

Maher Kachour*, Hassan S. Bakouch and Zohreh Mohammadi

A New INAR(1) Model for $\mathbb{Z}$ -Valued Time Series Using the Relative Binomial Thinning Operator

<https://doi.org/10.1515/jbnst-2022-0059>

Received September 26, 2022; accepted April 21, 2023

Abstract: A new first-order integer-valued autoregressive process (INAR(1)) with extended Poisson innovations is introduced based on a signed version of the thinning operator, called relative binomial thinning operator, which can be considered as an extension of standard binomial thinning operator introduced by Steutel, F.W. and van Harn, K. (1979. Discrete analogues of self-decomposability and stability. Ann. Probab. 7: 893–899). It is appropriate for modeling $\mathbb{Z}$ -valued time series and either positive or negative correlations. Some properties of the process are established. Conditional least squares, Yule–Walker and conditional maximum likelihood methods are considered for the parameter estimation of the model. Moreover, simulation experiments are carried out to attest to the performance of the estimation methods. The applicability of the proposed model is investigated through a practical data set of the Saudi stock market.

Keywords: time series, signed thinning operator, extended Poisson distribution, simulation, Pearson residuals

JEL Classification: C13, C15, C22, C53

1 Introduction

Discrete variable time series data are fairly common in practice, which has attracted the attention of many researchers. The early studies about modeling this type of data have been conducted by McKenzie (1985) and Al-Osh and Alzaid (1987), who

***Corresponding author: Maher Kachour**, ESSCA School of Management, Lyon, France,
E-mail: maher.kachour@essca.fr

Hassan S. Bakouch, Department of Mathematics, College of Science, Qassim University, Buraydah, Saudi Arabia; and Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt
Zohreh Mohammadi, Department of Statistics, Jahrom University, Jahrom, Iran

suggested the first-order integer-valued autoregressive (INAR(1)) models based on the binomial thinning operator introduced by Steutel and van Harn (1979). Afterward, the researchers have introduced alternative INAR(1) process with different marginal and innovation distributions based on this operator. Moreover, various modification of the binomial thinning operator have been proposed for modeling count time series. A good review on INAR models can be found in Weiß (2008) and Scotto et al. (2015).

All the above models mainly are applicable to analyze a time series with non-negative integer-valued, but in practice, one can encounter also integer-valued time series data that include negative values. For example, in stock market, we analyze intra-daily stock prices the changes belongs on a discrete valued set, which can be represented by $\mathbb{Z}$, known also as ticks (the price can go up or down on certain predefined ranges of value). In image analysis, we model the distribution of intensity differences between two neighborhood pixels. Thus, since intensities are discrete actually, we have the difference of two discrete variables. In football, the results of a match can be expressed in $\mathbb{Z}$ as the difference in the number of goals, a value often used in sport betting. In medicine, for some diseases we care about the difference in some count variables, such as number of pimples, before and after some treatment is applied. Moreover, in some fields, e.g. meteorology, by sake of simplicity and to facilitate the reading and interpret, data recorded are rounded, which lead to data with $\mathbb{Z}$ as support. Furthermore, when we analyze a non-stationary integer-valued time series with non-negative values, we use the differencing operator to achieve stationary. Therefore, we may obtain a time series on $\mathbb{Z}$.

To the best of our knowledge, limited studies have been conducted on modeling time series data defined on the set $\mathbb{Z}$, with an autoregressive structure. Kim and Park (2008) have proposed an integer-valued autoregressive process of order $p \geq 1$ with signed binomial thinning operator. Kachour and Yao (2009) have introduced the first-order rounded integer-valued autoregressive process, based on the rounding operator. A more general setup has been introduced since then by Kachour (2014). Recently, Liu et al. (2021) have proposed a semiparametric autoregressive model with a log-concave innovation, based on the model proposed by Kachour (2014). Zhang et al. (2010) have presented an integer-valued autoregressive processes of order $p \geq 1$ with signed generalized power series thinning operator. Kachour and Truquet (2011) have proposed an extension of INAR models using a modified version of the generalized thinning operator, which is different from that has been introduced by Kim and Park (2008). Zhang et al. (2012) have introduced a random coefficient process called generalized random coefficient first-order integer-valued autoregressive

process with signed thinning operator. In literature, there exists a family of models defined on $\mathbb{Z}$ that arise as the difference between two discrete distributions. For example, Freeland (2010) has defined the true integer-valued autoregressive process of order one as the difference of two Poisson INAR(1) processes, Barreto-Souza and Bourguignon (2015) have proposed the skew INAR(1) process, which is defined as the difference of two INAR(1) process with geometric marginal distribution, Bourguignon and Vasconcellos (2016) have defined the new skew integer-valued process as the difference between a Poisson INAR(1) process and a geometric INAR(1) process, and Taveira da Cunha et al. (2018) have proposed a new integer-valued autoregressive process with generalized Poisson difference marginal distributions based on difference of two quasi-binomial thinning operators. It is also important to mention the existence of certain studies that deal with the multivariate $\mathbb{Z}$ -valued autoregressive models. For example, Bulla et al. (2011) have introduced the bivariate autoregressive integer-valued time-series models, based on the signed thinning operator. More recently, Chen et al. (2023) have proposed a new bivariate $\mathbb{Z}$ -valued autoregressive model, based on the bivariate Skellam distribution.

In this paper we introduce a new integer-valued process, denoted by EP-RBINAR(1), by considering a parametric assumption on the common distribution of the counting sequence of the signed integer-valued autoregressive process proposed by Kachour and Truquet (2011). Thus, the EP-RBINAR(1) can fit integer-valued times series with possible negative values. Moreover, similar to an AR(1) process, the autocorrelation function of EP-RBINAR(1) can also have negative values. Indeed, unlike models that arise as the difference between two discrete distributions, EP-RBINAR(1) process can fit $\mathbb{Z}$ -valued time series without take into account this specific construction. Moreover, in comparison with the model proposed by Kachour and Yao (2009), the EP-RBINAR(1) doesn't have the issue related to small lack of identifiability on the parameter. Note that, the thinning part of the new process can be considered as the sum of a "discrete random walk", where one can go a one step forward, or one step back, or keep still. Furthermore, innovations of EP-RBINAR(1) process follow the extended Poisson distribution introduced by Bakouch et al. (2016). This distribution is defined on $\mathbb{Z}$, having a dispersion flexibility, and under some assumptions, it can be approximated by the Gaussian distribution.

The paper is structured as follows. EP-RBINAR(1) process is formally defined in Section 2 and some of its properties are outlined. In Section 3, estimation methods for the process parameters are proposed. Section 4 discusses some simulation results for the estimation methods. Moreover, the EP-RBINAR(1) model is applied to a practical data set of the Saudi stock market. Finally, the proofs of all propositions and theorems are contained in the appendix.

2 The EP-RBINAR(1) Process

2.1 Definition

In this section, we first consider a special version of the signed thinning operator, originally proposed by Latour and Truquet (2008). It is referred to “relative binomial thinning operator” and is an extension of the classical Steutel and van Harn operator to $\mathbb{Z}$ -valued random variables. Then, we propose a new signed integer-valued autoregressive process based on this operator with extended Poisson innovations.

Definition 1. (Relative binomial thinning operator) Let $\{Y_i\}_{i \in \mathbb{N}}$ be a sequence of independent identically distributed (i.i.d) integer-valued random variables with the common distribution F , and independent of an integer-valued random variable X . The relative binomial thinning operator, denoted by $F \circ$, is defined as

$$F \circ X = \begin{cases} \text{sign}(X) \sum_{i=1}^{|X|} Y_i, & \text{if } X \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where for any integer $x \neq 0$, $\text{sign}(x) = 1$ if $x > 0$ and -1 if $x < 0$, and F is defined as follow

$$\mathbb{P}(Y_i = y) = \begin{cases} \alpha^2, & \text{if } y = 1 \\ 2\alpha(1 - \alpha), & \text{if } y = 0 \\ (1 - \alpha)^2, & \text{if } y = -1 \end{cases} \quad (2)$$

with $0 \leq \alpha \leq 1$ (F is called a relative Bernoulli distribution).

Remark 1. For classical Steutel and van Harn operator, X is positive and $\{Y_i\}_{i \in \mathbb{N}}$ is a sequence of Bernoulli variables (i.e. $Y_i(\Omega) = \{0, 1\}$), where “ $Y_i = 1$ ” represents “success” (survival) and “ $Y_i = 0$ ” represents “failure” (death). However, for the relative binomial thinning operator, Y_i can be seen as a “discrete random walk”, where “ $Y_i = 1$ ” can be interpreted as a one step forward, “ $Y_i = -1$ ” as one step back, and “ $Y_i = 0$ ” as keep still.

Remark 2. Note that originally Chesneau and Kachour (2012) introduced the relative binomial distribution, without giving it this name. Moreover, one can see that $Y_i \stackrel{d}{=} Z_i - 1$ where $Z_i \sim B(2, \alpha)$ with $\mathbb{E}(Y_i) = 2\alpha - 1$ and $\mathbb{V}(Y_i) = 2\alpha(1 - \alpha)$. Moreover, let l be a positive integer, such that $l > 2$. Thus, one can deduce that

$$\alpha^2 + (-1)^l (1 - \alpha)^2 = \mathbb{E}(Y^l) < +\infty$$

In the following lemma, some useful basic properties of the relative binomial thinning operator are presented.

Lemma 1. Let $x \in \mathbb{Z} \setminus \{0\}$, and set $F \circ X(X = x) = F \circ x$, then the relative binomial thinning operator in Definition 1 has the following properties:

(i) $\forall y \in \mathbb{Z}$,

$$\mathbb{P}(F \circ x = y) = \binom{2|x|}{|x| + \text{sign}(x)y} \alpha^{|x| + \text{sign}(x)y} (1 - \alpha)^{|x| - \text{sign}(x)y} \mathbb{I}_A(y),$$

where $\mathbb{I}_A(y)$ denotes the indicator function, which has the value 1 when y takes values in A and has the value 0, otherwise and $A = \{-|x|, \dots, 0, \dots, |x|\}$.

(ii) $\mathbb{E}(F \circ x) = (2\alpha - 1)x$,

(iii) $\mathbb{V}(F \circ x) = 2\alpha(1 - \alpha)|x|$,

(iv) $\mathbb{E}(s^{F \circ x}) = s^{-x} (1 - \alpha + \alpha s^{\text{sign}(x)})^{2|x|}$,

(v) $\max((2\alpha - 1)|x|, 0) \leq \mathbb{E}(|F \circ x|) \leq |x|(2\alpha + 1)$.

Using the relative binomial thinning operator in Definition 1, we now introduce a new process to be used for modeling $\mathbb{Z}$ -valued time series.

Definition 2. A sequence $\{X_t\}_{t \in \mathbb{Z}}$ is said to be a EP-RBINAR(1) (first-order Extended-Poisson Relative Binomial Intergrated-valued Autoregressive) process if it has the following representation:

$$X_t = F \circ X_{t-1} + \epsilon_t, \quad t \in \mathbb{Z} \quad (3)$$

where F is defined in (2) and $\{\epsilon_t\}_{t \in \mathbb{Z}}$ is a sequence of i.i.d random variables, called innovations, following an extended Poisson distribution introduced by Bakouch et al. (2016). It is denoted by $\epsilon_t \sim \text{E-Po}(p, \lambda)$ and defined by the following probability mass function (pmf):

$$\mathbb{P}(\epsilon_t = k) = \begin{cases} e^{-\lambda} & \text{if } k = 0 \\ p e^{-\lambda} \frac{\lambda^k}{k!} & \text{if } k = 1, 2, \dots \\ (1 - p) e^{-\lambda} \frac{\lambda^{|k|}}{|k|!} & \text{if } k = \dots, -2, -1, \end{cases} \quad (4)$$

where $\lambda > 0$ and $0 \leq p \leq 1$. All $\{Y_i\}_{i \in \mathbb{N}}$ in counting sequence are independent of ϵ_t . Moreover, ϵ_t and the sigma-algebra $\sigma(X_{t-1}, X_{t-2}, \dots)$ are supposed to be mutually independent.

Remark 3. Given that innovations (i.e. $\{\epsilon_t\}_{t \in \mathbb{Z}}$) have a known distribution, our model defined in (3), can be considered as a special case of P-SINAR(1) process introduced by Chesneau and Kachour (2012), which is a particular case of SINAR(1) process presented by Kachour and Truquet (2011).

Remark 4. The use of E-Po(p, λ) distribution guarantees innovations with possible negative values and having more dispersion flexibility (i.e. the dispersion of E-Po(p, λ) distribution depends on the value of p).

Remark 5. From Bakouch et al. (2016), the corresponding formulas for mean, variance and probability generating function (pgf) of E-Po(p, λ) distribution with the pmf (4) are:

$$\begin{aligned}\mathbb{E}(\epsilon_t) &= (2p - 1)\lambda \\ \mathbb{V}(\epsilon_t) &= \lambda + 4p(1 - p)\lambda^2 \\ \Phi_{\epsilon_t}(s) &= e^{-\lambda} [pe^{s\lambda} + (1 - p)e^{\frac{\lambda}{2}}].\end{aligned}\tag{5}$$

Obviously, if $\epsilon_t \sim \text{E-Po}(p, \lambda)$, then $|\epsilon_t| \sim \text{Po}(\lambda)$. As a result, we have

$$\mathbb{E}(|\epsilon_t|) = \mathbb{V}(|\epsilon_t|) = \lambda.$$

Moreover, for any positive integer $l > 2$, we have

$$\mathbb{E}(|\epsilon_t|^l) \leq \lambda^l e^{\frac{\lambda}{2}} < +\infty.$$

Remark 6. Under the previous parametric assumptions, the EP-RBINAR(1) process can be seen as an extension on $\mathbb{Z}$ of the Poisson INAR(1) process, originally proposed Al-Osh and Alzaid (1987).

Remark 7. Based on its construction, one can deduce that, under specific parameters values, the EP-RBINAR(1) process can provide zero-inflated integer (positive and negative) observations. For example, suppose that (by sake of simplicity) that the innovation distribution is symmetric (i.e. $p = 0.5$) and concentrated on zero (i.e. $\lambda < 1$), and α is chosen so that zero be the mode of distribution F . Thus, in this case, it is very likely that zero is the most represented value of the process. Figure 1 shows plot and distribution of $n = 1000$ observations simulated from EP-RBINAR(1) process, where actual values are $\alpha = 0.45$, $p = 0.5$, and $\lambda = 0.3$. One can see that the empirical frequency associated with zero equals 60 % and there is an almost balance between the distribution of positive and negative values.

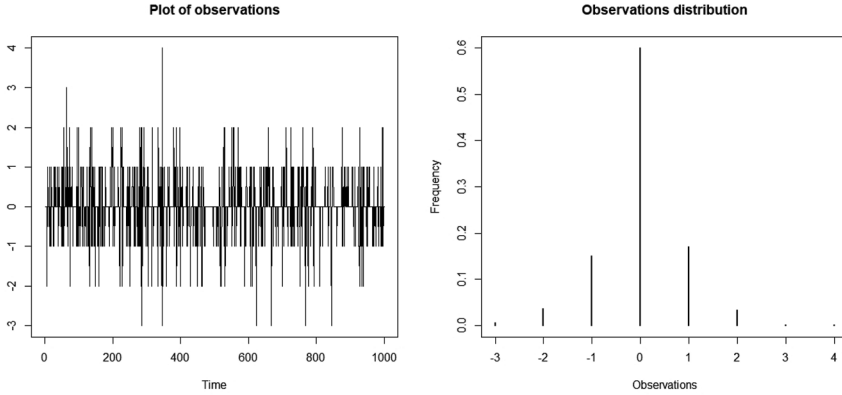


Figure 1: From left to right: plot and distribution of $n = 1000$ observations simulated from EP-RBINAR(1) process, where actual values are $\alpha = 0.45$, $p = 0.5$, and $\lambda = 0.3$.

2.2 Properties

The stationarity of the proposed EP-RBINAR(1) process is established in the following theorem.

Theorem 1. Suppose that $\alpha \in]0, 1[\setminus \{\frac{1}{2}\}$ and $0 < p < 1$. Hence, one can deduce that the process X_t defined in (3) has a unique stationary solution. Moreover, for any positive integer $l > 2$, we have

$$\mathbb{E}(|X_0|^l) < +\infty.$$

In the sequel, under stationary conditions and using Lemma 1, we derive some properties of the EP-RBINAR(1) process.

Proposition 1. Let F_j represent the filtration $\sigma(X_{j-i} | i \in \mathbb{N})$. Under stationary conditions presented in Theorem 1, one can deduce that

- (i) $\mathbb{E}(X_t | F_{t-1}) = (2\alpha - 1)X_{t-1} + (2p - 1)\lambda$.
- (ii) $\mathbb{V}(X_t | F_{t-1}) = 2\alpha(1 - \alpha)|X_{t-1}| + (\lambda + 4p(1 - p)\lambda^2)$.
- (iii) $\mathbb{E}(X_t) = \frac{(2p-1)\lambda}{2(1-\alpha)}$,
- (iv) $\mathbb{V}(X_t) = \frac{2\alpha(1-\alpha)\mathbb{E}(|X_t|) + \lambda(1+4p(1-p)\lambda)}{1 - (2\alpha-1)^2}$,
- (v) The probability generating function of the stationary EP-RBINAR(1) process is obtained as

$$\Phi_{X_t}(s) = \mathbb{E}(s^{F \circ X_{t-1}}) \Phi_{\epsilon_t}(s) \quad (6)$$

where

$$\begin{aligned} \mathbb{E}(s^{F \circ X_{t-1}}) &= \mathbb{P}(X_t = 0) + \sum_{x=-\infty}^{-1} s^{|x|} \left(1 - \alpha + \frac{\alpha}{s}\right)^{2|x|} \mathbb{P}(X_t = x) \\ &\quad + \sum_{x=1}^{\infty} \left(\frac{1}{s}\right)^x (1 - \alpha + \alpha s)^{2x} \mathbb{P}(X_t = x) \end{aligned}$$

and $\Phi_{\epsilon_t}(s)$ is the pgf of $\{\epsilon_t\}$ given in (5).

(vi) $\rho(1) = \text{corr}(X_t, X_{t-1}) = 2\alpha - 1$

Remark 8. The stationary conditions associated with the EP-RBINAR(1) process are similar to that of a classical AR(1) process. Indeed, one can see that

$$X_t = (2\alpha - 1)X_{t-1} + \xi_t,$$

where ξ_t is a stationary process defined by

$$\xi_t = \epsilon_t + F \circ X_{t-1} - (2\alpha - 1)X_{t-1}.$$

Moreover, one can see, based on properties of signed thinning operator that $\xi = (\xi_t)$ is an uncorrelated process. Thus, the EP-RBINAR(1) process has the same autocorrelation structure as a classic AR(1) process. In other words, the autocorrelation function of the EP-RBINAR(1) process is obtained as

$$\rho(k) = \text{corr}(X_t, X_{t+k}) = (2\alpha - 1)^k \quad \forall k \geq 1, \text{ with } \alpha \in]0, 1[\setminus \left\{\frac{1}{2}\right\}.$$

Therefore, the spectral density function of the model, for $\omega \in (-\pi, +\pi)$; is obtained as

$$\begin{aligned} f_{X,X}(\omega) &= \frac{1}{2\pi} \sum_{k=-\infty}^{+\infty} \text{cov}(X_t, X_{t-k}) e^{-i\omega k}, \\ &= \frac{\mathbb{V}(X_t)}{2\pi} \sum_{k=-\infty}^{+\infty} (2\alpha - 1)^k e^{-i\omega k}, \\ &= \frac{\mathbb{V}(X_t)}{2\pi} \frac{1}{(1 + (2\alpha - 1)^2 - 2(2\alpha - 1)\cos(\omega))}. \end{aligned}$$

Under stationary conditions, the marginal probability function of the EP-RBINAR(1) process is given in the following proposition.

Proposition 2. Let $\{X_t\}$ be an EP-RBINAR(1) process, with $0 < \alpha < 1$, $\alpha \neq \frac{1}{2}$, and $0 < p < 1$. Thus, for all $a \in \mathbb{Z}$, we have

$$\begin{aligned}\mathbb{P}(X_t = a) &= \sum_{i \in \mathbb{Z}^*} \sum_{j=-|i|}^{|i|} \binom{2|i|}{|j| + \text{sign}(j)i} \alpha^{|j| + \text{sign}(j)i} (1 - \alpha)^{|j| - \text{sign}(j)i} \\ &\quad \times \mathbb{P}(X_{t-1} = j) \mathbb{P}(\epsilon_t = a - i) + \mathbb{P}(X_t = 0) \mathbb{P}(\epsilon_t = a),\end{aligned}$$

where

$$\mathbb{P}(\epsilon_t = a - i) = \begin{cases} e^{-\lambda} & \text{if } a = i \\ pe^{-\lambda} \frac{\lambda^{a-i}}{(a-i)!} & \text{if } i < a \\ (1-p)e^{-\lambda} \frac{\lambda^{|a-i|}}{|a-i|!} & \text{if } i > a. \end{cases}$$

The EP-RBINAR(1) process's one-step transition probability, denoted by $\mathbb{P}_{i,j} = \mathbb{P}(X_{t+1} = j | X_t = i)$, is derived as

$$\mathbb{P}_{0,j} = \mathbb{P}(\epsilon_t = j) \quad (7)$$

and for $i \neq 0$

$$\mathbb{P}_{i,j} = \sum_{k=-|i|}^{|i|} \binom{2|i|}{|i| + \text{sign}(i)k} \alpha^{\text{sign}(i)(i+k)} (1 - \alpha)^{\text{sign}(i)(i-k)} \mathbb{P}(\epsilon_t = j - k), \quad (8)$$

where $\mathbb{P}(\epsilon_t = j)$ is the pmf of $\{\epsilon_t\}$ defined by (4).

Moreover, using the first-order dependence of the process, the joint probability can be obtained as

$$\begin{aligned}\mathbb{P}(X_1 = i_1, X_2 = i_2, \dots, X_m = i_m) &= \mathbb{P}(X_1 = i_1) \mathbb{P}(X_2 = i_2 | X_1 = i_1) \\ &\quad \times \mathbb{P}(X_m = i_m | X_{m-1} = i_{m-1}) \\ &= \mathbb{P}(X_1 = i_1) \prod_{s=1}^{m-1} \mathbb{P}_{i_s, i_{s+1}}.\end{aligned} \quad (9)$$

The k step-ahead conditional mean (usually used for time series forecast issues) of the process is derived in the following proposition.

Proposition 3. Let $\{X_t\}$ be an EP-RBINAR(1) process with $0 < \alpha < 1$, $\alpha \neq \frac{1}{2}$, and $0 < p < 1$. Thus, the k step-ahead conditional mean is given by

$$\mathbb{E}(X_{t+k} | X_t) = (2\alpha - 1)^k X_t + (2p - 1)\lambda \left(\frac{1 - (2\alpha - 1)^k}{2(1 - \alpha)} \right). \quad (10)$$

Remark 9. Using Equation (10), we find that

$$\lim_{k \rightarrow \infty} \mathbb{E}(X_{t+k}|X_t) = \frac{(2p-1)\lambda}{2(1-\alpha)},$$

which is the unconditional mean of the EP-RBINAR(1) process.

3 Parameters Estimation

Let $(X_1, \dots, X_m)^t$ be a vector of observations from our model defined in (3), and $\theta \in \{(\alpha, p, \lambda)^t; 0 < \alpha < 1, \alpha \neq \frac{1}{2}, 0 < p < 1, \lambda > 0\}$ denote the unknown parameter vector. In order to estimate θ , we propose three estimation methods, namely, Yule–Walker (YW), conditional least square (CLS), and conditional maximum likelihood (CML).

3.1 Yule–Walker Estimation

Let $\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i$, $|\bar{X}| = \frac{1}{m} \sum_{i=1}^m |X_i|$ and

$$\hat{\rho}(1) = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)} = \frac{\sum_{k=1}^{m-1} (X_t - \bar{X})(X_{t+1} - \bar{X})}{\sum_{k=1}^m (X_t - \bar{X})^2}.$$

In order to obtain estimations in this method, we solve a set of equations that are resulted from equating the theoretical and empirical aspects at the same time. Thus, the YW estimation of α is obtained by

$$\hat{\alpha}_{YW} = \frac{\hat{\rho}(1) + 1}{2}, \quad (11)$$

and the YW estimation of p and λ are obtained by solving the following equations

$$\begin{cases} \hat{\gamma}(0)(1 - \hat{\rho}^2(1)) = 2\hat{\alpha}_{YW}(1 - \hat{\alpha}_{YW})|\bar{X}| + \lambda(1 + 4p(1 - p)\lambda), \\ 2\bar{X}(1 - \hat{\alpha}_{YW}) = \lambda(2p - 1). \end{cases} \quad (12)$$

Remark 10. A special case of our model, when we suppose that $\alpha = p$. In this case, the YW estimation of α (and p) is

$$\hat{\alpha}_{YW} = \frac{\hat{\rho}(1) + 1}{2},$$

and the YW estimation of λ is given by

$$\hat{\lambda}_{YW} = \frac{2(1 - \hat{\alpha}_{YW})}{2\hat{\alpha}_{YW} - 1} \bar{X} = \left(\frac{1}{\hat{\rho}(1)} - 1 \right) \bar{X}.$$

3.2 Conditional Least Square Estimation

Let $\gamma = 2\alpha - 1$ and $\mu = (2p - 1)\lambda$, then the CLS estimator of $\theta^* = (\gamma, \mu)$ is defined by

$$\hat{\theta}_m^* = \arg \min_{\theta' \in \Theta} S_m(\theta'), \quad (13)$$

where $\hat{\theta}_m^* = (\hat{\gamma}_m, \hat{\mu}_m)$ and

$$S_m(\theta^*) = \frac{1}{m} \sum_{t=2}^m (X_t - \mathbb{E}(X_t | F_{t-1}))^2 = \frac{1}{m} \sum_{t=2}^m (X_t - \gamma X_{t-1} - \mu)^2.$$

Remark 11. Based on Theorem 1, we have $\mathbb{E}(X_0^2) < \infty$ and $\mathbb{E}(|X_0|^3) < \infty$. Thus, based on Theorem 2 of Kachour and Truquet (2011), one can deduce that

- $\hat{\theta}_m^*$ is strongly consistent (i.e. $\lim_{m \rightarrow \infty} \hat{\theta}_m^* = \theta^*$ a.s.)
- The CLS estimation is asymptotically normal.

Remark 12. A special case of the proposed model, when we suppose that $\alpha = p$. In this case, the CLS estimation of α (and p) is given as

$$\hat{\alpha}_{CLS} = \frac{\hat{\gamma}_m + 1}{2},$$

and the CLS estimation of λ is given by

$$\hat{\lambda}_{CLS} = \frac{\hat{\mu}_m}{2\hat{\alpha}_m - 1} = \frac{\hat{\mu}_m}{\hat{\gamma}_m}.$$

3.3 Conditional Maximum Likelihood Estimation

From the joint probability function (9), the conditional log-likelihood function for the EP-RBINAR(1) model can be written as

$$\begin{aligned}
 l_m(\theta|X_1) &= \sum_{s=1}^{m-1} \log[\mathbb{P}(\epsilon_t = X_{s+1}) \mathbb{I}_{\{0\}}(X_s) + \mathbb{I}_{\mathbb{Z}^*}(X_s) \\
 &\quad \times \sum_{k=-|X_s|}^{|X_s|} \binom{2|X_s|}{|X_s| + \text{sign}(X_s)k} \alpha^{\text{sign}(X_s)(X_s+k)} \\
 &\quad \times (1 - \alpha)^{\text{sign}(X_s)(X_s-k)} \mathbb{P}(\epsilon_t = X_{s+1} - k)],
 \end{aligned} \tag{14}$$

where $\mathbb{P}(\epsilon_t = j)$ is the pmf of $\{\epsilon_t\}$ defined by (4) and $\mathbb{I}_A(X_s)$ denotes the indicator function, which has the value 1 when X_s takes values in A and has the value 0, otherwise.

The conditional maximum likelihood (CML) estimator $\hat{\theta}_{CML} = (\hat{\alpha}_{CML}, \hat{p}_{CML}, \hat{\lambda}_{CML})'$ of $\theta = (\alpha, p, \lambda)'$ is defined as the value of θ that maximize the conditional log-likelihood function in (14). Since equating the first-order conditional log-likelihood derivatives to zero leads us to a complicated system of equations, the CML estimates are achieved using numerical methods.

4 Empirical Study

This section includes two subsections. In the first part, the performance of estimation methods that we used for parameters of the proposed process is compared through a simulation study. To ensure the applicability of the proposed model, the second part is devoted to analyzing a practical data set of the Saudi stock market.

4.1 Simulation

4.1.1 The Case When $\alpha = p$

In this section, we aim to test the efficiency of the parameter's estimation discussed in the previous section. For the sake of simplicity, we suppose that $\alpha = p$. Thus, we simulate (using the R programming language) 1000 paths of length 100, 500, 1000 and 10,000. These paths are simulated using Equation (3) with three sets of parameters: (a) $\alpha = p = 0.1$ and $\lambda = 1$; (b) $\alpha = p = 0.6$ and $\lambda = 2$; (c) $\alpha = p = 0.9$ and $\lambda = 3$. Mean values of YW, CLS, and CML estimates for each set of parameters are given in Table 1. The standards deviations of the estimates are stated in brackets under the estimated values.

Remark 13. In the case when $\alpha = p$, one can use results presented in Remarks 10 and 12 to calculate YW and CLS estimates. However, CML estimates are calculate by using

Table 1: Estimated parameters and the corresponding standard errors (in brackets) stated under the estimates for YW, CLS and CML methods (when $\alpha = p$).

Size	$\hat{\alpha}_{YW}$	$\hat{\lambda}_{YW}$	$\hat{\alpha}_{CLS}$	$\hat{\lambda}_{CLS}$	$\hat{\alpha}_{CML}$	$\hat{\lambda}_{CML}$
a) $\alpha = p = 0.1$ and $\lambda = 1$						
$n = 100$	0.110 (0.034)	1.009 (0.176)	0.109 (0.034)	1.010 (0.175)	0.101 (0.028)	0.977 (0.120)
$n = 500$	0.102 (0.015)	1.002 (0.076)	0.101 (0.015)	1.002 (0.076)	0.102 (0.011)	1.000 (0.050)
$n = 1000$	0.101 (0.010)	1.002 (0.054)	0.101 (0.010)	1.002 (0.054)	0.099 (0.009)	1.006 (0.041)
$n = 10,000$	0.100 (0.003)	0.999 (0.016)	0.100 (0.003)	0.999 (0.016)	0.9999 (0.002)	1.000 (0.011)
b) $\alpha = p = 0.6$ and $\lambda = 2$						
$n = 100$	0.592 (0.050)	1.584 (21.931)	0.592 (0.050)	1.513 (23.147)	0.593 (0.033)	2.019 (0.172)
$n = 500$	0.598 (0.023)	2.251 (1.048)	0.597 (0.023)	2.252 (1.047)	0.599 (0.015)	1.987 (0.082)
$n = 1000$	0.599 (0.016)	2.101 (0.660)	0.599 (0.016)	2.101 (0.659)	0.597 (0.011)	1.999 (0.061)
$n = 10,000$	0.599 (0.005)	2.004 (0.184)	0.599 (0.005)	2.005 (0.004)	0.598 (0.052)	1.999
c) $\alpha = p = 0.9$ and $\lambda = 3$						
$n = 100$	0.887 (0.317)	3.454 (1.377)	0.883 (0.032)	3.704 (1.445)	0.899 (0.012)	2.999 (0.283)
$n = 500$	0.897 (0.013)	3.100 (0.535)	0.896 (0.136)	3.152 (0.544)	0.899 (0.005)	2.996 (0.129)
$n = 1000$	0.898 (0.009)	3.040 (0.379)	0.898 (0.009)	3.065 (0.382)	0.899 (0.003)	3.000 (0.089)
$n = 10,000$	0.899 (0.003)	3.004 (0.114)	0.899 (0.003)	3.007 (0.114)	0.899 (0.002)	3.002 (0.064)

numerical methods. Explicitly, we use the `nlm` function (Non-Linear Minimization, from package “stats”, software R), to find values that maximize the conditional log-likelihood function, denoted by (14) (where we consider that $\alpha = p$ in this equation).

In general way, from the obtained results, one can deduce that YW, CLS and CML methods provide estimates that are quite close to the actual values. Moreover, the performances of YW and CLS methods are very similar. For these methods, one can

see that when the length of series is over 500, they converge quickly to the actual parameter values. However, results of CML method are more efficient even for a small length of series. For all the proposed methods, we can notice that the standard deviation of the estimates decreases as the length series increases. Furthermore, in most of studied cases, the CLM method provides the lower standard deviation of estimates.

However, for YW and CLS methods, it is important to point the convergence rate is less quickly for λ . Indeed, this can be explained by the fact that, for both methods, the estimate of λ depends on the estimated value of α . Thus, for these methods, we find that a large current value of λ implies the standard deviation of the estimates is high, especially when the size of the series is small. This remark is illustrated thanks to the simulation results obtained for parameters set (c) (for details, see Table 1). The boxplots for the parameter set (c) are given in Figures 2–4 for CLS, YW and CML

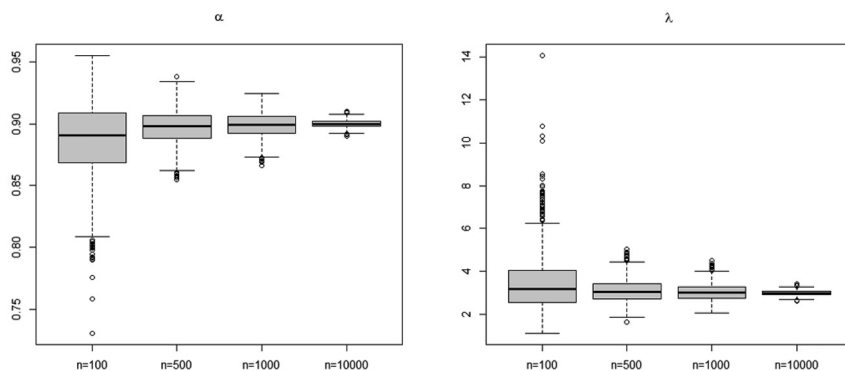


Figure 2: The boxplot for CLS estimates for parameters set (c), where actual values are $\alpha = \rho = 0.9$ and $\lambda = 3$.

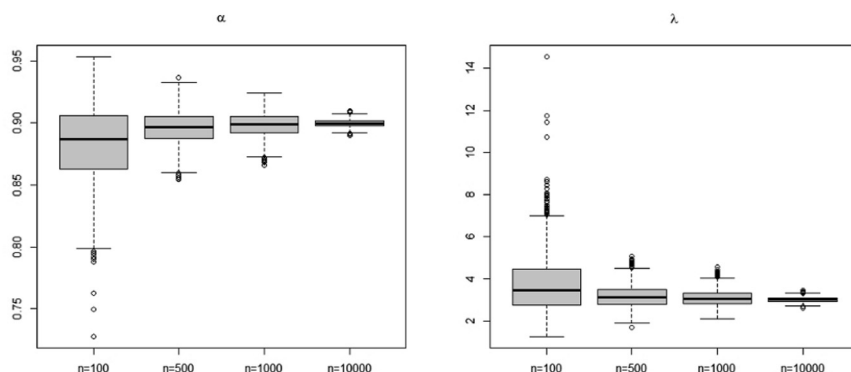


Figure 3: The boxplot for YW estimates for parameters set (c), where actual values are $\alpha = \rho = 0.9$ and $\lambda = 3$.

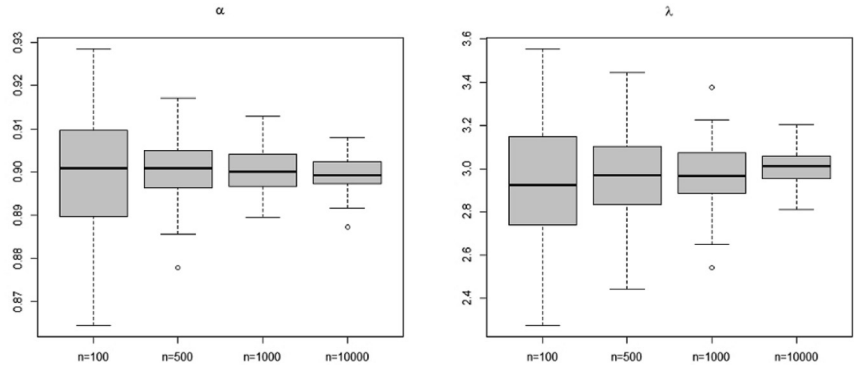


Figure 4: The boxplot for CML estimates for parameters set (c), where actual values are $\alpha = p = 0.9$ and $\lambda = 3$.

methods, respectively. One can see that, contrary to the other methods, the estimates from the CML method have few or no outliers.

On the other hand, one can see, based on the simulation results of YW and CLS methods, obtained for parameter set (b) when $n = 100$, that the standard deviation of λ estimates is very high for both methods (see Table 1). In fact, these results are not surprising cause that $\alpha (= p)$ actual value is close to 0.5 (which is an excluded value of the parameter). So, with a small size of series the estimate value of α may be close to 0.5, which can “explode” the estimate value of λ . However, this issue is not present when we use the CML method (in this case, even with a small length series, the standard deviation is weak, see Table 1). Finally, fitting to the Gaussian distribution is illustrated in Figures 5 and 6 for YW, CLS and CML estimators (with the parameter

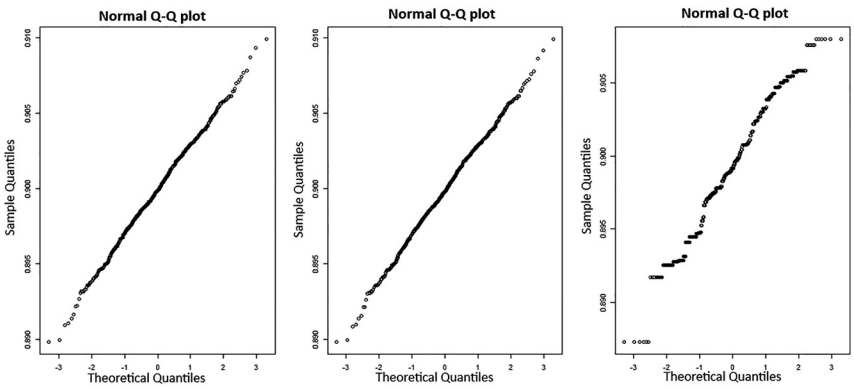


Figure 5: Normal Q-Q plots for errors $(\hat{a}_{YW} - 0.9)$, $(\hat{a}_{CLS} - 0.9)$ and $(\hat{a}_{CML} - 0.9)$, when $n = 10,000$. The horizontal axis indicates theoretical quantiles and the vertical axis indicates sample quantiles.

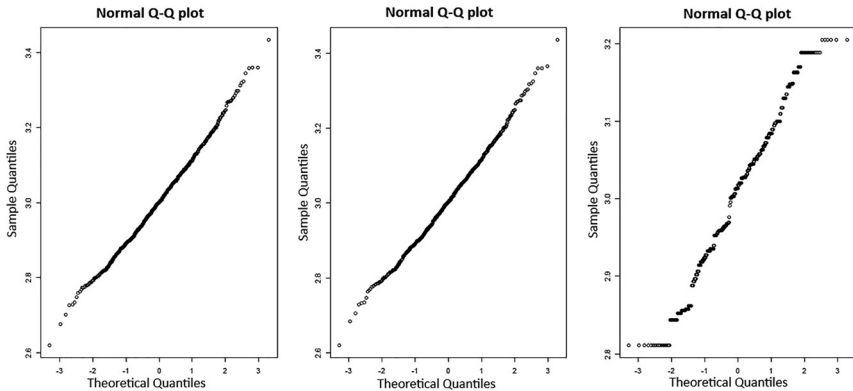


Figure 6: Normal Q-Q plots for errors $(\hat{\lambda}_{YW} - 3)$, $(\hat{\lambda}_{CLS} - 3)$ and $(\hat{\lambda}_{CML} - 3)$, when $n = 10,000$. The horizontal axis indicates theoretical quantiles and the vertical axis indicates sample quantiles.

set (c) and $n = 10,000$). These figures show numerically the normal asymptotical distribution of the proposed estimators.

4.1.2 The Case When $\alpha \neq p$

As we mentioned in the previous section, the performance of the YW and CLS methods are quite close. Moreover, we have also seen that if $\alpha \neq p$, then the estimates of these methods do not have an explicit form. Thus, in this section, we will compare the performances resulting from the YW and CML methods based on one set of parameters: $\alpha = 0.75$, $p = 0.4$, and $\lambda = 2$. We simulate (using the R programming language) 10,000 paths of length 100, 250, and 1000. These paths are simulated using Equation (3) with the above set of parameters. To calculate the CML estimates we use the *nlm* function (Non-Linear Minimization, from package “stats”, software R) to find values that maximize the conditional log-likelihood function, denoted by (14). On the other hand, the YW estimates of α is given by (11). Once we estimate the value of parameter α (based on (11)) and substitute this value into Equation (12), we get the system of two equations with two unknown variables λ and p . To find YW estimates of these parameters, this system has to be solved numerically. Thus, the numerical procedure will be conducted in programming language R, where we will use the BB package. Mean values of YW and CML estimates for each parameters are given in Table 2. The standards deviations of the estimates are stated in brackets under the estimated values. Once again, we can see that the precision of these estimates (from both methods) increases when the size n increases.

Table 2: Estimated parameters and the corresponding standard errors stated (in brackets) under the estimates for YW and CML methods (when $\alpha \neq p$).

$\alpha = 0.75, \lambda = 2, \text{ and } p = 0.4$						
Size	$\hat{\alpha}_{YW}$	$\hat{\lambda}_{YW}$	$\hat{p}_{YW}$	$\hat{\alpha}_{CML}$	$\hat{\lambda}_{CML}$	$\hat{p}_{CML}$
$n = 100$	0.739 (0.046)	1.981 (0.186)	0.389 (0.074)	0.739 (0.078)	1.997 (0.171)	0.397 (0.065)
$n = 250$	0.748 (0.015)	1.982 (0.129)	0.412 (0.045)	0.746 (0.069)	2.012 (0.103)	0.402 (0.039)
$n = 1000$	0.749 (0.015)	2.002 (0.054)	0.401 (0.019)	0.749 (0.019)	1.999 (0.052)	0.399 (0.019)

Remark 14. In order to study the compatibility of the proposed model and the estimation results, we generated, based on Equation (1), $n = 1000$ observations of the EP-RBINAR(1) process, with $\alpha = 0.75$, $p = 0.4$, and $\lambda = 2$. Basic descriptive statistics associated with the simulated data are represented in the following table:

Indicator	Mean	Variance	First autocorrelation	Second autocorrelation	Third autocorrelation
Empirical value	−0.936	8.590494	0.48065090	0.20075170	0.06104609

Using the simulated data and the procedure presented in the above section, we compute the CML estimates of the process parameters: $\hat{\alpha} = 0.7431100$, $\hat{p} = 0.3854399$ and $\hat{\lambda} = 1.9496662$. Now, using these estimates and based on the properties of the EP-RBINAR(1) process (see Proposition 1 and Remark 8), we calculate the indicators presented in the above table

Indicator	Mean	Variance	First autocorrelation	Second autocorrelation	Third autocorrelation
Estimate value	−0.8694537	8.487022	0.48622	0.2364099	0.1149472

Thus, one can see that empirical values of these indicators are very close to that of estimate values (calculated based on estimate parameters and properties of the process).

4.2 A Practical Data Example

Here, we present an application of the model EP-RBINAR(1) based on real-world data from Saudi stock market. In 2007, the minimum amount of change was 0.25 SR (Saudi Riyal) for all stocks. The daily close price as number of ticks (ticks = close price $\times$ 4) in 2007 for Saudi Telecommunication Company (STC) stock is considered. Note that these data were originally introduced by Alzaid and Omair (2014) as an application of a new integer-valued model with a Poisson difference marginal distribution. This model can be used as a tool to model non-stationary count data. Basic descriptive statistics concerning these data are presented in Table 3.

After, examination of the time series plot of the data, authors (Alzaid and Omair (2014)) notice a non-stationarity in the mean (this issue was confirmed based on a sustained large autocorrelation function (ACF) and exceptionally large first lag partial autocorrelation function (PACF)). Thus, authors considered that differencing is needed. Explicitly, if $\{X_t\}$ is the process behind the above data, then authors propose to study the lag one differenced process, denoted by $\{Y_t\}$, where $Y_t = X_t - X_{t-1}$. Indeed, one can consider that data observed from process $\{Y_t\}$ as the difference between two consecutive daily ticks (associated with the close price). Thus, a data with negative values represents the number of ticks “lost” at the close between two consecutive days (i.e. a drop in price), while a positive value reveals the number of ticks “gained” at the close between two consecutive days (i.e. a price increase). Finally, if the observed data is zero, this means that the number of ticks at closing has remained “stable” between two consecutive days (i.e. the price has not changed).

The time series plot of the differenced stock is illustrated in Figure 7 and basic descriptive statistics concerning the differenced data are presented in Table 4. Thus, Figure 7 shows the stationarity in the mean is now verified. The ACF and PACF for differenced stock are shown in Figure 8. These figures show that the lag one correlation is positive and significant, which implies that the differenced process have the same autocorrelation structure of an AR(1) process. Moreover, after the lag one differenced stage, the data is integer with negative values.

Remark 15. Bourguignon and Vasconcellos (2016) have used data, dating from 2012, providing also from Saudi Telekom. The choice of these data is consistent with the

Table 3: Basic descriptive statistics for the STC stock data set.

Length	Mean	Minimum	Maximum	First quartile	Median	Third quartile	Standard deviation
248	49.82	45	66	46	48	52	4.6968

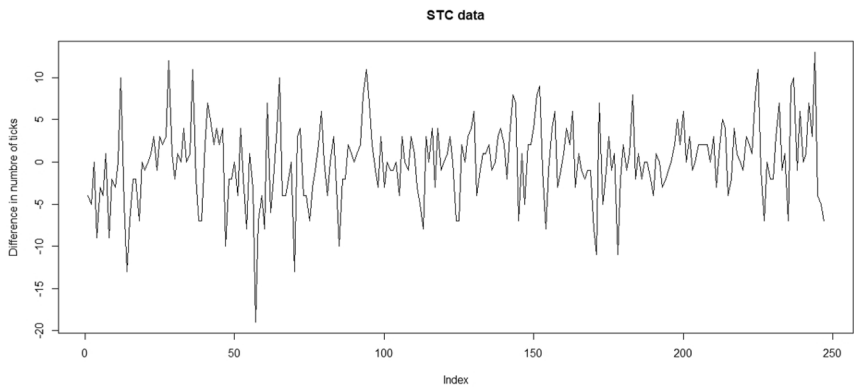


Figure 7: The time series plot of the differenced STC stock data set.

Table 4: Basic descriptive statistics for the differenced STC data set.

Length	Mean	Minimum	Maximum	First quartile	Median	Third quartile	Standard deviation
247	0.004049	−19	13	−2	0	3	4.715049

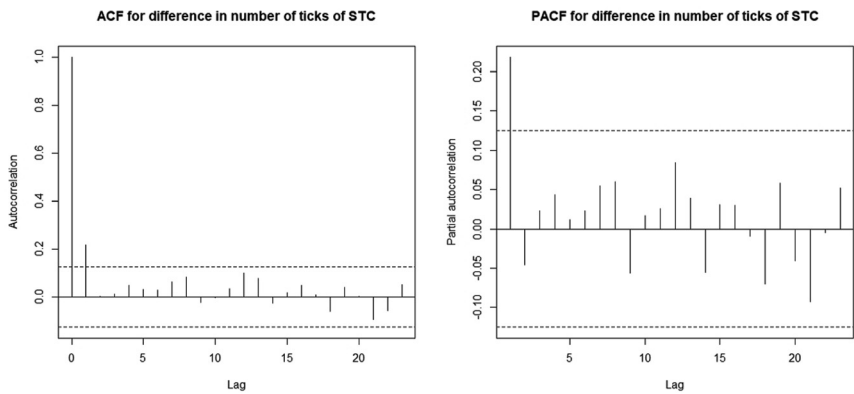


Figure 8: The ACF and PACF of the differenced STC stock data set.

model proposed by the authors (where they introduced a new process (defined on $\mathbb{Z}$) constructed as being the difference between a Poisson INAR(1) process and a geometric INAR(1) process). At the first glance, it is quite possible to consider the model proposed by Bourguignon and Vasconcellos (2016) to fit the differenced data of our application. However, we cannot say that the differenced data used in our

application comes from the difference between two sets of independent positive integer-valued data and therefore it goes against the philosophy of the model proposed by Bourguignon and Vasconcellos (2016). However, since our model belongs to SINAR(1) process family, it can be considered to fit data used by Bourguignon and Vasconcellos (2016) without take into account how data have been constructed.

To fit the differenced data, we propose our EP-RBINAR(1) process. Explicitly, we consider that

$$Y_t = F \circ Y_{t-1} + \epsilon_t,$$

where F is as defined in (2) and ϵ_t follows an E-Po(p, λ), where $0 < p < 1$, $\lambda > 0$ and $\alpha \in]0, \frac{1}{2}[\cup]\frac{1}{2}, 1[$. The estimated values of unknown parameters are obtained by using the CML method. Thus, we obtain $\hat{\alpha}_{CML} = 0.6448645$, $\hat{p}_{CML} = 0.5345621$ and $\hat{\lambda}_{CML} = 3.402455$.

Remark 16. Based on properties of the process and using the estimates of the unknown parameters, one can deduce that the mean is equal to 0.3311299 (which is not close enough to 0.004048583 (mean of the data) however, one can consider that both values are near to zero), the standard deviation equals 4.246681 (which is close to 4.715049, the standard deviation calculated from the data), and the lag-1 autocorrelation is equal to 0.289729 (which is close to 0.2194211, lag-1 autocorrelation calculated from the data). These first results support our choice of the EP-RBINAR(1) model to fit the data.

Moreover, based on the conditional one-step ahead mean, the estimated residuals are computed as

$$\begin{aligned}\hat{\epsilon}_t &= y_t - \hat{y}_t \\ &= y_t - ((2\hat{\alpha}_{CML} - 1)y_{t-1} + \hat{\lambda}_{CML}(2\hat{p}_{CML} - 1))\end{aligned}$$

Remark 17. In general, $\hat{y}_t$ are real-valued, a mapping into the discrete support of the series is obtained by rounding to the nearest integer.

To verify the adequacy of the considered model, the ACF and the PACF of the estimated residuals are plotted in Figure 9. Clearly, this graph shows that the residuals can be considered as observations from a white noise. Moreover, in Figure 10, we make a comparison between the empirical frequencies of $\hat{\epsilon}_t$ and their associated theoretical probabilities resulting from E-Po(0.5345621, 3.402455). This comparison shows that the frequencies and probabilities associated with the observed modalities are relatively close. This is one of reasons to justify our choice to use the EP-RBINAR(1) model to fit the data.

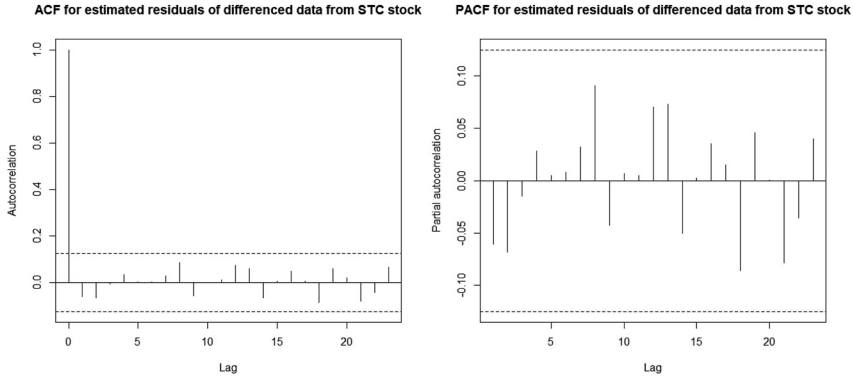


Figure 9: The ACF and PACF of the estimated residuals of the differenced STC stock data set.

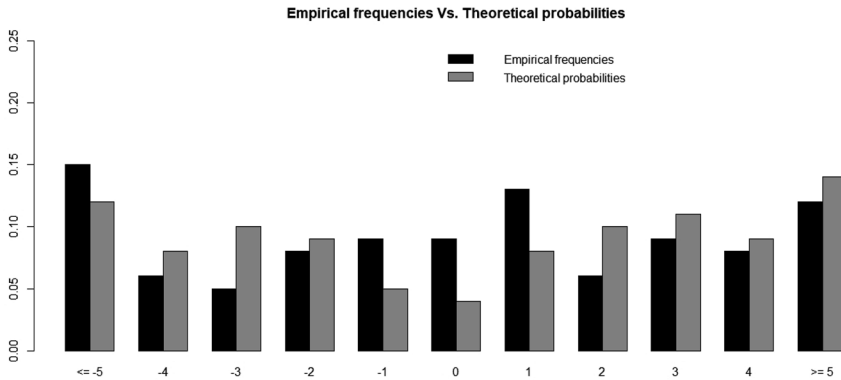


Figure 10: Comparison between the empirical frequencies of estimated residuals and their associated theoretical probabilities resulting from $E\text{-}Po(0.5345621, 3.402455)$.

Remark 18. Based on the conditional one-step ahead standard deviation, the estimated standardized Pearson residuals are computed as

$$\hat{k}_t = \frac{y_t - ((2\hat{a}_{CML} - 1)y_{t-1} + \hat{\lambda}_{CML}(2\hat{p}_{CML} - 1))}{\sqrt{2\hat{a}_{CML}(1 - \hat{a}_{CML})|y_{t-1}| + (\hat{\lambda}_{CML} + 4\hat{p}_{CML}(1 - \hat{p}_{CML})(\hat{\lambda}_{CML})^2)}}$$

For details concerning the standardized Pearson residuals and its role to check model adequacy for count time series, see Weiß et al. (2019). Note that the empirical mean, median, and standard deviation, of the estimated standardized Pearson residuals, are, respectively, equals: -0.05694 , 0 , and 1.127245 . Thus, one can see that the empirical mean is very close to zero and empirical standard deviation is close to one.

Moreover, the ACF, the PACF, and the fitting to the Gaussian distribution (Normal Q-Q plot) of the estimated standardized Pearson residuals are plotted in Figure 11. These results can be considered as additional arguments whose objective is to check the model adequacy to fit the data.

Remark 19. As mentioned before, the data processed in this section have been studied by Alzaid and Omair (2014). Indeed, authors proposed their model, called PDINAR(1) process, to fit the data. However, this process has two variants, denoted by $\text{PDINAR}^+(1)$ and $\text{PDINAR}^-(1)$, depending on the sign of the correlation. In practice, to fix which variant should be used to fit the data, it is necessary to have a preliminary idea of the sign of $\hat{\rho}(1)$. Thus, since the empirical first-order autocorrelation is positive, Alzaid and Omair (2014) considered the $\text{PDINAR}^+(1)$ to study the data. However, for the use of our model, we don't need to know the sign of $\hat{\rho}(1)$. On the other hand, Alzaid and Omair (2014) show that one-step ahead least squares predictions can be calculated by

$$\hat{y}_t = 0.2201y_{t-1} + 10.4088 - 10.3953.$$

These predicted values need also to be rounded to the nearest integer to be adequate with the discrete support of the data. From Figure 12, one can deduce that the predictions (based on conditional one-step ahead mean) results provided from $\text{PDINAR}^+(1)$ process and those from the EP-RBINAR(1) process, are very similar. This implies that both models have a prediction performances very close. Indeed, the Root Mean Squared Error (RMSE) equals 4.604699 (resp. 4.590552), for forecasting values (based on conditional mean) provided from $\text{PDINAR}^+(1)$ (resp. EP-RBINAR(1))

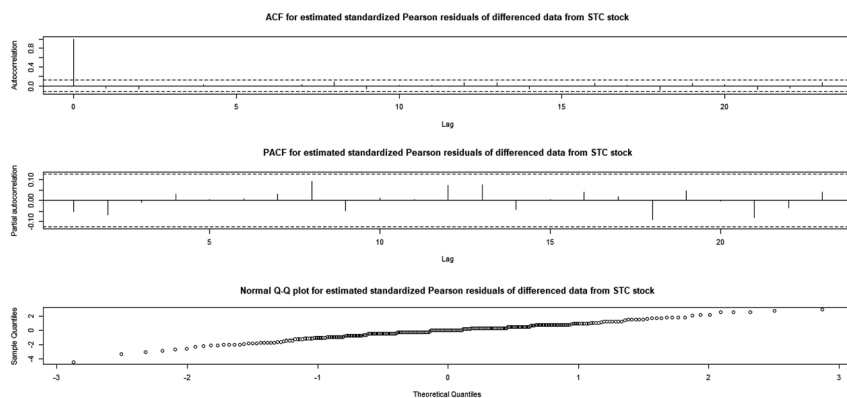


Figure 11: From top to bottom: ACF, PACF, and normal Q-Q plot of estimated standardized Pearson residuals of the differenced STC stock data set.

process. However, one can see also that for both models there is a rather discrepancy between the forecast and the actual data.

Remark 20. As mentioned above, conditional one-step ahead mean leads to real-valued forecasts, which is incoherent with the support of the integer-valued time series. Recently, coherent techniques have been introduced in order to provide integer-valued forecasts. One of the most used technique is the conditional median (for details, see Homburg et al. (2019)), which is obtained by calculating the conditional probability of each possible integer value, then selecting a forecast value with a cumulative conditional probability greater than 0.5. From Figure 13, one can see that discrepancy between conditional median forecasts and actual data have been

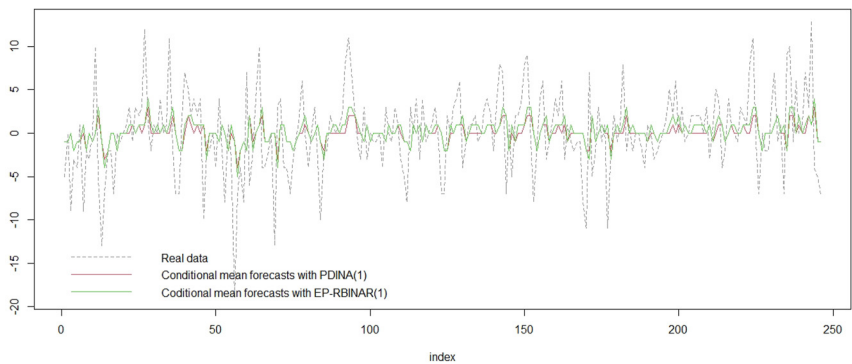


Figure 12: Comparison of conditional mean predictions results from PDINAR(1) and EP-RBINAR(1).

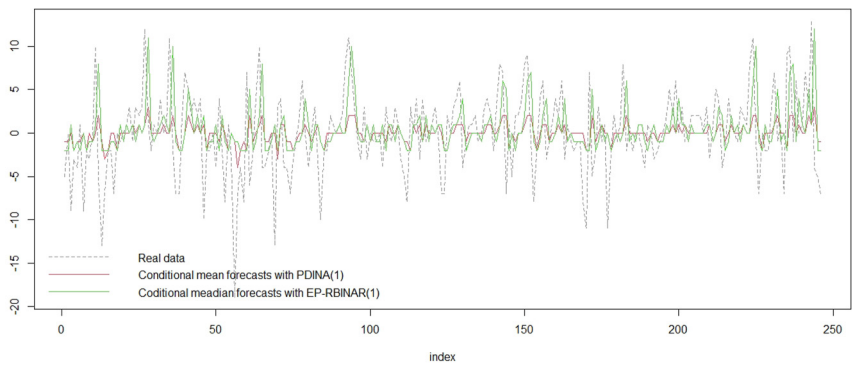


Figure 13: Comparison between conditional mean predictions results of PDINAR(1) and conditional median predictions of EP-RBINAR(1).

reduced, specially for high positive values. Moreover, one can find that the Mean Absolute Error (MAE) (based on the conditional median forecasts) equals 3.768293 and Median Absolute Error (MedAE) equals 3.

5 Concluding Remarks

In this paper, we introduce a new stationary first-order integer-valued autoregressive process, denoted by EP-RBINAR(1), where the thinning part of the process can be considered as the sum of a random walk. The new model has several advantages: possible negative values for time series, possible negative values for autocorrelation, and an innovation structure that can be interpreted as an extension on $\mathbb{Z}$ of the Poisson distribution. Moreover, under specific parameters values, EP-RBINAR(1) can provide zero-inflated integer (positive and negative) observations. The main properties of the model are derived. Then we considered the problem of parameter estimation and derived properties of the Yule–Walker, conditional least square, and conditional maximum likelihood estimators. An example of application to a real-world data set, based on number of ticks (associated with close price) for Saudi Telecommunication Company (STC) stock (in 2007), illustrates the importance and potentiality of the new model. We are convinced that this new process may attract many other wider applications in time series analysis. As part of future research, it would be of interest to extend the proposed process, studying the process of order p .

Appendix

Proof of Lemma 1. In proving all cases, we use this fact that $Y_i \stackrel{d}{=} Z_i - 1$ where $Z_i \sim B(2, \alpha)$.

(i) For any $y \in \mathbb{Z}$, we get

$$\mathbb{P}(F \circ x = y) = \mathbb{P}\left(\text{sign}(x) \sum_{i=1}^{|x|} Y_i = y\right) = \mathbb{P}\left(\sum_{i=1}^{|x|} Z_i = |x| + \text{sign}(x)y\right),$$

where $\sum_{i=1}^{|x|} Z_i$ has a binomial distribution with parameters $2|x|$ and α , $B(2|x|, \alpha)$. This completes the proof.

(ii)

$$\mathbb{E}(F \circ x) = \sum_{y \in \mathbb{Z}} y \mathbb{P}(F \circ x = y) = \sum_{z=0}^{2|x|} \text{sign}(x) (z - |x|) P(Z = z),$$

where $Z \sim B(2|x|, \alpha)$ and $\text{sign}(x)|x| = x$. So, we have

$$\mathbb{E}(F \circ x) = \text{sign}(x)(\mathbb{E}(Z) - |x|) = \text{sign}(x)(2\alpha|x| - |x|) = (2\alpha - 1)x.$$

$$\begin{aligned} \text{(iii)} \quad \mathbb{V}(F \circ x) &= \mathbb{E}((F \circ x)^2) - (\mathbb{E}(F \circ x))^2 = \sum_{y \in \mathbb{Z}} y^2 \mathbb{P}(F \circ x = y) - (2\alpha - 1)^2 x^2 \\ &= \mathbb{E}(Z^2) - 2|x|\mathbb{E}(Z) + |x|^2 - (2\alpha - 1)^2 x^2 \\ &= 2\alpha(1 - \alpha)|x|. \end{aligned}$$

$$\begin{aligned} \text{(iv)} \quad \mathbb{E}(s^{F \circ x}) &= \sum_{y \in \mathbb{Z}} s^y \mathbb{P}(F \circ x = y) \\ &= \sum_{z=0}^{2|x|} s^{\text{sign}(x)(z - |x|)} \mathbb{P}(Z = z) \\ &= s^{-x} (1 - \alpha + \alpha s^{\text{sign}(x)})^{2|x|}. \end{aligned}$$

(v) First note that

$$\begin{aligned} \mathbb{E}(|F \circ x|) &= \sum_{y \in \mathbb{Z}} |y| \mathbb{P}(F \circ x = y) \\ &= \sum_{z=0}^{2|x|} |\text{sign}(x)(z - x)| \mathbb{P}(Z = z) = \mathbb{E}(|Z - x|). \end{aligned}$$

By utilizing the fact that $\max(|a| - |b|, 0) \leq |a - b| \leq |a| + |b|$, we obtain

$$\max(\mathbb{E}(Z) - |x|, 0) \leq \mathbb{E}(|Z - x|) \leq \mathbb{E}(Z) + |x|,$$

the proof is completed by substituting $\mathbb{E}(Z) = 2|x|\alpha$. $\square$

Proof of Theorem 1. Under the assumptions of Theorem 1, one can deduce that

$$(A1) \quad F\{0\} = \mathbb{P}(Y_i = 0) > 0,$$

$$(A2) \quad \forall a \in \mathbb{Z}; \mathbb{P}(\epsilon_0 = a) > 0,$$

(A3) The root of the polynomial function

$$P(z) = 1 - (2\alpha - 1)z$$

is outside the unit disc (i.e. $\frac{1}{2\alpha-1} > 1$).

Consequently, assumptions of Theorem 1 of Kachour and Truquet (2011) are verified. Thus, one can confirm that the process X_t defined in (3) has a unique stationary solution. In other hand, from Remarks 2 and 5, for all positive integer $l > 2$, we have that

- The l th moment of F is finite,
- $\mathbb{E}(|\epsilon_t|) < +\infty$.

Once again, based on Theorem 1 of Kachour and Truquet (2011), one can deduce that $\mathbb{E}(|X_0|^l) < +\infty$. $\square$

Proof of Proposition 1. Let F_j represent the filtration $\sigma(X_{j-i} | i \in \mathbb{N})$. Thus, based on Lemma 1, we have

$$\begin{aligned}
 \text{(i)} \quad \mathbb{E}(X_t | F_{t-1}) &= \mathbb{E}(F \circ X_{t-1} | F_{t-1}) + \mathbb{E}(\epsilon_t) \\
 &= (2\alpha - 1)X_{t-1} + (2p - 1)\lambda. \\
 \text{(ii)} \quad \mathbb{V}(X_t | F_{t-1}) &= \mathbb{V}(F \circ X_{t-1} | F_{t-1}) + \mathbb{V}(\epsilon_t) \\
 &= 2\alpha(1 - \alpha)|X_{t-1}| + (\lambda + 4p(1 - p)\lambda^2). \\
 \text{(iii)} \quad \mathbb{E}(X_t) &= \mathbb{E}(F \circ X_{t-1}) + \mathbb{E}(\epsilon_t) = \mathbb{E}(\mathbb{E}(F \circ X_{t-1} | F_{t-1})) + (2p - 1)\lambda \\
 &= (2\alpha - 1)\mathbb{E}(X_t) + (2p - 1)\lambda. \\
 \text{(iv)} \quad \mathbb{V}(X_t) &= \mathbb{V}(F \circ X_{t-1}) + \mathbb{V}(\epsilon_t) \\
 &= \mathbb{V}(\mathbb{E}(F \circ X_{t-1} | F_{t-1})) + \mathbb{E}(\mathbb{V}(F \circ X_{t-1} | F_{t-1})) + \mathbb{V}(\epsilon_t) \\
 &= (2\alpha - 1)^2 \mathbb{V}(X_t) + 2\alpha(1 - \alpha)\mathbb{E}(|X_t|) + \lambda(1 + 4p(1 - p)\lambda). \\
 \text{(v)} \quad \mathbb{E}(s^{F \circ X_{t-1}}) &= \mathbb{E}(\mathbb{E}(s^{F \circ X_{t-1}} | F_{t-1})) \\
 &= s^0 \mathbb{P}(X_t = 0) + \sum_{x \in \mathbb{Z}} \mathbb{E}(s^{F \circ x}) \mathbb{P}(X_t = x). \\
 \text{(vi)} \quad \text{cov}(X_t, X_{t-1}) &= \text{cov}(F \circ X_{t-1}, X_{t-1}) \\
 &= \mathbb{E}(\mathbb{E}((F \circ X_{t-1})X_{t-1} | F_{t-1})) - (2\alpha - 1)\mathbb{E}^2(X_{t-1}) \\
 &= \mathbb{E}(X_{t-1}\mathbb{E}(F \circ X_{t-1} | F_{t-1})) - (2\alpha - 1)\mathbb{E}^2(X_{t-1}) \\
 &= (2\alpha - 1)\mathbb{E}(X_{t-1}^2) - (2\alpha - 1)\mathbb{E}^2(X_{t-1}).
 \end{aligned}$$

□

Proof of Proposition 2. Note that, for all $a \in \mathbb{Z}$, the marginal probability function of X_t can be calculated as

$$\mathbb{P}(X_t = a) = \sum_{i \in \mathbb{Z}} \mathbb{P}(F \circ X_{t-1} = i) \mathbb{P}(\epsilon_t = a - i), \quad (15)$$

where

$$\begin{aligned}
 \mathbb{P}(F \circ X_{t-1} = i) &= \sum_{j \in \mathbb{Z}} \mathbb{P}(F \circ X_{t-1} = i, X_{t-1} = j) \\
 &= \sum_{j \in \mathbb{Z}^*} \mathbb{P}(F \circ X_{t-1} = i | X_{t-1} = j) \mathbb{P}(X_{t-1} = j) \\
 &\quad + \mathbb{P}(F \circ X_{t-1} = i | X_{t-1} = 0) \mathbb{P}(X_{t-1} = 0).
 \end{aligned} \quad (16)$$

Now, by substituting (16) into (15), we have

$$\begin{aligned}\mathbb{P}(X_t = a) &= \sum_{i \in \mathbb{Z}} \sum_{j \in \mathbb{Z}^*} \mathbb{P}(F \circ j = i) \mathbb{P}(X_{t-1} = j) \mathbb{P}(\epsilon_t = a - i) \\ &\quad + \sum_{i \in \mathbb{Z}} \mathbb{P}(F \circ X_{t-1} = i | X_{t-1} = 0) \mathbb{P}(X_{t-1} = 0) \mathbb{P}(\epsilon_t = a - i),\end{aligned}\quad (17)$$

where

$$\mathbb{P}(F \circ X_{t-1} = i | X_{t-1} = 0) = \begin{cases} 1 & \text{if } i = 0, \\ 0 & \text{otherwise.} \end{cases} \quad (18)$$

The proof is completed. $\square$

Proof of Proposition 3. Using Proposition 1, the one step-ahead conditional mean equals

$$\mathbb{E}(X_{t+1}|X_t) = \mathbb{E}(F \circ X_t + \epsilon_t|X_t) = (2\alpha - 1)X_t + (2p - 1)\lambda.$$

Based on the above equation, we have

$$\begin{aligned}\mathbb{E}(X_{t+2}|X_t) &= \mathbb{E}(\mathbb{E}(X_{t+2}|X_{t+1})|X_t) \\ &= \mathbb{E}((2\alpha - 1)X_{t+1} + (2p - 1)\lambda|X_t) \\ &= (2\alpha - 1)^2 X_t + (2p - 1)\lambda[1 + (2\alpha - 1)],\end{aligned}$$

and in the same manner, one can see that

$$\begin{aligned}\mathbb{E}(X_{t+3}|X_t) &= \mathbb{E}(\mathbb{E}(X_{t+3}|X_{t+1})|X_t) \\ &= \mathbb{E}((2\alpha - 1)^2 X_{t+1} + (2p - 1)\lambda[1 + (2\alpha - 1)]|X_t) \\ &= (2\alpha - 1)^2 [(2\alpha - 1)X_t + (2p - 1)\lambda] + (2p - 1)\lambda[1 + (2\alpha - 1)] \\ &= (2\alpha - 1)^3 X_t + (2p - 1)\lambda[1 + (2\alpha - 1) + (2\alpha - 1)^2].\end{aligned}$$

Hence by induction, one can conclude that

$$\begin{aligned}\mathbb{E}(X_{t+k}|X_t) &= (2\alpha - 1)^k X_t + (2p - 1)\lambda[1 + (2\alpha - 1) + \dots + (2\alpha - 1)^{k-1}] \\ &= (2\alpha - 1)^k + (2p - 1)\lambda \left(\frac{1 - (2\alpha - 1)^k}{2(1 - \alpha)} \right).\end{aligned}$$

$\square$

References

- Al-Osh, M. and Alzaid, A.A. (1987). First-order integer-valued autoregressive (INAR(1)) process. J. Time Ser. Anal. 8: 261–275.
- Alzaid, A.A. and Omair, M.A. (2014). Poisson difference integer valued autoregressive model of order one. Bull. Malaysian Math. Sci. Soc. 37: 465–485.
- Bakouch, H.S., Kachour, M., and Nadarajah, S. (2016). An extended Poisson distribution. Commun. Stat. Theor. Methods 45: 6746–6764.

- Barreto-Souza, W. and Bourguignon, M. (2015). A skew INAR(1) process on Z. *AStA Adv. Stat. Anal.* 99: 189–208.
- Bourguignon, M. and Vasconcellos, K.L. (2016). A new skew integer valued time series process. *Stat. Methodol.* 31: 8–19.
- Bulla, J., Chesneau, C., and Kachour, M. (2011). A bivariate first-order signed integer-valued autoregressive process. *Commun. Stat. Theory Methods* 46: 6590–6604.
- Chen, H., Zhu, F., and Liu, X. (2023). Two-step conditional least squares estimation for the bivariate Z-valued INAR(1) model with bivariate Skellam innovations. *Commun. Stat. Theor. Methods*, <https://doi.org/10.1080/03610926.2023.2172587>.
- Chesneau, C. and Kachour, M. (2012). A parametric study for the first-order signed integer-valued autoregressive process. *J. Stat. Theory Pract.* 6: 760–782.
- Freeland, R.K. (2010). True integer value time series. *AStA Adv. Stat. Anal.* 94: 217–229.
- Homburg, A., Weiß, C.H., Alwan, L.C., Frahm, G., and Göb, R. (2019). Evaluating approximate point forecasting of count processes. *Econometrics* 7: 30.
- Kachour, M. (2014). On the rounded integer-valued autoregressive process. *Commun. Stat. Theor. Methods* 43: 355–376.
- Kachour, M. and Truquet, L. (2011). A p-order signed integer-valued autoregressive (SINAR(p)) model. *J. Time Ser. Anal.* 32: 223–236.
- Kachour, M. and Yao, J.F. (2009). First-order rounded integer-valued autoregressive (RINAR (1)) process. *J. Time Ser. Anal.* 30: 417–448.
- Kim, H.Y. and Park, Y. (2008). A non-stationary integer-valued autoregressive model. *Stat. Pap.* 49: 485–502.
- Latour, A. and Truquet, L. (2008). An integer-valued bilinear type model, Available at: <http://hal.archives-ouvertes.fr/hal-00373409/fr/>.
- Liu, Z., Li, Q., and Zhu, F. (2021). Semiparametric integer-valued autoregressive models on. *Can. J. Stat.* 49: 1317–1337.
- McKenzie, E. (1985). Some simple models for discrete variate time series. *Water Resour. Bull.* 21: 645–650.
- Scotto, M.G., Weiß, C.H., and Gouveia, S. (2015). Thinning-based models in the analysis of integer-valued time series: a review. *Stat. Model. Int. J.* 15: 590–618.
- Steutel, F.W. and van Harn, K. (1979). Discrete analogues of self-decomposability and stability. *Ann. Probab.* 7: 893–899.
- Taveira da Cunha, E., Vasconcellos, K.L., and Bourguignon, M. (2018). A skew integer-valued time-series process with generalized Poisson difference marginal distribution. *J. Stat. Theory Pract.* 12: 718–743.
- Weiß, C.H. (2008). Thinning operations for modeling time series of counts—a survey. *AStA Adv. Stat. Anal.* 92: 319–343.
- Weiß, C., Scherer, L., Aleksandrov, B., and Feld, M. (2019). Checking model adequacy for count time series by using Pearson residuals. *J. Time Econom.* 12: 2018–0018.
- Zhang, H., Wang, D., and Zhu, F. (2010). Inference for INAR (p) processes with signed generalized power series thinning operator. *J. Stat. Plann. Inference* 140: 667–683.
- Zhang, H., Wang, D., and Zhu, F. (2012). Generalized RCINAR (1) process with signed thinning operator. *Commun. Stat. Theor. Methods* 41: 1750–1770.