REVSTAT – Statistical Journal
Volume 21, Number 4, October 2023, 469–490
https://doi.org/10.57805/revstat.v21i4.430

Random Environment Integer-Valued Autoregressive Process with Discrete Laplace Marginal Distributions

Authors: Bogdan A. Pirković ✉
– Department of Mathematics and Informatics, Faculty of Science, University of Kragujevac, Kragujevac, Serbia
bogdanpirkovic@gmail.com

Miroslav M. Ristić
– Department of Mathematics, Faculty of Sciences and Mathematics, University of Niš, Niš, Serbia
miristic72@gmail.com

Aleksandar S. Nastić
– Department of Mathematics, Faculty of Sciences and Mathematics, University of Niš, Niš, Serbia
anastic78@gmail.com

Received: June 2021 Revised: February 2022 Accepted: February 2022

Abstract:

- A new random environment integer-valued autoregressive process of order 1 with discrete Laplace marginal distributions and with r states (abbrev. $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$) is introduced. It is shown that this process is distributed as a difference of two independent generalized random environment integer-valued autoregressive processes, when their orders are equal to 1. Other distributional and correlation properties of the $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ process are discussed. Strongly consistent Yule-Walker estimates are defined. The method of moments is implemented for different cases of simulated samples. Finally, the proposed model is applied to real-life data and the obtained results show its effectiveness.

Keywords:

- *random environment;* $\mathrm{INAR}(1)$, $\mathrm{rDLINAR}_1(\mathcal{M},\mathcal{A})$*;* $\mathrm{DLINAR}(1)$*; discrete Laplace distribution.*

AMS Subject Classification:

- 62M10.

✉ Corresponding author.

1. INTRODUCTION

The major breakthrough concerning integer-valued autoregressive (INAR) processes was made, independently of each other, by [16] and [1]. In the period that followed, many generalizations and modifications of these models were published, and many authors were trying to create their own models in order to better describe the data obtained from some natural processes. Some of them introduced new thinning operators dependent of a single parameter, as it was done by [12], [22, 23] and [13]. More recently, a thinning operator with two parameters has appeared in [15]. On contrary to this, others researchers discussed marginal distributions, as given in [2] and [3]. In recent years, authors have been trying to model some specific count data. Thus, data sets with excess zeros and excess ones are modeled in [20], while a model for modeling heavily-tailed count data is proposed in [21]. Most of the introduced models were applied to non-negative data, although in many real-life situations there are processes which may consist of integer values including both positive and negative numbers. A step forward in this direction was made by [9], who introduced a true integer-valued process, defined in distribution as a difference of two non-negative, independent INAR processes with Poisson marginal distributions. In the same way, [17] introduced an INAR model with discrete Laplace marginal distributions (DLINAR(1)), defined in distribution as a difference of two non-negative, independent INAR processes with the same geometric marginal distributions. Some generalizations of this idea emerged in the work of [4] and [7]. Lately, [6] and [14] have also come up with innovative ideas for creating models on $\mathbb{Z}$. Although each of these models deserves attention, the DLINAR(1) model introduced by [17] is of particular importance for this paper.

All the processes mentioned here are stationary, since stationarity simplifies the calculation when describing the model and determining the estimates of model parameters. But, we can say that stationary processes are rigid, since some of their properties are conserved in time. Nevertheless, the real data are not usually like that. The first non-stationary model appeared in [18] and is flexible towards the environment conditions changes. Namely, quantitative properties of phenomena from nature depend on environment, so it's logical, as well, to assume the distribution to depend on environment. It is supposed that environment conditions can be divided into r different types, called states. Each state is associated with a fixed distribution, and an element of the process has the distribution of its state. Various authors have tried to generalize or modify this idea in recent years. They assumed that the order of the model, or even the thinning parameter value, are also determined by the environment state at a particular moment, as it was done by [11] and [19].

Although the DLINAR(1) model successfully estimates the data it was tested on, due to its property of stationarity, the model shows substantial difficulties in adjusting to the elements that deviate significantly from zero. In particular, the model is struggling to estimate the highest and the lowest peaks, with a large difference between real values and their estimated values. This fact leaves room for model improvement. The main goal of this article is to make the DLINAR(1) model more flexible, using the idea given in [18]. In other words, the goal is to construct the DLINAR(1) process dependent on the environment states. Nevertheless, the idea mentioned in [18] cannot be fully taken over and certain adaptations have to be made. So, in Section 2 of this article, the construction of such a process, which overcomes problems mentioned above, is given alongside with its main distributional properties.

Section 3 provides the k-step ahead conditional expectation and a correlation structure. Yule-Walker (YW) estimates of the parameters of the defined model are given in Section 4. In Section 5, the quality of YW estimates is examined on simulated data. Section 6 deals with forecasting and provides a criterion to compare the prediction results between different models. An application of the introduced model to some real-life data is presented in Section 7, and results are compared for different models.

2. CONSTRUCTION OF THE PROCESS

As mentioned in the previous section, the first attempt to increase the flexibility of the DLINAR(1) process followed the idea given in [18]. An attempt to construct such an improved process brought some difficulties, because the newly acquired process had the same shape of one-step ahead conditional expectation as it was the case with DLINAR(1) process. To avoid this issue, the flexibility of the DLINAR(1) process is improved using the concept given in [11], although in a bit simpler form. Namely, it is assumed that information about the environment state is not only carried by the marginal distribution parameter, but it can also be expressed through the thinning parameter value. In other words, we assume that the value of the marginal distribution parameter and the value of the thinning parameter, both in moment n, depend on environment state in the same moment. A new INAR process with discrete Laplace marginal distributions, that meets the aforementioned assumptions, is defined in this section and some of its properties are discussed.

In order to make the reading of the manuscript easygoing, definitions of RrNGINAR$(\mathcal{M},\mathcal{A},\mathcal{P})$ and DLINAR(1) processes are given, since the paper relies heavily on those. As mentioned in [11], we call $\{X_n(z_n)\}$ the RrNGINAR$(\mathcal{M},\mathcal{A},\mathcal{P})$ process if its element $X_n(z_n)$ at moment $n\in\mathbb{N}$ is determined by the recursive relation

$$X_n=\begin{cases}\alpha_{z_n}*X_{n-1}(z_{n-1})+\varepsilon_n(z_n,z_n-1) & \text{w.p. } \phi_{1,P_n}^{z_n},\\ \alpha_{z_n}*X_{n-2}(z_{n-2})+\varepsilon_n(z_n,z_n-2) & \text{w.p. } \phi_{2,P_n}^{z_n},\\ \vdots & \\ \alpha_{z_n}*X_{n-P_n}(z_{n-P_n})+\varepsilon_n(z_n,z_n-P_n) & \text{w.p. } \phi_{P_n,P_n}^{z_n},\end{cases}$$

where $\{z_n\}_{n=1}^{\infty}$ is the realization of the random environment process $\{Z_n\}_{n=1}^{\infty}$ (which is a Markov chain) whose elements take values in $E_r=\{1,...,r\}$, $r\in\mathbb{N}$, for r being the number of different environment states. Probabilities $\phi_{i,P_n}^{z_n}$, $i=1,2,...,P_n$, are all in $[0,1]$ and $\sum_{i=1}^{P_n}\phi_{i,P_n}^{z_n}=1$. In addition, "$\alpha*$", $\alpha\in(0,1)$, denotes the negative binomial thinning operator defined as $\alpha*X=\sum_{i=1}^{X}U_i$. Such defined thinning operator assigns to each integer-valued random variable X the sum of X independent random variables having the same geometric distribution with the mean α. Sets $\mathcal{M}=\{\mu_1,...,\mu_r\}$, $\mathcal{A}=\{\alpha_1,...,\alpha_r\}$, $\mathcal{P}=\{p_1,...,p_r\}$ contain parameter values of the model, μ_{z_n} is the mean of the marginal geometric distribution of $X_n(z_n)$, α_{z_n} is the thinning parameter value and p_{z_n} represents the maximal value that the order P_n may take for a fixed state $z_n\in\{1,...,r\}$.

Now, let us define the thinning operator "$\alpha\odot$" as it was done in [17]. Let Y be a random variable with discrete Laplace distribution DL$(\mu/(1+\mu))$, $\mu>0$, with probability

mass function given by

$$P(Y=y)=\frac{1}{1+2\mu}\left(\frac{\mu}{1+\mu}\right)^{|y|},\quad y=0,\pm1,\pm2,...,$$

and let $X^{(1)}$ and $X^{(2)}$ be two independent random variables with the same Geom $(\mu/(1+\mu))$ distribution. In that case, operator "$\alpha\odot$" is defined as

$$\alpha\odot Y|Y\overset{d}{=}\left(\alpha*X^{(1)}-\alpha*X^{(2)}\right)|\left(X^{(1)}-X^{(2)}\right), \tag{2.1}$$

where "$\alpha*$", $\alpha\in(0,1)$, represents the negative binomial thinning operator. In addition, the counting sequences involved in $\alpha*X^{(1)}$ and $\alpha*X^{(2)}$ are mutually independent and independent of random variables $X^{(1)}$ and $X^{(2)}$.

Using this newly defined thinning operator, [17] defined the DLINAR(1) process in the following way:

$$Y_n=\alpha\odot Y_{n-1}+e_n,\ n\in\mathbb{N},$$

where $\{Y_n\}$ represents a discrete Laplace distributed process, while $\{e_n\}$ is an innovation sequence of independent and identically distributed (i.i.d.) random variables, such that e_n and Y_{n-l} are mutually independent for all $l>0$.

For the purpose of better understanding the content which follows, it is convenient to introduce here as well the skewed discrete Laplace distribution SDL$(\mu/(1+\mu),\nu/(1+\nu))$, $\mu>0,\ \nu>0$, with probability mass function given by

$$P(Y=y)=\begin{cases}\dfrac{1}{1+\mu+\nu}\left(\dfrac{\mu}{1+\mu}\right)^{y}, & y\geq 0,\\[2ex] \dfrac{1}{1+\mu+\nu}\left(\dfrac{\nu}{1+\nu}\right)^{-y}, & y<0.\end{cases}$$

Following notations given in the definition of RrNGINAR$(\mathcal{M},\mathcal{A},\mathcal{P})$, let $E_r=\{1,2,...,r\}$ be the set of all possible environment states, where $r\in\mathbb{N}$, and let $\{z_n\},\ n\in\mathbb{N}_0$, be a realization of the r states random environment process $\{Z_n\}$. For $i,j\in E_r$, let $\{e_n(i,j),\ n\in\mathbb{N}\}$ be the sequences of i.i.d. random variables. Notation $Y_n(z_n)$ will be used to tag an element of a new process, where z_n represents the realized value of the random environment process in moment $n\geq 0$. Regarding this, let us introduce the following notations:

$$\begin{aligned} Y_n(Z_n) &= \sum_{z=1}^{r} Y_n(z)I_{\{Z_n=z\}},\\ e_n(Z_{n-1},Z_n) &= \sum_{z_1=1}^{r}\sum_{z_2=1}^{r} e_n(z_1,z_2)I_{\{Z_{n-1}=z_1,Z_n=z_2\}},\\ \alpha_{Z_n} &= \sum_{z=1}^{r}\alpha_z I_{\{Z_n=z\}}, \end{aligned}$$

whereby $I_{\{Z_n=z\}}$ represents an indicator random variable associated with the event $Z_n=z$.

Before introducing the definition of a new process that will be in the focus of this research, it is necessary to define a random environment INAR process based on the thinning operator "$\alpha\odot$", with variable marginal distribution and inconstant thinning parameter value.

Definition 2.1. Let $\{Z_n\}$ be a random environment process with r possible states from the set $E_r = \{1, 2, ..., r\}$, $r \in \mathbb{N}$. Let $\mathcal{M} = \{\mu_1, \mu_2, ..., \mu_r\}$ and $\mathcal{A} = \{\alpha_1, \alpha_2, ..., \alpha_r\}$, with $\mu_i > 0$ and $\alpha_i \in (0, 1)$, for all $i \in E_r$. We say that $\{Y_n(Z_n)\}$ is a random environment INAR process of order 1 based on the thinning operator "$\alpha \odot$", with r states, distribution parameters set $\mathcal{M}$ and thinning parameters set $\mathcal{A}$ ($\mathrm{RrINAR}_1(\mathcal{M}, \mathcal{A})$), if the random variable $Y_n(Z_n)$ is defined for $n \geq 1$ as

$$Y_n(Z_n) = \alpha_{Z_n} \odot Y_{n-1}(Z_{n-1}) + e_n(Z_{n-1}, Z_n), \tag{2.2}$$

where "$\alpha_{Z_n} \odot$" is defined by (2.1) and the following conditions are satisfied:

1. For fixed $i, j \in E_r$, the sequence $\{e_n(i, j)\}_{n\in\mathbb{N}}$ is a sequence of i.i.d. random variables;
2. The sequences of random variables $\{Z_n\}, \{e_n(1, 1)\}, \{e_n(1, 2)\}, ..., \{e_n(r, r)\}$ are mutually independent;
3. Random variables Z_m and $e_m(i, j)$ are independent of $Y_n(l)$ for all $n < m$ and all $i, j, l \in E_r$.

It is convenient now to define a random environment INAR process of order 1 with discrete Laplace marginals. In order to simplify the process, we can assume we know a realization $\{z_n\}$ of the random environment process $\{Z_n\}$. This assumption is plausible, since the estimate of $\{z_n\}$ can be easily obtained by applying the appropriate clustering procedure.

Definition 2.2. Let $\{z_n\}$ be a realization of the random environment process $\{Z_n\}$ with r possible states from the set $E_r = \{1, 2, ..., r\}$, $r \in \mathbb{N}$, and let $\mathcal{M} = \{\mu_1, \mu_2, ..., \mu_r\}$ and $\mathcal{A} = \{\alpha_1, \alpha_2, ..., \alpha_r\}$, with $\mu_i > 0$ and $\alpha_i \in (0, 1)$, for all $i \in E_r$. We say that $\{Y_n(z_n)\}$ is a random environment discrete Laplace INAR process of order 1 with r states, distribution parameters set $\mathcal{M}$ and thinning parameters set $\mathcal{A}$ ($\mathrm{RrDLINAR}_1(\mathcal{M}, \mathcal{A})$), if the random variable $Y_n(z_n)$ satisfies

$$Y_n(z_n) = \alpha_{z_n} \odot Y_{n-1}(z_{n-1}) + e_n(z_{n-1}, z_n), \tag{2.3}$$

for $n \geq 1$, where conditions 1-3 from Definition 2.1 are satisfied and the random variable $Y_n(z_n)$ has $\mathrm{DL}(\mu_{z_n}/(1 + \mu_{z_n}))$ distribution, for all $n \in \mathbb{N}_0$.

The introduced process is fully determined if the distributions of random variables $e_n(i, j)$ are known for all $n \geq 1$ and all $i, j \in E_r$. The following theorem reveals distributions of these random variables.

Theorem 2.1. *Let $\{Y_n(z_n)\}$ be a RrDLINAR$_1(\mathcal{M}, \mathcal{A})$ process. Let us suppose that $z_n = j$ and $z_{n-1} = k$ for some k and $j \in E_r$. If $0 < \alpha_j \leq \frac{\mu_j}{1+\max_{i\in E_r} \mu_i}$, then the distribution of the random variable $e_n(k, j)$ can be written as a mixture of discrete Laplace and skewed discrete Laplace distributed random variables in the following form:*

$$e_n(k, j) \stackrel{d}{=} \begin{cases} DL\left(\frac{\mu_j}{1+\mu_j}\right), & w.p.\ \left(1 - \frac{\alpha_j\mu_k}{\mu_j - \alpha_j}\right)^2, \\ SDL\left(\frac{\mu_j}{1+\mu_j}, \frac{\alpha_j}{1+\alpha_j}\right), & w.p.\ \frac{\alpha_j\mu_k}{\mu_j - \alpha_j}\left(1 - \frac{\alpha_j\mu_k}{\mu_j - \alpha_j}\right), \\ SDL\left(\frac{\alpha_j}{1+\alpha_j}, \frac{\mu_j}{1+\mu_j},\right), & w.p.\ \frac{\alpha_j\mu_k}{\mu_j - \alpha_j}\left(1 - \frac{\alpha_j\mu_k}{\mu_j - \alpha_j}\right), \\ DL\left(\frac{\alpha_j}{1+\alpha_j}\right), & w.p.\ \left(\frac{\alpha_j\mu_k}{\mu_j - \alpha_j}\right)^2. \end{cases} \tag{2.4}$$

Proof: Let $\varphi_{e_n(k,j)}(t)$ represents the characteristic function of the random variable $e_n(k,j)$. Based on the definition and properties of the process and the assumption that $z_{n-1}=k$ and $z_n=j$, it holds

$$\varphi_{e_n(k,j)}(t)=\frac{\varphi_{Y_n(j)}(t)}{\varphi_{\alpha_j\odot Y_{n-1}(k)}(t)}.$$

As can be seen in [17],

$$\begin{aligned}\varphi_{Y_n(j)}(t) &= \frac{1}{(1+\mu_j-\mu_je^{it})(1+\mu_j-\mu_je^{-it})},\\ \varphi_{\alpha_j\odot Y_{n-1}(k)}(t) &= \frac{(1+\alpha_j-\alpha_je^{it})(1+\alpha_j-\alpha_je^{-it})}{(1+\alpha_j(1+\mu_k)-\alpha_j(1+\mu_k)e^{it})(1+\alpha_j(1+\mu_k)-\alpha_j(1+\mu_k)e^{-it})}.\end{aligned}$$

Using these facts, we obtain that

$$\begin{aligned}\varphi_{e_n(k,j)}(t) &= \frac{[1+\alpha_j(1+\mu_k)-\alpha_j(1+\mu_k)e^{it}][1+\alpha_j(1+\mu_k)-\alpha_j(1+\mu_k)e^{-it}]}{(1+\alpha_j-\alpha_je^{it})(1+\alpha_j-\alpha_je^{-it})(1+\mu_j-\mu_je^{it})(1+\mu_j-\mu_je^{-it})}\\ &= \frac{A}{(1+\alpha_j-\alpha_je^{it})(1+\alpha_j-\alpha_je^{-it})}+\frac{B}{(1+\alpha_j-\alpha_je^{it})(1+\mu_j-\mu_je^{-it})}\\ &\quad+\frac{C}{(1+\mu_j-\mu_je^{it})(1+\mu_j-\mu_je^{-it})}+\frac{D}{(1+\mu_j-\mu_je^{it})(1+\alpha_j-\alpha_je^{-it})}.\end{aligned}$$

By solving the system

$$\begin{aligned}AM^2+BKM+CK^2+DMK &= (1+\alpha_j(1+\mu_k))^2,\\ AM\mu_j+BK\mu_j+CK\alpha_j+DM\alpha_j &= \alpha_j(1+\mu_k)(1+\alpha_j(1+\mu_k)),\\ AM\mu_j+BM\alpha_j+CK\alpha_j+DK\mu_j &= \alpha_j(1+\mu_k)(1+\alpha_j(1+\mu_k)),\\ A\mu_j^2+B\alpha_j\mu_j+C\alpha_j^2+D\alpha_j\mu_j &= (\alpha_j(1+\mu_k))^2,\end{aligned}$$

where $K=1+\alpha_j$ and $M=1+\mu_j$, we obtain that

$$A=\left(\frac{\alpha_j\mu_k}{\mu_j-\alpha_j}\right)^2,\ B=D=\frac{\alpha_j\mu_k}{\mu_j-\alpha_j}\left(1-\frac{\alpha_j\mu_k}{\mu_j-\alpha_j}\right),\ C=\left(1-\frac{\alpha_j\mu_k}{\mu_j-\alpha_j}\right)^2.$$

Knowing that the characteristic functions of random variables with $\mathrm{DL}\left(\frac{\mu_j}{1+\mu_j}\right)$, $\mathrm{SDL}\left(\frac{\mu_j}{1+\mu_j},\frac{\alpha_j}{1+\alpha_j}\right)$, $\mathrm{SDL}\left(\frac{\alpha_j}{1+\alpha_j},\frac{\mu_j}{1+\mu_j}\right)$ and $\mathrm{DL}\left(\frac{\alpha_j}{1+\alpha_j}\right)$ distributions are of the form

$$\begin{aligned}\varphi_1(t) &= \frac{1}{(1+\mu_j-\mu_je^{it})(1+\mu_j-\mu_je^{-it})},\quad \varphi_2(t)=\frac{1}{(1+\mu_j-\mu_je^{it})(1+\alpha_j-\alpha_je^{-it})},\\ \varphi_3(t) &= \frac{1}{(1+\alpha_j-\alpha_je^{it})(1+\mu_j-\mu_je^{-it})},\quad \varphi_4(t)=\frac{1}{(1+\alpha_j-\alpha_je^{it})(1+\alpha_j-\alpha_je^{-it})}\end{aligned}$$

respectively, it becomes obvious that (2.4) holds.

It is left to provide that A,B,C and D are probabilities, i.e. that $A+B+C+D=1$ and all of them belong to $[0,1]$. First condition is easily confirmable. To provide the second one, it is enough to confirm that $0\le\frac{\alpha_j\mu_k}{\mu_j-\alpha_j}\le 1$. By solving this double inequality, we get $\alpha_j\le\frac{\mu_j}{1+\mu_k}$. Since this condition must hold for an arbitrary k and j, and $\alpha_j\in(0,1)$, we have that $0<\alpha_j\le\frac{\mu_j}{1+\max_{i\in E_r}\mu_i}$. This fact completes the proof. □

According to the previous theorem and the fact that discrete Laplace and skewed discrete Laplace distributed random variables can be represented as a difference of two random variables with geometric distributions, it is possible to make an interesting conclusion.

Corollary 2.1. *If* $0<\alpha_j\leq\frac{\mu_j}{1+\max_{i\in E_r}\mu_i}$, *then* $e_n(i,j)\overset{d}{=}\varepsilon_n(i,j)-\eta_n(i,j)$, *where* $\varepsilon_n(i,j)$ *and* $\eta_n(i,j)$ *are two i.i.d. random variables with the distribution given as*

$$\begin{cases} Geom\left(\dfrac{\mu_j}{1+\mu_j}\right), & w.p.\ \left(1-\dfrac{\alpha_j\mu_i}{\mu_j-\alpha_j}\right), \\ Geom\left(\dfrac{\alpha_j}{1+\alpha_j}\right), & w.p.\ \dfrac{\alpha_j\mu_i}{\mu_j-\alpha_j}. \end{cases} \tag{2.5}$$

Presenting the distribution of the innovation time series $\{e_n(i,j)\},\ n\in\mathbb{N}$, in this shape may simplify the calculation of many properties of the process itself, as can be seen in the following corollary.

Corollary 2.2. *Let us suppose that* $z_n=j$ *and* $z_{n-1}=i$, *for some* i *and* $j\in E_r$. *Then, we have:*

$$\begin{aligned} E(e_n(i,j)) &= 0, \\ \mathrm{Var}(e_n(i,j)) &= 2(\mu_j(1+\mu_j)-\alpha_j\mu_i(1+2\alpha_j+\alpha_j\mu_i)). \end{aligned}$$

Proof: The proof of the first equality is trivial, given that $\varepsilon_n(i,j)$ and $\eta_n(i,j)$ have the same distribution.
Bearing in mind the shape of the distribution of $\varepsilon_n(i,j)$ and $\eta_n(i,j)$ and using properties of the probability generating function (p.g.f.), it is easy to prove that

$$\begin{aligned} \mathrm{Var}(\eta_n(i,j))=\mathrm{Var}(\varepsilon_n(i,j)) &= \Phi''_{\varepsilon_n(i,j)}(1)+\Phi'_{\varepsilon_n(i,j)}(1)-\left[\Phi'_{\varepsilon_n(i,j)}(1)\right]^2 \\ &= \mu_j(1+\mu_j)-\alpha_j\mu_i(1+2\alpha_j+\alpha_j\mu_i). \end{aligned}$$

Now, it is obvious that

$$\mathrm{Var}(e_n(i,j))=\mathrm{Var}(\varepsilon_n(i,j))+\mathrm{Var}(\eta_n(i,j))=2(\mu_j(1+\mu_j)-\alpha_j\mu_i(1+2\alpha_j+\alpha_j\mu_i)).$$

□

Remark 2.1. Let us highlight here two interesting facts:

- For $z_n=j$ and $z_{n-1}=i$, $\varepsilon_n(i,j)$ and $\eta_n(i,j)$ have the same distribution as an innovation process given in RrNGINAR$(\mathcal{M},\mathcal{A},\mathcal{P})$ model (see [11]);
- For $j=i$, the distribution of the innovation process $\{e_n(i,j)\}$ coincides with the distribution of the innovation process of the DLINAR(1) model (see [17]).

3. PROPERTIES OF THE PROCESS

In this section, the most important properties of the $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ model will be derived and analyzed. It is interesting to notice that many properties can be derived by observing $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ process as a difference of two independent $\mathrm{RrNGINAR}(\mathcal{M},\mathcal{A},\mathcal{P})$ processes, in case of $\mathcal{P}=\{1\}$.

To that purpose, for given sets $\mathcal{M},\mathcal{A}$ and $\mathcal{P}=\{1\}$, let us define two $\mathrm{RrNGINAR}(\mathcal{M},\mathcal{A},\mathcal{P})$ time series

$$\begin{aligned} X_n^{(1)}(z_n) &= \alpha_{z_n} * X_{n-1}^{(1)}(z_{n-1}) + \varepsilon_n(z_{n-1},z_n), \quad n\geq 1,\\ X_n^{(2)}(z_n) &= \alpha_{z_n} * X_{n-1}^{(2)}(z_{n-1}) + \eta_n(z_{n-1},z_n), \quad n\geq 1. \end{aligned}$$

Time series $\{X_n^{(1)}(z_n)\}$ and $\{X_n^{(2)}(z_n)\}$ are mutually independent and, for fixed $z_n=j$, $X_n^{(1)}(j)$ and $X_n^{(2)}(j)$ have the same $\mathrm{Geom}\Big(\frac{\mu_j}{1+\mu_j}\Big)$, $\mu_j\in\mathcal{M}$, distribution. Also, for fixed values $z_n=j$ and $z_{n-1}=i$, $\{\varepsilon_n(i,j)\}$ and $\{\eta_n(i,j)\}$ are two mutually independent time series with the same marginal distribution given in Corollary 2.1. Based on the definition of the $\mathrm{RrNGINAR}(\mathcal{M},\mathcal{A},\mathcal{P})$ process, $X_{n-l}^{(1)}(k)$ and $\varepsilon_n(i,j)$, as well as $X_{n-l}^{(2)}(k)$ and $\eta_n(i,j)$, are mutually independent for all $l\geq 1$ and for all $i,j,i\in E_r$.

Let $Y_n(z_n)$ be a $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ process with $\mathrm{DL}\Big(\frac{\mu_j}{1+\mu_j}\Big)$ marginals, given $z_n=j$. Now, using Corollary 2.1 and Corollary 2.1 from [17], we have

$$\begin{aligned} X_n^{(1)}(z_n) - X_n^{(2)}(z_n) &= \Big(\alpha_{z_n} * X_{n-1}^{(1)}(z_{n-1}) - \alpha_{z_n} * X_{n-1}^{(2)}(z_{n-1})\Big)\\ &\quad + \big(\varepsilon_n(z_{n-1},z_n) - \eta_n(z_{n-1},z_n)\big)\\ &\overset{d}{=} \alpha_{z_n}\odot Y_{n-1}(z_{n-1}) + e_n(z_{n-1},z_n) = Y_n(z_n). \end{aligned} \tag{3.1}$$

Now, it is easy to prove that $E(Y_n(z_n))=0$ and

$$\mathrm{Var}(Y_n(z_n)) = 2\,\mathrm{Var}\Big(X_n^{(1)}(z_n)\Big) = 2\mu_{z_n}(1+\mu_{z_n}).$$

One important property holds for $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ process. Namely, according to the Theorem 2.4 given in [17],

$$\alpha\odot Y \overset{d}{=} \mathrm{sgn}(Y)(\alpha * |Y|) + \sum_{j=1}^{\min\left(X^{(1)},X^{(2)}\right)} D_j,$$

whereby the following conditions are satisfied:

a) $Y\sim \mathrm{DL}\Big(\frac{\mu}{1+\mu}\Big)$, $X^{(1)}\sim \mathrm{Geom}\Big(\frac{\mu}{1+\mu}\Big)$, $X^{(2)}\sim \mathrm{Geom}\Big(\frac{\mu}{1+\mu}\Big)$;

b) $D_j\sim \mathrm{DL}\Big(\frac{\alpha}{1+\alpha}\Big)$;

c) random variables $Y, X^{(1)}, X^{(2)}, D_j$, $j\geq 1$, and random variables involved in "$\alpha*$" are independent.

For $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ process, the following result holds.

Theorem 3.1. *The RrDLINAR$_1(\mathcal{M},\mathcal{A})$ process $\{Y_n(z_n)\}$ is a Markov process.*

Proof: Let us define sets A and B as $A=\{Y_s(z_s)=y_s,\ s=0,1,...,n-2\}$ and $B=A\cup\{Y_{n-1}(z_{n-1})=y_{n-1}\}$. According to the property of "$\alpha\odot$" mentioned above, it holds

$$\alpha_{z_n}\odot Y_{n-1}(z_{n-1}) = \operatorname{sgn}(Y_{n-1}(z_{n-1}))(\alpha_{z_n}*|Y_{n-1}(z_{n-1})|) + \sum_{j=1}^{\min\left(X_{n-1}^{(1)}(z_{n-1}),X_{n-1}^{(2)}(z_{n-1})\right)} D_j(z_n),$$

whereby $X_{n-1}^{(1)}(z_{n-1})$ and $X_{n-1}^{(2)}(z_{n-1})$ have the same $\text{Geom}\left(\frac{\mu_{z_{n-1}}}{1+\mu_{z_{n-1}}}\right)$ distribution and $D_j(z_n)$ has the $\text{DL}\left(\frac{\alpha_{z_n}}{1+\alpha_{z_n}}\right)$ distribution. Now, we have

$$P(Y_n(z_n)=y_n|B) = P\Bigg(\operatorname{sgn}(Y_{n-1}(z_{n-1}))(\alpha_{z_n}*|Y_{n-1}(z_{n-1})|) + \sum_{j=1}^{\min\left(X_{n-1}^{(1)}(z_{n-1}),X_{n-1}^{(2)}(z_{n-1})\right)} D_j(z_n)+e_n(z_{n-1},z_n)=y_n\Big|B\Bigg).$$

Bearing in mind property (c) mentioned above and condition (3) of the Definition 2.1, it becomes obvious that

$$\begin{aligned}P(Y_n(z_n)=y_n|B) &= \sum_{j=-\infty}^{+\infty} P(\operatorname{sgn}(Y_{n-1}(z_{n-1}))(\alpha_{z_n}*|Y_{n-1}(z_{n-1})|)=j|B)\\ &\quad\times P\Bigg(\sum_{j=1}^{\min\left(X_{n-1}^{(1)}(z_{n-1}),X_{n-1}^{(2)}(z_{n-1})\right)} D_j(z_n)+e_n(z_{n-1},z_n)=y_n-j\Bigg)\\ &= \sum_{j=-\infty}^{+\infty}\binom{|y_{n-1}|+|j|-1}{|j|}\frac{\alpha_{z_n}^{|j|}}{(1+\alpha_{z_n})^{|y_{n-1}|+|j|}}\\ &\quad\times P\Bigg(\sum_{j=1}^{\min\left(X_{n-1}^{(1)}(z_{n-1}),X_{n-1}^{(2)}(z_{n-1})\right)} D_j(z_n)+e_n(z_{n-1},z_n)=y_n-j\Bigg).\end{aligned}$$

As the last expression depends only on y_{n-1}, it is obvious that RrDLINAR$_1(\mathcal{M},\mathcal{A})$ is a Markov process. □

3.1. The k-step ahead conditional expectation

Theorem 3.2. *Let $\{Y_n(z_n)\}$ be a RrDLINAR$_1(\mathcal{M},\mathcal{A})$ process. Then for $k\geq 1$,*

$$E(Y_{n+k}(z_{n+k})|Y_n(z_n)) = \left(\prod_{j=1}^{k}\alpha_{z_{n+j}}\right)Y_n(z_n). \tag{3.2}$$

Proof: The proof will be derived by induction. Let $k = 1$. Using Theorem 2.3 from [17], we have

$$\begin{aligned} E(Y_{n+1}(z_{n+1})|Y_n(z_n)) &= E(\alpha_{z_{n+1}} \odot Y_n(z_n)|Y_n(z_n)) + E(e_{n+1}(z_n, z_{n+1})) \\ &= \alpha_{z_{n+1}} Y_n(z_n). \end{aligned} \tag{3.3}$$

Suppose the equality (3.2) holds for $k < m$. Bearing in mind the Markov property of the RrDLINAR$_1(\mathcal{M}, \mathcal{A})$ process, we will prove that (3.2) holds for $k = m$ as well. Namely,

$$\begin{aligned} E(Y_{n+m}(z_{n+m})|Y_n(z_n)) &= E\big[E(Y_{n+m}(z_{n+m})|Y_{n+m-1}(z_{n+m-1}), ..., Y_n(z_n))|Y_n(z_n)\big] \\ &= E\big[E(Y_{n+m}(z_{n+m})|Y_{n+m-1}(z_{n+m-1}))|Y_n(z_n)\big] \\ &= E(\alpha_{z_{n+m}} Y_{n+m-1}(z_{n+m-1})|Y_n(z_n)) \\ &= \alpha_{z_{n+m}} \left(\prod_{j=1}^{m-1} \alpha_{z_{n+j}} \right) Y_n(z_n) \\ &= \left(\prod_{j=1}^{m} \alpha_{z_{n+j}} \right) Y_n(z_n). \end{aligned}$$

This completes the proof of this theorem. □

3.2. Correlation structure

Theorem 3.3. *A RrDLINAR$_1(\mathcal{M}, \mathcal{A})$ process $\{Y_n(z_n)\}$ given by (2.3) is the correlated process with*

$$\operatorname{Corr}(Y_n(z_n), Y_{n-k}(z_{n-k})) = \begin{cases} \left(\prod_{j=0}^{k-1} \alpha_{z_{n-j}} \right) \sqrt{\dfrac{\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}{\mu_{z_n}(1+\mu_{z_n})}}, & k \geq 0, \\ \left(\prod_{j=1}^{-k} \alpha_{z_{n+j}} \right) \sqrt{\dfrac{\mu_{z_n}(1+\mu_{z_n})}{\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}}, & k < 0. \end{cases} \tag{3.4}$$

Proof: Since $\{Y_n(z_n)\}$ is a process with the k-step ahead conditional expectation of the form $E(Y_{n+k}(z_{n+k})|Y_n(z_n)) = \left(\prod_{j=1}^{k} \alpha_{z_{n+j}}\right) Y_n(z_n)$, unconditional expectation $E(Y_n(z_n)) = 0$ and finite variance $\operatorname{Var}(Y_n(z_n)) = 2\mu_{z_n}(1 + \mu_{z_n})$, for $k \geq 0$ it becomes easy to obtain

$$\begin{aligned} \operatorname{Cov}(Y_n(z_n), Y_{n-k}(z_{n-k})) &= \operatorname{Cov}(E(Y_n(z_n)|Y_{n-k}(z_{n-k})), Y_{n-k}(z_{n-k})) \\ &= \operatorname{Cov}\left(\left(\prod_{j=0}^{k-1} \alpha_{z_{n-j}} \right) Y_{n-k}(z_{n-k}), Y_{n-k}(z_{n-k}) \right) \\ &= \left(\prod_{j=0}^{k-1} \alpha_{z_{n-j}} \right) \operatorname{Var}(Y_{n-k}(z_{n-k})) \\ &= 2 \left(\prod_{j=0}^{k-1} \alpha_{z_{n-j}} \right) \mu_{z_{n-k}}(1 + \mu_{z_{n-k}}), \end{aligned}$$

whence we have

$$\begin{aligned}\mathrm{Corr}(Y_n(z_n), Y_{n-k}(z_{n-k})) &= \frac{2\left(\prod_{j=0}^{k-1}\alpha_{z_{n-j}}\right)\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}{\sqrt{2\mu_{z_n}(1+\mu_{z_n})2\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}} \\ &= \left(\prod_{j=0}^{k-1}\alpha_{z_{n-j}}\right)\sqrt{\frac{\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}{\mu_{z_n}(1+\mu_{z_n})}}.\end{aligned}$$

Similar to this, for $k<0$ we obtain

$$\begin{aligned}\mathrm{Cov}(Y_n(z_n), Y_{n-k}(z_{n-k})) &= E(Y_n(z_n)\cdot Y_{n-k}(z_{n-k})) \\ &= E[E(Y_n(z_n)Y_{n-k}(z_{n-k})|Y_n(z_n))] \\ &= E\left(Y_n(z_n)\left(\prod_{j=1}^{-k}\alpha_{z_{n+j}}\right)Y_n(z_n)\right) \\ &= \left(\prod_{j=1}^{-k}\alpha_{z_{n+j}}\right)\mathrm{Var}(Y_n(z_n)) \\ &= \left(\prod_{j=1}^{-k}\alpha_{z_{n+j}}\right)2\mu_{z_n}(1+\mu_{z_n}),\end{aligned}$$

whence we have

$$\mathrm{Corr}(Y_n(z_n), Y_{n-k}(z_{n-k})) = \left(\prod_{j=1}^{-k}\alpha_{z_{n+j}}\right)\sqrt{\frac{\mu_{z_n}(1+\mu_{z_n})}{\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}}.$$

□

Remark 3.1. If $z_n = z_{n-1} = \cdots = z_{n-k} = j$, then it holds that $\mathrm{Corr}(Y_n(z_n), Y_{n-k}(z_{n-k})) = \alpha_j^{|k|}$, which matches with correlation function of the DLINAR(1) process.

Bearing in mind the equality (3.4) and the facts that $\mu_{z_{n-k}} > 0$, $\mu_{z_n} > 0$ and $\alpha_{z_{n-j}} > 0$ for all $j = 0, 1, ..., k-1$, it is obvious that in case of $k \geq 0$, $\mathrm{Corr}(Y_n(z_n), Y_{n-k}(z_{n-k})) > 0$. Let us prove now the validity of the relation $\mathrm{Corr}(Y_n(z_n), Y_{n-k}(z_{n-k})) < 1$.

For all $j = 0, 1, ..., k-1$, $\alpha_{z_{n-j}} \leq \frac{\mu_{z_{n-j}}}{1+\max_{i\in E_r}\mu_{z_i}}$, so, obviously

$$\alpha_{z_{n-j}} \leq \frac{\mu_{z_{n-j}}}{1+\mu_{z_{n-j-1}}} < \frac{\mu_{z_{n-j}}}{\mu_{z_{n-j-1}}} < \frac{1+\mu_{z_{n-j}}}{\mu_{z_{n-j-1}}}.$$

Then,

$$\alpha_{z_{n-j}}^2 < \frac{\mu_{z_{n-j}}}{1+\mu_{z_{n-j-1}}}\cdot\frac{1+\mu_{z_{n-j}}}{\mu_{z_{n-j-1}}},$$

so we can conclude that $\alpha_{z_{n-j}} < \sqrt{\frac{\mu_{z_{n-j}}(1+\mu_{z_{n-j}})}{\mu_{z_{n-j-1}}(1+\mu_{z_{n-j-1}})}}$, and further, that

$$\begin{aligned}\prod_{j=0}^{k-1}\alpha_{z_{n-j}} &< \sqrt{\frac{\mu_{z_n}(1+\mu_{z_n})}{\mu_{z_{n-1}}(1+\mu_{z_{n-1}})}}\sqrt{\frac{\mu_{z_{n-1}}(1+\mu_{z_{n-1}})}{\mu_{z_{n-2}}(1+\mu_{z_{n-2}})}}\cdots\sqrt{\frac{\mu_{z_{n-k+1}}(1+\mu_{z_{n-k+1}})}{\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}} \\ &= \sqrt{\frac{\mu_{z_n}(1+\mu_{z_n})}{\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}}.\end{aligned}$$

Finally, it holds

$$\mathrm{Corr}(Y_n(z_n), Y_{n-k}(z_{n-k})) < \sqrt{\frac{\mu_{z_n}(1+\mu_{z_n})}{\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}}\sqrt{\frac{\mu_{z_{n-k}}(1+\mu_{z_{n-k}})}{\mu_{z_n}(1+\mu_{z_n})}} = 1.$$

Similarly, it can be shown that $0 < \mathrm{Corr}(Y_n(z_n), Y_{n-k}(z_{n-k})) < 1$, for $k < 0$.

4. YULE-WALKER ESTIMATION

In this section, the YW estimators will be provided and their strong consistency will be proven. To that purpose, we will use procedure similar to the one described in [18].

Thus, let $Y_1(z_1), Y_2(z_2), ..., Y_N(z_N)$ be a sample of the RrDLINAR$_1(\mathcal{M}, \mathcal{A})$ process $\{Y_n(z_n)\}$ of size N. The main idea of the procedure described in [18] is to estimate parameters μ_k and α_k only by using elements corresponding to the state k. Thus, let us divide the initial sample into r subsamples S_k, $k = 1, 2, ..., r$, where S_k is a subsample which contains all the elements corresponding to the state k and doesn't contain elements corresponding to any other state. This division can be performed in the following way:

$$\begin{aligned} I_k &= \{i \in \{1, 2, ..., N\} | z_i = k\}, \quad k \in \{1, 2, ..., r\}, \\ \bigcup_{k=1}^{r} I_k &= \{1, 2, ..., N\}, \quad |I_k| = n_k, \quad n_1 + n_2 + \cdots + n_r = N, \\ S_k &= \{Y_{k_1}(k), Y_{k_2}(k), ..., Y_{k_{n_k}}(k)\}, \quad k_j \in I_k, \quad k_j < k_{j+1}, \quad \forall j \in \{1, 2, ..., n_k - 1\}. \end{aligned}$$

In more detail, each S_k, $k = 1, 2, ..., r$, represents a disjoint union of subsamples $S_{k,1}, S_{k,2}, ..., S_{k,i_k}$, which we call 'maximal' subsamples. For an arbitrary subsample $S_{k,l}$, $l = 1, 2, ..., i_k$, we can find natural numbers m_l and n_l, $m_l < n_l$, such that $z_{m_l} \neq k$, $z_{m_l+1} = z_{m_l+2} = \cdots = z_{n_l} = k$, $z_{n_l+1} \neq k$. In that case, the subsample $S_{k,l} = \{Y_{m_l+1}(z_{m_l+1}), Y_{m_l+2}(z_{m_l+2}), ..., Y_{n_l}(z_{n_l})\}$ corresponds to the state k and is maximal in the sense that it cannot be expanded neither to the left nor right side in the way that all of its elements correspond to the state k. Now, each of those maximal subsamples $S_{k,l}$, $l = 1, 2, ..., i_k$ may be observed as a sample of some DLINAR(1) process with the marginal distribution parameter μ_k. Let us introduce the following notation: $J_{k,l} = \{i \in \{1, 2, ..., N\} | Y_i(z_i) \in S_{k,l}\}$, $|J_{k,l}| = n_{k,l}$ for all $l = 1, 2, ..., i_k$ and $n_{k,1} + n_{k,2} + \cdots + n_{k,i_k} = n_k$. As shown in [17], the DLINAR(1) process is stationary and ergodic, and the corresponding sample variance and the first-order sample covariance are strongly consistent estimates of the variance and the first-order covariance of the process. Finally, in case of subsample $S_{k,l}$, these estimators are of the form

$$\widehat{\gamma}_{0,l}^{(k)} = \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} Y_i^2(k) \quad \text{and} \quad \widehat{\gamma}_{1,l}^{(k)} = \frac{1}{n_{k,l}} \sum_{i, i+1 \in J_{k,l}} Y_i(k) Y_{i+1}(k).$$

Let us define now the corresponding estimators without taking maximal subsamples into account.

Definition 4.1. Estimators obtained from the subsample S_k corresponding to the state k are defined as

$$\widehat{\gamma}_0^{(k)} = \frac{1}{n_k} \sum_{i \in I_k} Y_i^2(k), \quad \widehat{\gamma}_1^{(k)} = \frac{1}{n_k} \sum_{i, i+1 \in I_k} Y_i(k) Y_{i+1}(k). \tag{4.1}$$

Theorem 4.1. *Estimators* $\widehat{\gamma}_0^{(k)}$ *and* $\widehat{\gamma}_1^{(k)}$ *from Definition 4.1 are strongly consistent.*

Proof: This theorem shall be proven in a similar way as it was done by [18]. First of all, the strong consistency property for $\widehat{\gamma}_0^{(k)}$ shall be proven. Because $\widehat{\gamma}_{0,l}^{(k)}$ is strongly consistent for all $l \in \{1, 2, ..., i_k\}$ it holds that $\widehat{\gamma}_{0,l}^{(k)} \to \gamma_0^k$, $n_{k,l} \to \infty$ everywhere except on the set $\Omega_{k,l}$, where $P(\Omega_{k,l}) = 0$. Now, it holds that

$$\widehat{\gamma}_0^{(k)} = \frac{1}{n_k} \sum_{i \in I_k} Y_i^2(k) = \frac{1}{n_k} \sum_{l=1}^{i_k} \sum_{i \in J_{k,l}} Y_i^2(k) = \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} \frac{1}{n_{k,l}} \sum_{i \in J_{k,l}} Y_i^2(k) = \sum_{l=1}^{i_k} \frac{n_{k,l}}{n_k} \widehat{\gamma}_{0,l}^{(k)}.$$

Let $n_k \to \infty$. Following the technique introduced by [18], it is easy to show that

$$\lim_{n_k \to \infty} \widehat{\gamma}_0^{(k)} = \lim_{n_k \to \infty} \sum_{l=1}^{d} \frac{n_{k,l}}{n_k} \widehat{\gamma}_{0,l}^{(k)},$$

where $n_{k,l}$, $l = 1, 2, ..., d$, represent those maximal sample sizes which approach infinity when n_k does so. Thus, we have

$$\begin{aligned} \lim_{n_k \to \infty} \widehat{\gamma}_0^{(k)} &= \lim_{n_{k,l} \to \infty,\ \forall l \in \{1,2,...,d\}} \sum_{l=1}^{d} \frac{n_{k,l}}{n_k} \widehat{\gamma}_{0,l}^{(k)} \\ &= \gamma_0^{(k)} \lim_{n_{k,l} \to \infty,\ \forall l \in \{1,2,...,d\}} \sum_{l=1}^{d} \frac{n_{k,l}}{n_k} = \gamma_0^{(k)}. \end{aligned} \tag{4.2}$$

We mentioned earlier that $\lim_{n_{k,l} \to \infty} \widehat{\gamma}_{0,l}^{(k)} = \gamma_0^{(k)}$ everywhere except on the set $\Omega_{k,l}$, where $P(\Omega_{k,l}) = 0$. Thus, the equality (4.2) holds everywhere except on the set $\Omega_k = \bigcup_{l=1}^{d} \Omega_{k,l}$, where

$$P(\Omega_k) = P\left(\bigcup_{l=1}^{d} \Omega_{k,l} \right) \leq \sum_{l=1}^{d} P(\Omega_{k,l}) = 0.$$

From the non-negativity of the probability, it follows that $P(\Omega_k) = 0$. Hence, $\widehat{\gamma}_0^{(k)}$ is a strongly consistent estimator of the variance $\gamma_0^{(k)}$.

Proof for $\widehat{\gamma}_1^{(k)}$ is analogous to the one proposed for $\widehat{\gamma}_0^{(k)}$. □

It has remained to estimate parameters μ_k and α_k on the subsample S_k. According to Theorem 3.3 from [17], it holds

$$\gamma_0^{(k)} = 2\mu_k(1 + \mu_k), \quad \gamma_1^{(k)} = 2\alpha_k\mu_k(1 + \mu_k).$$

It follows that

$$\widehat{\mu}_k^{YW} = -\frac{1}{2} + \frac{1}{2}\sqrt{1 + 2\widehat{\gamma}_0^{(k)}}, \quad \widehat{\alpha}_k^{YW} = \frac{\widehat{\gamma}_1^{(k)}}{\widehat{\gamma}_0^{(k)}}.$$

Now, it is obvious that $f(x) = -\frac{1}{2} + \frac{1}{2}\sqrt{1 + 2x}$, $x \geq 0$ is a continuous function. According to Proposition 6.3.4 provided by [5], we conclude that $\widehat{\mu}_k^{YW}$ is strongly consistent. The strong consistency of $\widehat{\alpha}_k^{YW}$ follows from the Theorem 4.1 given in [8].

5. MODEL SIMULATIONS

We now focus on the YW estimating procedures on simulated data series, in order to clarify the utility of the observed non-stationary modeling and to justify the quality of the presented estimation method. Both processes, the random environment process $\{Z_n\}$ and the $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ process $\{Y_n(z_n)\}$ have been simultaneously simulated in 100 replicates, each of size 10000. Sequences $\{Y_n(z_n)\}$ are simulated using the fact that the newly defined process is distributed the same as a difference between two independent $\mathrm{RrNGINAR}(\mathcal{M},\mathcal{A},\mathcal{P})$ processes $\{X_n^{(1)}(z_n)\}$ and $\{X_n^{(2)}(z_n)\}$, with $\mathcal{P}=\{1\}$. Thus, first we simulated $\{X_n^{(1)}(z_n)\}$, and independently of it, $\{X_n^{(2)}(z_n)\}$ and derive $\{Y_n(z_n)\}$ as $Y_n(z_n)=X_n^{(1)}(z_n)-X_n^{(2)}(z_n),\ n\geq 1$. Considering the number of possible random states, we have decided to discuss, in our opinion, two of the most plausible cases in practice, as follows.

5.1. The case of two states

Here we have assumed that the random environment process is performed in two different states. Also, we have considered two different combinations of the model parameter values. Bearing in mind that $\alpha_j \in \left(0, \frac{\mu_j}{1+\max_{i\in E_r}\mu_i}\right]$, first of all we have analyzed the case when parameters $\alpha_j,\ j=1,2$, were close to their upper limits. So, we have used the following true values: $\mu=(1,3)$ and $\alpha=(0.25,0.7)$. It has been assumed that the choice of the initial random state is fair, which led us to have $p_{vec}=(0.5,0.5)$. It has remained for us to set the random environment process transition probability matrix. In order to preserve the simulated $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ process in one state as long as possible, we have chosen to prefer the present state of the random environment process, i.e. probabilities that the random environment process changes its state are significantly smaller. Thus, $p_{mat}=\begin{bmatrix}0.6 & 0.4\\ 0.2 & 0.8\end{bmatrix}$.

In the second case, we have chosen parameters $\alpha_j,\ j=1,2$, to be smaller than the midpoints of intervals $\left(0, \frac{\mu_j}{1+\max_{i\in E_r}\mu_i}\right],\ j=1,2$. In order not to shrink these intervals too much, parameters $\mu_j,\ j=1,2$, must have relatively close values. In that manner, we have chosen the following true parameter values: $\mu=(2,3)$ and $\alpha=(0.2,0.3)$. An initial state is nearly fair, due to the value of its distribution $p_{vec}=(0.45,0.55)$, and the random states transition probabilities are given as $p_{mat}=\begin{bmatrix}0.7 & 0.3\\ 0.3 & 0.7\end{bmatrix}$.

5.2. The case of three states

Similarly, as aforementioned, in the case of three different random states we have also considered two different combinations of the true parameter values. And yet again, the case when parameters α_j, $j=1,2,3$, approach their upper limits has been analyzed in the first place. The following true parameter values have been used: $\mu=(1,2,5)$ and $\alpha=(0.1,0.25,0.7)$.

An initial state is nearly fair, due to the value of its distribution $p_{vec} = (0.3, 0.4, 0.3)$, and the random environment process transition probability matrix is $p_{mat} = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.2 & 0.1 & 0.7 \end{bmatrix}$.

At the very end, the second case represents the simulation when parameters α_j, $j = 1, 2, 3$, are smaller than the midpoints of corresponding intervals $\left(0, \frac{\mu_j}{1+\max_{i \in E_r} \mu_i}\right]$, $j = 1, 2, 3$. Thus, we have chosen the following true parameter values: $\mu = (2, 3, 5)$ and $\alpha = (0.1, 0.2, 0.4)$. An initial state is fair, due to the value of its distribution $p_{vec} = (0.33, 0.34, 0.33)$ and the random environment process transition probability matrix is $p_{mat} = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$.

5.3. Estimation results

For both presented cases, $r = 2$ and $r = 3$, we estimate parameters μ_j and α_j, $j = 1, 2, ..., r$, of the RrDLINAR$_1(\mathcal{M}, \mathcal{A})$ model using the YW method. Compared to the DLINAR(1) model, the newly defined model has a greater number of unknown parameters, which leads to better fitting to the data, because it is more "flexible". However, since each state has its own parameters that can be estimated only based on the part of the sample corresponding to that state, it is expected to need a bigger sample for the same precision of the estimation.

Table 1: The case of 2 states.

a) True values $\mu = (1, 3)$, $\alpha = (0.25, 0.7)$, $p_{vec} = (0.5, 0.5)$, $p_{mat} = \begin{bmatrix} 0.6 & 0.4 \\ 0.2 & 0.8 \end{bmatrix}$.

N_1	$\widehat{\mu}_1^{YW}$	$\widehat{\mu}_2^{YW}$	$\widehat{\alpha}_1^{YW}$	$\widehat{\alpha}_2^{YW}$
200	0.967	2.865	0.229	0.662
St. errors	(0.164)	(0.498)	(0.136)	(0.114)
500	0.980	2.943	0.236	0.686
St. errors	(0.113)	(0.350)	(0.092)	(0.068)
1000	0.984	3.003	0.241	0.700
St. errors	(0.082)	(0.238)	(0.066)	(0.052)
5000	0.997	3.002	0.249	0.700
St. errors	(0.038)	(0.117)	(0.033)	(0.023)
10000	0.990	2.998	0.250	0.700
St. errors	(0.025)	(0.087)	(0.024)	(0.015)

b) True values $\mu = (2, 3)$, $\alpha = (0.2, 0.3)$, $p_{vec} = (0.45, 0.55)$, $p_{mat} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$.

N_1	$\widehat{\mu}_1^{YW}$	$\widehat{\mu}_2^{YW}$	$\widehat{\alpha}_1^{YW}$	$\widehat{\alpha}_2^{YW}$
200	2.052	3.033	0.210	0.291
St. errors	(0.274)	(0.368)	(0.114)	(0.121)
500	2.026	2.992	0.197	0.294
St. errors	(0.166)	(0.256)	(0.080)	(0.073)
1000	2.024	2.992	0.202	0.298
St. errors	(0.108)	(0.209)	(0.057)	(0.056)
5000	2.000	3.002	0.198	0.298
St. errors	(0.055)	(0.082)	(0.027)	(0.023)
10000	2.000	3.002	0.198	0.298
St. errors	(0.041)	(0.056)	(0.017)	(0.017)

Thus, parameter estimates are derived for subsamples of sizes $200, 500, 1000, 5000$ and 10000. In each mentioned case, 100 simulated data series are used, and the corresponding standard errors are calculated. With the sample size increment, all estimates are convergent with the standard errors decreasing towards 0. It is also visible that the standard errors for μ_j are smaller for smaller values of μ_j. These results are presented in Table 1 and Table 2, corresponding to the cases of two and three random states, respectively. Transition probabilities are not estimated this time, since those are not parameters of the $\mathrm{RrDLINAR}_1(\mathcal{M}, \mathcal{A})$ model itself.

Table 2: The case of 3 states.

a) True values $\mu = (1, 2, 5)$, $\alpha = (0.1, 0.25, 0.7)$, $p_{vec} = (0.3, 0.4, 0.3)$, $p_{mat} = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.2 & 0.1 & 0.7 \end{bmatrix}$.

N_1	$\widehat{\mu}_1^{YW}$	$\widehat{\mu}_2^{YW}$	$\widehat{\mu}_3^{YW}$	$\widehat{\alpha}_1^{YW}$	$\widehat{\alpha}_2^{YW}$	$\widehat{\alpha}_3^{YW}$
200	1.030	2.103	4.957	0.115	0.263	0.695
St. errors	(0.202)	(0.407)	(0.945)	(0.130)	(0.130)	(0.141)
500	1.007	2.038	5.051	0.107	0.259	0.705
St. errors	(0.133)	(0.250)	(0.696)	(0.085)	(0.092)	(0.104)
1000	0.993	2.037	5.018	0.104	0.259	0.705
St. errors	(0.090)	(0.174)	(0.521)	(0.062)	(0.061)	(0.078)
5000	1.000	2.010	5.009	0.103	0.250	0.701
St. errors	(0.035)	(0.073)	(0.216)	(0.030)	(0.030)	(0.034)
10000	1.000	2.004	4.995	0.103	0.250	0.700
St. errors	(0.025)	(0.053)	(0.164)	(0.020)	(0.020)	(0.026)

b) True values $\mu = (2, 3, 5)$, $\alpha = (0.1, 0.2, 0.4)$, $p_{vec} = (0.33, 0.34, 0.33)$, $p_{mat} = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.1 & 0.7 & 0.2 \\ 0.1 & 0.1 & 0.8 \end{bmatrix}$.

N_1	$\widehat{\mu}_1^{YW}$	$\widehat{\mu}_2^{YW}$	$\widehat{\mu}_3^{YW}$	$\widehat{\alpha}_1^{YW}$	$\widehat{\alpha}_2^{YW}$	$\widehat{\alpha}_3^{YW}$
200	1.978	3.016	4.937	0.110	0.181	0.390
St. errors	(0.359)	(0.652)	(0.737)	(0.122)	(0.141)	(0.119)
500	2.016	3.016	4.946	0.103	0.207	0.390
St. errors	(0.216)	(0.366)	(0.471)	(0.083)	(0.121)	(0.083)
1000	1.991	2.994	4.954	0.103	0.201	0.393
St. errors	(0.157)	(0.248)	(0.312)	(0.060)	(0.080)	(0.054)
5000	1.991	3.006	4.989	0.099	0.199	0.396
St. errors	(0.078)	(0.120)	(0.146)	(0.028)	(0.036)	(0.027)
10000	1.991	3.006	4.991	0.099	0.201	0.399
St. errors	(0.048)	(0.087)	(0.112)	(0.018)	(0.023)	(0.018)

6. FORECASTING

Accuracy of forecasting in real-life data analysis is as important as evaluating the fit of the model. We introduce here a criterion suitable to compare the prediction results of different models. However, before we introduce the criterion itself, the forecasting procedure will be described in brief. At the beginning, the data sample of size $N = n_1 + n_2$ is divided into two sets, the training set and the prediction set. The training set contains the first n_1 sample elements and the prediction set contains the last n_2 elements of the sample.

We use training set to estimate model parameters and evaluate the fitting quality. In order to evaluate the forecasting accuracy, we generate m sequences of predictions from the estimated model parameters, each of size n_2. Using the proposed criterion, we compare generated predictions with the prediction set and finally determine the accuracy of forecasting.

We decided to use the forecasting log-score criterion (FLSC) already introduced in [14]. The criterion represents a modification of the log-score criterion (LSC) proposed by [10]. The FLSC criterion is given by the formula

$$\text{FLSC} = \sum_{h=1}^{n_2} \log \hat{p}_{n_1+h}(x_{n_1+h}),$$

where $h = 1, 2, ..., n_2$ and $\hat{p}_{n_1+h}(x_{n_1+h})$ represents the estimated probability of correctly predicting the value x_{n_1+h} from the prediction set, that is,

$$\hat{p}_{n_1+h}(x_{n_1+h}) = \frac{\text{number of correct predictions}}{m}.$$

Hihger values of the FLSC imply better forecasting.

7. APPLICATION

Regarding the application to the real-life data, we want to examine whether there is any progress compared to other INAR models that deal with both positive and negative values. For that purpose, we consider the number of motor vehicle thefts reported on a monthly basis to police stations number 1608 and 2811, in Pittsburgh, Pennsylvania, USA, between January 1990. and December 2001. The data were collected by the City of Pittsburgh Bureau of Police and reported under the FBI Uniform Crime Report. The differences between motor vehicle thefts on a monthly basis reported to these two police stations are calculated and provided in Table 3.

Table 3: Differences between motor vehicle thefts reported on a monthly basis to police stations number 1608 and 2811.

12	−1	2	3	8	−2	−3	4	4	6	5	5	5	4	4
5	4	5	4	0	1	0	1	2	3	−6	0	−1	−1	1
0	2	−1	0	1	−4	−5	−13	−4	−4	−5	−4	−6	−5	−8
−5	−5	−4	−4	−6	−5	0	1	−3	3	0	1	−2	0	0
−3	−1	−3	−3	−1	3	1	−1	0	0	−1	−1	2	1	1
1	3	0	2	1	0	0	2	1	1	−2	−2	−1	0	1
0	0	−3	0	1	−2	0	−2	2	−2	−3	2	2	2	3
2	1	−2	0	0	2	3	−3	0	−2	3	3	1	0	0
2	3	1	0	−3	−2	1	−3	−3	−3	2	3	−2	−2	1
3	1	2	0	3	2	3	2	−3						

Based on the sample size of $N = 144$, the sample mean of the differences between motor vehicle thefts in police stations number 1608 and 2811 is $\overline{y} = 0.048$, which proves the fact that the mean values of the number of motor vehicle thefts in both stations are approximately the same.

This condition is crucial here, since it had been claimed earlier that both processes $\{X_n^{(1)}(z_n)\}$ and $\{X_n^{(2)}(z_n)\}$ must have the same distribution.

The usual first step in standard INAR modeling is to obtain the plots of autocorrelation and partial autocorrelation functions. Those are given in Figure 1 and successfully justify the usage of the INAR(1) modeling. Bearing that in mind, we decide to compare our RrDLINAR$_1(\mathcal{M}, \mathcal{A})$ model on the training set with TINAR(1) (introduced by [9]), DLINAR(1) (proposed by [17]) and STINAR(1) (given in [4]). All mentioned models involve two i.i.d. latent AR components and are therefore suitable to compete with the RrDLINAR$_1(\mathcal{M}, \mathcal{A})$ model. As the criteria of the model validity, we take the following goodness-of-fit statistics: the root mean square error (RMSE), the mean absolute error (MAE) and the median absolute error (MdAE). Here, errors are defined as differences between observed values and corresponding predictions. In general, the model that provides better fitting to the data is expected to show lower values of these statistics.

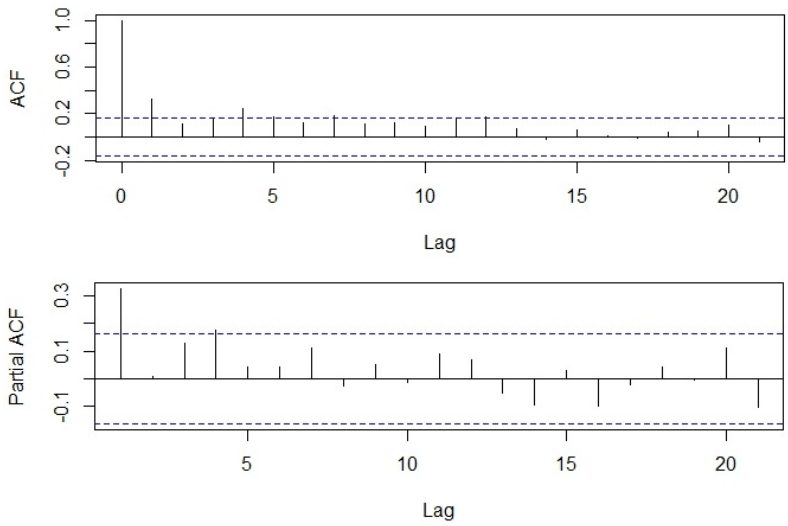

Figure 1: Autocorrelations and partial autocorrelations.

In the next step, the observed data values are clustered. This is how we actually obtain realized values of the corresponding random environment process. For all $n \in \{1, 2, ..., 144\}$, if the theft difference in the n-th month is in the i-th cluster (where $i \in E_r$ and $E_r = \{1, 2, ..., r\}$ is a set of possible random states), then $z_n = i$ is a n-th realized value of the corresponding r states random environment process. In that way, realizations z_n are determined. In this particular case, we decided to divide the theft difference realizations into two clusters, using the K-means clustering technique. These clustering results are given in Figure 2. Based on this figure, the decision to divide the theft difference realizations into two clusters is justified. As can be noticed, the differences that do not deviate too much from zero are located in the first cluster (triangles). The second cluster mainly consists of realizations from two time intervals (August 1990–July 1991 and December 1993–March 1995). These realizations deviate significantly from zero (circles), indicating that changes in environment state may have occurred. With each additional increase in the number of clusters, at least one cluster with very few realizations in it is created. This leads to frequent state changes, which is ruinous to any model in a random environment, including RrDLINAR$_1(\mathcal{M}, \mathcal{A})$.

Hence, $\mathrm{RrDLINAR}_1(\mathcal{M},\mathcal{A})$ models with more than 2 environment states should not be considered for this data series.

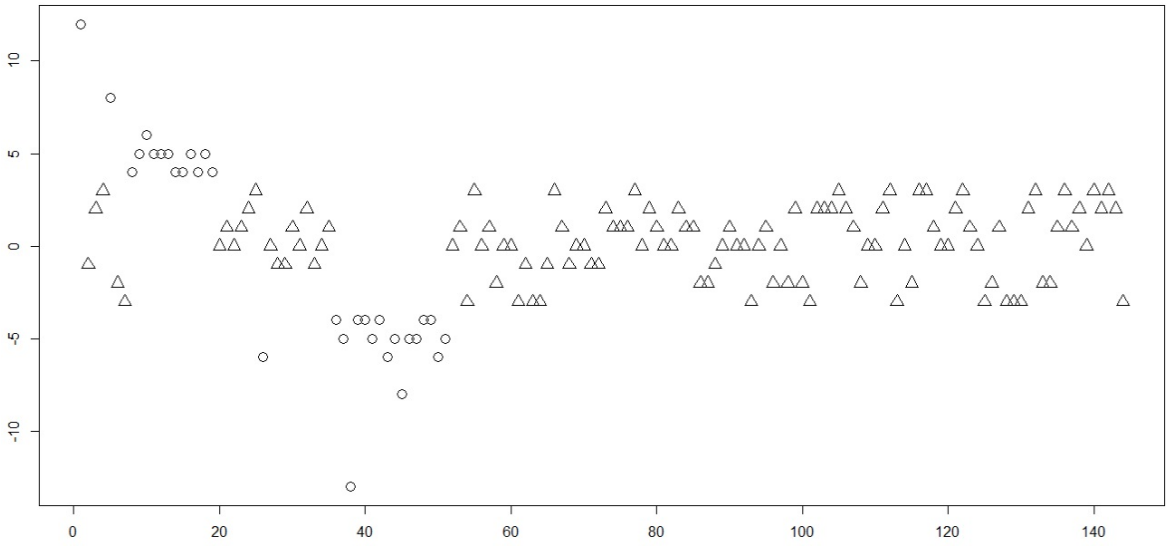

Figure 2: Clustering results obtained by applying the K-means technique to the motor vehicle theft differences.

The data set was divided into two parts, the training set with the first 120 sample elements and the prediction set with the last 24 sample elements. Based on the training set, we are able to construct the $\mathrm{R2DLINAR}_1(\mathcal{M},\mathcal{A})$ process $\{Y_n(z_n)\}$, of which arbitrary element $Y_n(z_n)$ has the $\mathrm{DL}\left(\frac{\widehat{\mu}_{z_n}}{1+\widehat{\mu}_{z_n}}\right)$ distribution, where $\widehat{\mu}_1,\widehat{\mu}_2$ and $\widehat{\alpha}_1,\widehat{\alpha}_2$ are previously obtained from the training set using corresponding YW estimates.

Furthermore, we perform on training set the fitting quality comparison of all applied models by calculating RMSE, MAE and MdAE for each particular model. These values, together with the Yule-Walker parameter estimates, are presented in Table 4. Obviously, the $\mathrm{R2DLINAR}_1(\mathcal{M},\mathcal{A})$ model based on two states random environment process showed the best performance when fitting this kind of real-life data sequences, providing the smallest values of all goodness-of-fit statistics. Regarding other competitors, STINAR(1) performed slightly better. However, even for this model, goodness-of-fit statistics are significantly higher then corresponding counterparts calculated for the $\mathrm{R2DLINAR}_1(\mathcal{M},\mathcal{A})$ model. One more fact is noticeable. Compared to the parameter α_2, parameter α_1 from $\mathrm{R2DLINAR}_1(\mathcal{M},\mathcal{A})$ model is more similar to the parameter α from the other three models. Bearing in mind the form of the one-step ahead conditional expectation given with (3.3), it can be concluded that all proposed models are suppose to give relatively similar approximations of data points from the first cluster. However, in the second cluster, the approximations should differ dramatically, which would explain the difference in RMSE, MAE and MdAE values.

In Figure 3, the realization of the theft difference process, as well as the predicted values of $\mathrm{R2DLINAR}_1(\mathcal{M},\mathcal{A})$ and STINAR(1) models, are shown. The TINAR(1) and DLINAR(1) model predictions are omitted here due to their similarity to STINAR(1) predictions.

Table 4: YW parameter estimates and goodness-of-fit statistics RMSE, MAE and MdAE for INAR(1) modeling of the theft difference.

Model	YW	RMSE	MAE	MdAE
TINAR(1)	$\widehat{\alpha} = 0.334$ $\widehat{\lambda} = 1.818$	2.691	2.139	2.010
DLINAR(1)	$\widehat{\alpha} = 0.332$ $\widehat{\mu} = 1.882$	2.690	2.135	2.001
STINAR(1)	$\widehat{\alpha} = 0.338$ $\widehat{\mu}_1 = 2.008$ $\widehat{\mu}_2 = 1.992$	2.681	2.131	1.993
R2DLINAR$_1(\mathcal{M}, \mathcal{A})$	$\widehat{\alpha}_1 = 0.190$ $\widehat{\alpha}_2 = 0.809$ $\widehat{\mu}_1 = 0.816$ $\widehat{\mu}_2 = 3.649$	2.188	1.665	1.001

As expected, both models (R2DLINAR$_1(\mathcal{M}, \mathcal{A})$ and STINAR(1)) approximate well the values which are not that far from zero. A large difference in quality fitting is noticeable in realizations that correspond to another state, i.e., that deviate significantly from zero. In this case, the newly defined R2DLINAR$_1(\mathcal{M}, \mathcal{A})$ model shows much better ability to adjust to the real-life realizations, which leads to better fitting. For high deviations, the difference between the predicted values of the R2DLINAR$_1(\mathcal{M}, \mathcal{A})$ and STINAR(1) models is larger. This adaptability is certainly a consequence of the non-stationary nature of the RrDLINAR$_1(\mathcal{M}, \mathcal{A})$. It is important to add that the trajectory of the R2DLINAR$_1(\mathcal{M}, \mathcal{A})$ model generally follows the trajectory of the realized process.

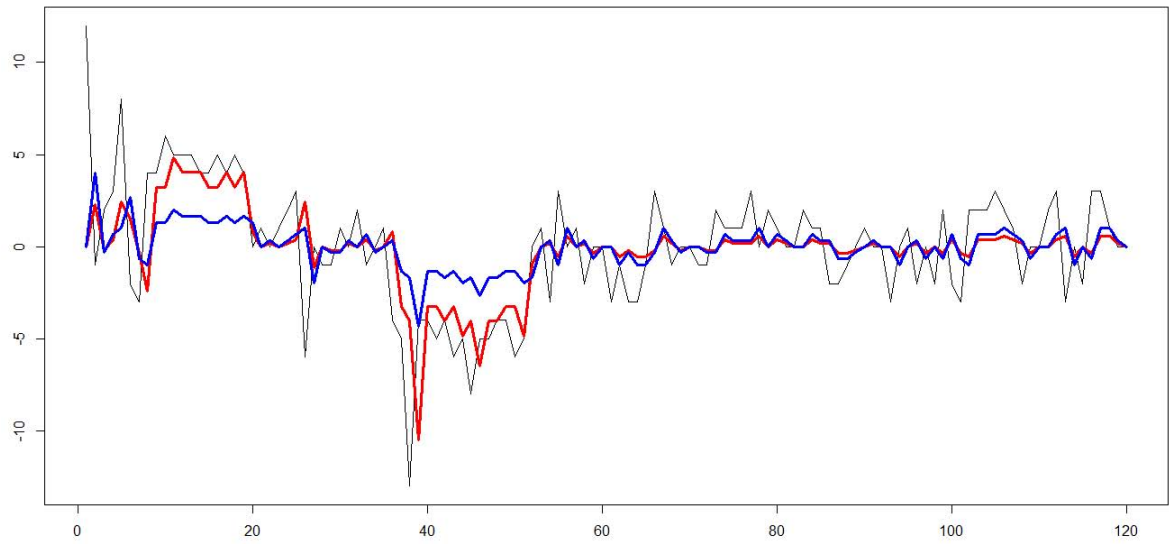

Figure 3: Black line – realization of the theft difference process;
red line – R2DLINAR$_1(\mathcal{M}, \mathcal{A})$ predicted values;
blue line – STINAR(1) predicted values.

Finally, we focus our attention on forecasting. To neutralize the impact of randomness on forecasting procedure, 10000 different sequences of size 24 are generated for each model, whereby corresponding model parameters are based on the training set. We compare these generated sequences of predictions with the prediction set and calculate FLSC. Table 5 shows results of the FLSC for all considered models. According to this table, the $\text{R2DLINAR}_1(\mathcal{M}, \mathcal{A})$ model has the largest FLSC among all considered models and thus, it provides the most accurate forecasting.

Table 5: FLSC for the prediction set of the theft difference.

	TINAR(1)	DLINAR(1)	STINAR(1)	$\text{R2DLINAR}_1(\mathcal{M}, \mathcal{A})$
FLSC	−65.712	−63.982	−64.151	−63.136

8. CONCLUSION

In this paper, we introduced a random environment integer-valued autoregressive process with discrete Laplace marginal distributions, $\text{RrDLINAR}_1(\mathcal{M}, \mathcal{A})$. Since the construction of this process was inspired by the work of [17], some of its features have been obtained in a similar way. Besides the definition of the $\text{RrDLINAR}_1(\mathcal{M}, \mathcal{A})$ process, we presented the full characterization of the process including its k-step ahead conditional expectation, correlation properties and the innovation process distribution. Parameter estimation was carried out using the method of moments and the strong consistency was proven. The YW estimates quality has been verified using subsamples of different sizes of 100 simulated data series, each of length 10000. At the very end, an application of the introduced model on real-life data series has been presented.

Further research might be performed in two directions. First of all, the model itself can be generalized to an order higher than 1, following the technique used in [11]. As for the second direction, an idea presented in this paper might be applied to the INAR process with asymmetric discrete Laplace marginal distribution, introduced by [7].

ACKNOWLEDGMENTS

We sincerely thank two reviewers for critically reading the manuscript and suggesting substantial improvements.

REFERENCES

[1] Al-Osh, M.A. and Alzaid, A.A. (1987). First-order integer-valued autoregressive (INAR(1)) process, *Journal of Time Series Analysis*, **8**, 261–275.

[2] Alzaid, A.A. and Al-Osh, M.A. (1993). Some autoregressive moving average processes with generalized Poisson marginal distributions, *Annals of the Institute of Statistical Mathematics*, **45**, 223–232.

[3] Bakouch, H.S. and Ristić, M.M. (2010). Zero truncated Poisson integer-valued AR(1) model, *Metrika*, **72**(2), 265–280.

[4] Barreto-Souza, W. and Bourguignon, M. (2015). A skew INAR(1) process on Z, *AStA Advances in Statistical Analysis*, **99**(2), 189–208.

[5] Brockwell, P.J. and Davis, R.A. (1987). *Time Series – Theory and Methods*, Springer, New York.

[6] Cui, Y.; Li, Q. and Zhu, F. (2021). Modeling Z-valued time series based on new versions of the Skellam INGARCH model, *Brazilian Journal of Probability and Statistics*, **35**, 293–314.

[7] Djordjević, M.S. (2017). An extension on INAR models with discrete Laplace marginal distributions, *Communications in Statistics – Theory and Methods*, **46**(12), 5896–5913.

[8] Du, J.G. and Li, Y. (1991). The integer-valued autoregressive (INAR(p)) model, *Journal of Time Series Analysis*, **12**, 129–142.

[9] Freeland, R.K. (2010). True integer value time series, *AStA Advances in Statistical Analysis*, **94**, 217–229.

[10] Gelman, A.; Hwang, J. and Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models, *Statistics and Computing*, **24**, 997–1016.

[11] Laketa, P.N.; Nastić, A.S. and Ristić, M.M. (2018). Generalized random environment INAR models of higher order, *Mediterranean Journal of Mathematics*, **15**, 9–30.

[12] Latour, A. (1998). Existence and stochastic structure of a non-negative integer-valued autoregressive process, *Journal of Time Series Analysis*, **19**, 439–455.

[13] Liu, Z.; Li, Q. and Zhu, F. (2020). Random environment binomial thinning integer-valued autoregressive process with Poisson or geometric marginal, *Brazilian Journal of Probability and Statistics*, **34**(2), 251–272.

[14] Liu, Z.; Li, Q. and Zhu, F. (2021). Semiparametric integer-valued autoregressive models on $\mathbb{Z}$, *Canadian Journal of Statistics*, **49**(4), 1317–1337.

[15] Liu, Z. and Zhu, F. (2021). A new extension of thinning-based integer-valued autoregressive models for count data, *Entropy*, **23**(1), 62.

[16] McKenzie, E. (1985). Some simple models for discrete variate time series, *Water Resources Bulletin*, **21**, 645–650.

[17] Nastić, A.S.; Ristić, M.M. and Djordjević, M.S. (2016). An INAR model with discrete Laplace marginal distribution, *Brazilian Journal of Probability and Statistics*, **30**(1), 107–126.

[18] Nastić, A.S.; Laketa, P.N. and Ristić, M.M. (2016). Random environment integer-valued autoregressive process, *Journal of Time Series Analysis*, **37**, 267–287.

[19] Nastić, A.S.; Laketa, P.N. and Ristić, M.M. (2019). Random environment INAR models of higher order, *RevStat – Statistical Journal*, **17**(1), 35–65.

[20] Qi, X.; Li, Q. and Zhu, F. (2019). Modeling time series of count with excess zeros and ones based on INAR(1) model with zero-and-one inflated Poisson innovations, *Journal of Computational and Applied Mathematics*, **346**, 572–590.

[21] Qian, L.; Li, Q. and Zhu, F. (2020). Modelling heavy-tailedness in count time series, *Applied Mathematical Modelling*, **82**, 766–784.

[22] Zheng, H.; Basawa, I.V. and Datta, S. (2006). Inference for pth-order random coefficient integer-valued autoregressive processes, *Journal of Time Series Analysis*, **27**, 411–440.

[23] Zheng, H.; Basawa, I.V. and Datta, S. (2007). First-order random coefficient integer-valued autoregressive processes, *Journal of Statistical Planning and Inference*, **137**, 212–229.