pittsburgh. We consider two real data sets from the forecasting principles site (http://www.forecastingprinciples.com). Each data set contains 144 observations which represents monthly observed corresponding criminal counts in the period between January 1990 and December 2001.

4.1. Computer Aided Dispatch (CAD) calls about drug dealing

The first considered real data set represents Computer Aided Dispatch (CAD) calls about drug dealing registered in Pittsburgh 1011th tract. The sample mean and the sample variance are 3.1944 and 13.3605, respectively, which indicate that we deal with overdispersed data. The sample autocorrelation is 0.6509, so the observations are strongly correlated. The sample path, ACF and PACF plots are presented in Figure 4.1.

From the PACF plot, we can conclude that the first order integer-valued time series model will be adequate for this real data set. Also, from the sample path plot we can conclude that there are low activities about the CAD drug calls before the first 60 months. After that, the CAD drug calls significantly increase. This indicates that there is a change in model after 60 months, so our model with structural break can be considered as competitive model in comparison with GINAR(1) and NGINAR(1) models. For each model we derive the maximum likelihood estimates, values of the Akaike (AIC) and Bayesian (BIC) criterions and the root mean square (RMS). Beside this, we estimate the structural break for our model. All results are presented in Table 4.1.

Model	ML estimates	AIC	BIC	RMS
GINAR(1)	$\widehat{\alpha} = 0.3901$, $\widehat{\mu} = 2.7882$	633.0897	639.0293	2.9385
NGINAR(1)	$\widehat{\beta} = 0.6588$, $\widehat{\mu} = 3.0756$	608.0767	614.0163	2.7664
Mixture	$\widehat{\tau} = 62$, $\widehat{\alpha} = 0.0736$, $\widehat{\beta} = 0.6309$, $\widehat{\mu}_1 = 0.868$, $\widehat{\mu}_2 = 4.2307$	593.0997	604.9790	2.6405

Table 4.1: ML estimates, AIC, BIC, RMS for the CAD drug calls in 1011th tract

From the obtained results, we can conclude that our model provides the smallest values for the AIC and BIC criterions, which leads to conclusion that our model has captured very well the change of models. The estimated structural break is $\widehat{\tau} = 62$ which corresponds to earlier conclusion obtained from the sample path. Also, we can see that the smallest RMS is obtained for our model which means that our model gives best fit among the considered models. Fits for all three models are given in Figure 4.2. The bolded lines represent the fit of each model. From this figure we can conclude that our model very well capture behavior of CAD drug calls.

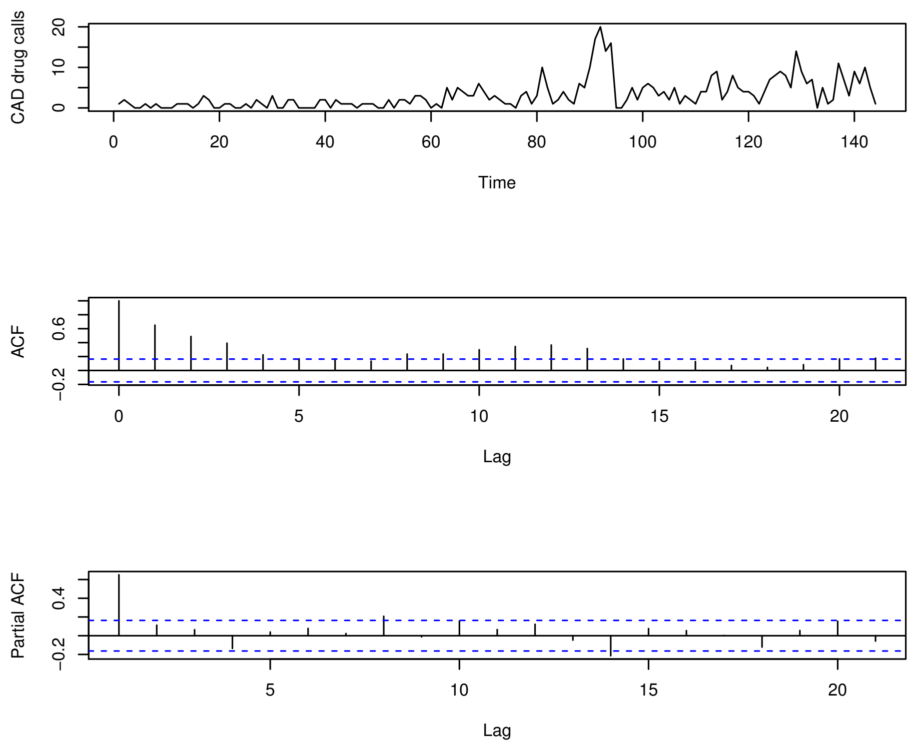

Fig. 4.1: The sample path, ACF and PACF for the CAD calls about drug dealing in 1011th tract

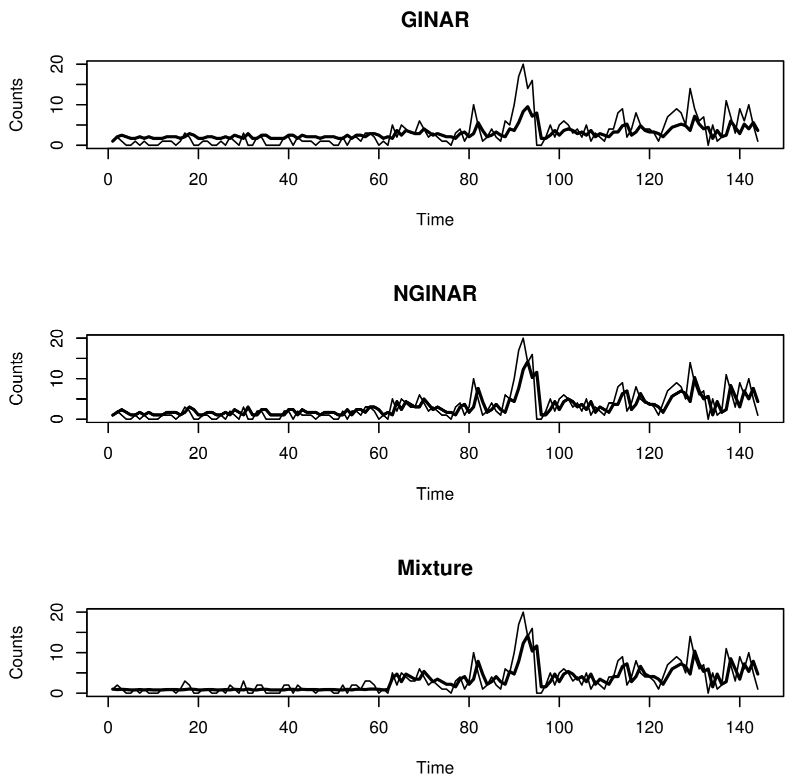

Fig. 4.2: The fits of all three models for the CAD drug calls data

4.2. Phone calls about registered shootings that can be reported also by civilians

The second considered actual dataset consists of phone calls about registered shootings that can be reported also by civilians registered in the 1017th Pittsburgh tract. The sample mean and the sample variance are 5.9722 and 32.3349, respectively. Again, we have overdispersed data. The sample autocorrelation is 0.4563 which indicates significant correlation between the observations. The Figure 4.3 represents the sample path, ACF and PACF plots.

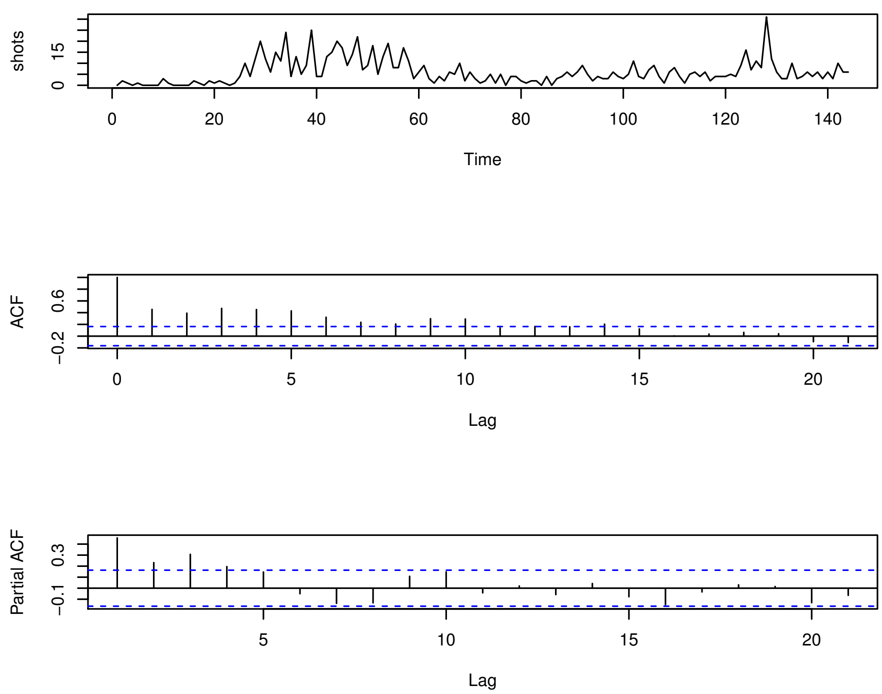

Fig. 4.3: The sample path, ACF and PACF for 1017 data

From the sample path we can observe that there is change in the behavior of the phone calls after around 20 months. This is justified by estimation in which we obtain that estimated structural break is 24 months. From the results presented in Table 4.2, we can conclude that our model very well fit the counts of phone calls.

Model	ML estimates	AIC	BIC	RMS
GINAR(1)	$\widehat{\alpha} = 0.3360$, $\widehat{\mu} = 5.2345$	787.0130	792.9526	5.1089
NGINAR(1)	$\widehat{\beta} = 0.4213$, $\widehat{\mu} = 5.4126$	782.5761	788.5157	5.0521
Mixture	$\widehat{\tau} = 24$, $\widehat{\alpha} = 0.1903$, $\widehat{\beta} = 0.3724$, $\widehat{\mu}_1 = 0.7662$, $\widehat{\mu}_2 = 5.9462$	768.2952	780.1744	4.8912

Table 4.2: ML estimates, AIC, BIC, RMS for the calls in 1017th tract

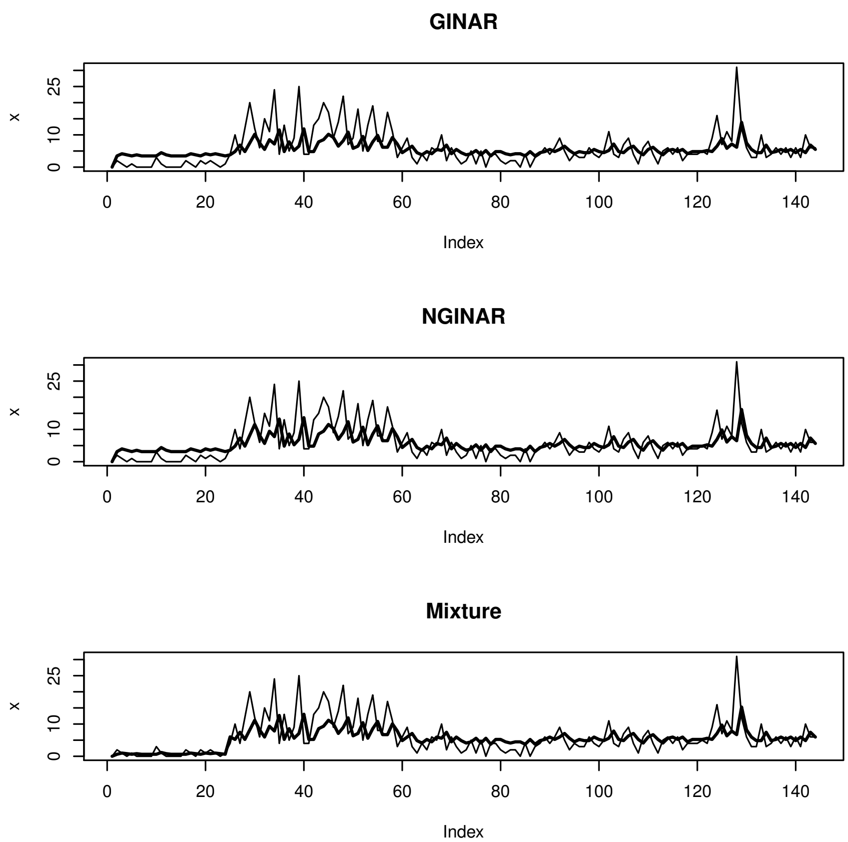

Fig. 4.4: The fits of all three models for the phone calls data

5. Concluding remarks

In this manuscript we have introduced an integer-valued autoregressive model of the first-order with a structural break as a mixture of two integer-valued autoregressive models with binomial and negative binomial thinning operators. The model has been constructed under motivations of the different behaviors of the considered objects before and after a break. Exactly, we have considered objects (virus, criminals etc.) which have low activities before a break leading to counts of small values and after a break have increasing activities leading to counts of large values. Because of that, we have used two different thinning operators, the binomial thinning for low activity and the negative binomial thinning for increasing activity. A model with different geometric marginals has been constructed and many of its properties are considered. Some of them are distribution of the innovations, conditional and unconditional properties, covariance and correlation structures. Two methods of estimations, conditional maximum likelihood and conditional least squares, are considered and the performances of their estimates have been checked by simulations. At the end, applicability of the model has been considered on two real data sets about criminal acts. It would be interesting to introduce some methods which can be used to detect the position of structural break. Standard CUSUM tests cannot be applied for our model because the change in thinning operators is not considered yet. Also, it would be interesting to consider more breaks and generalizes the model introduced in this paper.

REFERENCES

1. M.A. Al-Osh and A.A. Alzaid: *First-order integer-valued autoregressive (INAR(1)) process*, J. Time Ser. Anal. **8(3)** (1987), 261–275.
2. A. Aue and L. Horváth: *Structural breaks in time series*, J. Time Ser. Anal. **34(1)** (2013), 1–16.
3. P.J. Avery and D.A. Henderson: *Detecting a changed segment in DNA sequences*, J. Royal Stat. Soc.: Series C (Appl. Stat.) **48(4)** (1999), 489–503.
4. C.W.S. Chen and S. Lee: *Generalized Poisson autoregressive models for time series of counts*, Comput. Stat. Data Anal. **99** (2016), 51–67.
5. Y. Cui and R. Wu: *Test of parameter changes in a class of observation-driven models for count time series*, Commun. Stat. - Theory Methods (2019), 1–27.
6. J. Franke, C. Kirch and J.T. Kamgaing: *Changepoints in times series of counts*, J. Time Ser. Anal. **33(5)** (2012), 757–770.
7. D.V. Hinkley and E.A. Hinkley: *Inference about the change-point in a sequence of binomial variables*, Biometrika **57(3)** (1970), 477–488.
8. L. Horváth: *The maximum likelihood method for testing changes in the parameters of normal observations*, Ann. Stat. (1993), 671–680.
9. Š. Hudecová: *Structural changes in autoregressive models for binary time series*, J. Stat. Plan. Inference **143(10)** (2013), 1744–1752.
10. Š. Hudecová, M. Hušková and S. Meintanis: *Detection of changes in INAR models*, *Stochastic models, statistics and their applications* (2015), 11–18.

11. A.S. Kashikar, N. Rohan and T.V. Ramanathan: *Integer autoregressive models with structural breaks*, J. Appl. Stat. **40(12)** (2013), 2653–2669.

12. H. Kim and S. Lee: *On residual CUSUM statistic for PINAR(1) model in statistical design and diagnostic of control chart*, Commun. Stat. - Simul. Comput. (2019), 1–25.

13. E. McKenzie: *Some simple models for discrete variate time series*, JAWRA J. Amer. Water Res. Assoc. **21(4)** (1985), 645–650.

14. M.M. Ristić, H.S. Bakouch and A.S. Nastić: *A new geometric first-order integer-valued autoregressive (NGINAR(1)) process*, J. Stat. Plan. Inference **139(7)** (2009), 2218–2226.

15. K. Yu, H. Zou and D. Shi: *Integer-valued moving average models with structural changes*, Math. Probl. Eng. **2014** (2014).