Filomat 39:21 (2025), 7173–7193
https://doi.org/10.2298/FIL2521173C

Published by Faculty of Sciences and Mathematics,
University of Niš, Serbia
Available at: http://www.pmf.ni.ac.rs/filomat

Value-correlation controlled integer-valued autoregressive process in random environment

Teodora D. Čamagić[a], Aleksandar S. Nastić[a,*], Miodrag S. Đorđević[a]

[a]*Faculty of Sciences and Mathematics, University of Niš, Serbia*

Abstract. The manuscript presents a new integer-valued first-order autoregressive process in a random environment, which is governed by two control processes. The first process defines the marginal distribution, while the second regulates the correlation structure within the model. The properties of the proposed model are examined in detail, providing insights into its theoretical foundations and practical implications. Two methods for estimating the unknown parameters are introduced: the Yule-Walker estimator and the conditional maximum likelihood estimator. A series of simulations assesses the efficiency of these estimation techniques and demonstrates their performance across various scenarios. The effectiveness of the introduced model is further evaluated through its application to real-life data.

1. Introduction

Due to the pervasive and frequent applications of counting processes in both natural and applied sciences, as well as in social activities, they have become a highly active area of research for many statisticians.

In the mid-1980s, INAR (integer-valued autoregressive) models were developed based on the binomial thinning operator introduced independently by McKenzie in [8], [9] and Al-Osh and Alzaid in [2]. These models were suitable when describing data representing the counting of random events and elements of a population that can only persist or disappear over time. Shortly thereafter, generalizations concerning marginal distributions and thinning operators were made. Notable contributions in these areas can be found in several sources, like [1], [3], [4] and [5]. Additionally, first discussions on higher-order INAR models and those with random coefficients are present in [10], [14] and [15]. However, INAR models with binomial thinning were not suitable for describing the population-related counts in situations where they can change not only through the disappearance of elements but also due to interaction among existing elements. A significant breakthrough was made by [13] by introducing the negative binomial thinning operator and the new integer-valued process, now known as NGINAR time series, defined using a geometric marginal distribution. All the advancements in developing these time series aimed to make the models more adaptable to real-world data. Nevertheless, all of these models exhibited stationary characteristics.

2020 *Mathematics Subject Classification*. Primary 62M10.
Keywords. random environment, INAR(1), negative binomial thinning, geometric marginals, controlling processes
Received: 22 January 2025; Revised: 20 May 2025; Accepted: 26 May 2025
Communicated by Dragan S. Djordjević
* Corresponding author: Aleksandar S. Nastić
Email addresses: teodora.camagic@pmf.edu.rs (Teodora D. Čamagić), anastic78@gmail.com (Aleksandar S. Nastić), miodrag.djordjevic@pmf.edu.rs (Miodrag S. Đorđević)
ORCID iDs: https://orcid.org/0009-0004-6347-4331 (Teodora D. Čamagić), https://orcid.org/0000-0003-4703-6492 (Aleksandar S. Nastić), https://orcid.org/0000-0002-4642-0348 (Miodrag S. Đorđević)

However, as almost everything of our observing interest is interconnected, specific non-stationary characteristics often kept appearing in the graphs. Hence, a completely new concept of model construction in [11], [12] and [7] by the RrINAR (random environment integer-valued autoregressive) models incorporated a precisely defined influence of random environment conditions known as random states. It allowed the marginal distribution parameter to vary over time so that its changes align with variations in environment conditions. Additionally, the random environment also influenced the distribution of the counting sequence, generating a time-flexible thinning operator. The idea is to use the marginal distribution and the distribution of the counting sequence with parameters that could take different values corresponding to the states of a random environment. It was assumed that the number of environmental states is finite, making the set of possible parameter values finite. Additionally, the processes of random environments are constructed as first-order Markov chains, meaning their future states depend only on the present and not on the past.

In the mentioned works on RrINAR models, clustering the process values and assigning the obtained clusters to the corresponding values of Markov chains of controlling variables actually determined the random environment process. However, besides the parameter values of the marginal distribution, this procedure also set parameter α (and consequently the process correlation) as a counting sequence parameter, which was not so intuitively clear. This limitation highlights the need for a more flexible modeling framework, in which the dynamics of the mean and the correlation can be independently controlled. By introducing two distinct control processes, each governed by its own Markov chain, the proposed model enables a clear separation between the influence of the environment on the marginal distribution and its effect on the autocorrelation structure. This structural distinction not only enhances the interpretability of the model but also allows for more accurate modeling of real-world systems, where external conditions may affect the mean level and temporal dependence of the data in different ways. Therefore, in this paper, we introduce a new model with two control processes, in which a separate Markov process determines its correlation structure. For example, consider the number of people waiting in line to buy ice cream on the street over time. The number of people waiting in line is influenced, for instance, by the temperature. If the temperature is exceptionally high, fewer people are likely to be walking on the street, resulting in fewer people buying ice cream. Instead of using temperature values in exact degrees, we can simplify the situation by introducing possible intervals of temperature values. So specifically, we can observe three different temperature-based random states: high, optimal, and low. On the other hand, if there is a long line or customers loudly comment on the ice cream's quality, passersby may want to buy and join the line. In this way, the quality of the ice cream, which can be either good or bad, affects the correlation of the number of customers at successive time intervals. This leads to the concept of correlation control, which refers to a mechanism by which the strength of dependence between consecutive observations is modulated by a latent process. In our case, a separate Markov chain governs the values of the thinning parameter α, thus dynamically adjusting the autocorrelation structure of the series in response to changes in the underlying environment.

This paper focuses on developing and analyzing a new model, highlighting its construction, properties, and practical applications. The first part of the study details the construction process for a model with an arbitrary marginal distribution. In the second part of this section, the emphasis shifts to specifying the geometric marginal distribution. The third section examines the fundamental properties of the defined model, including its distributional characteristics, correlation structure, and conditional moment properties. In the fourth section, we describe in detail the new concept of model control through two distinct random environment processes. Then, in the next section we present the methods for estimating the model's unknown parameters and we also provide the simulation results that demonstrate the effectiveness of the obtained parameter estimators. Finally, the sixth section discusses a potential application of the developed model to real-life data, showcasing its practical value and potential uses in various fields. At the end some concluding remarks ar given.

2. Construction of the process

Definition 2.1. *A sequence of random variables $\{Z_n\}$, where $n \in N_0$, is called a random environment process if it forms a Markov chain with the state space $E_r = \{1, 2, \ldots, r\}$, $r \in N$.*

We consider two independent sequences of random variables $\{Z_n\}$ and $\{V_n\}$. Both sequences are Markov chains with state spaces $E_{r_1} = \{1, 2, \ldots, r_1\}$ and $E_{r_2} = \{1, 2, \ldots, r_2\}$, respectively. Using these sequences, we will construct the following model.

Definition 2.2. *A non-negative integer-valued sequence of random variables $\{X_n(Z_n, V_n)\}$, $n \in N_0$, is said to be the value-correlation controlled integer-valued autoregressive process of order 1 (V-CCINAR(1)), with states r_1 and r_2 if it is given by*

$$X_n(Z_n, V_n) = \sum_{i=1}^{X_{n-1}(Z_{n-1}, V_{n-1})} U_i + \varepsilon_n(Z_{n-1}, Z_n, V_{n-1}, V_n), \quad n \in N,$$

where

$$X_n(Z_n, V_n) = \sum_{z=1}^{r_1} \sum_{v=1}^{r_2} X_n(z, v) I_{\{Z_n = z\}} I_{\{V_n = v\}},$$

$$\varepsilon_n(Z_{n-1}, Z_n, V_{n-1}, V_n) = \sum_{z_1=1}^{r_1} \sum_{z_2=1}^{r_1} \sum_{v_1=1}^{r_2} \sum_{v_2=1}^{r_2} \varepsilon_n(z_1, z_2, v_1, v_2) I_{\{Z_{n-1}=z_1, Z_n=z_2\}} I_{\{V_{n-1}=v_1, V_n=v_2\}},$$

$\{U_i\}$, $i \in N$, is a counting sequence of independent and identically distributed (i.i.d.) random variables generating a thinning operator. Let $\{Z_n\}$ be a value controlling process (VC-process) and $\{V_n\}$ a correlation controlling process (CC-process). These are random environment processes with state spaces E_{r_1} and E_{r_2}, respectively, as defined in Definition 1. Here, Z_n determines the distribution of the random variable X_n, while the random variable V_n dictates the distribution of the counting sequence $\{U_i\}$. The corresponding sets of parameter values are given by $\mathcal{M} = \{\mu_1, \mu_2, \ldots, \mu_{r_1}\}$ and $\mathcal{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_{r_2}\}$.

The sequences $\{Z_n\}$ and $\{V_n\}$ are mutually independent and consist of i.i.d. random variables that satisfy the following conditions:
(1) $\{Z_n\}$, $\{V_n\}$, for $n \in N_0$ and $\{\varepsilon_n(i, j, i', j')\}$, $n \in N$, $i, j \in E_{r_1}$, $i', j' \in E_{r_2}$, are mutually independent for $n \geq 0$,
(2) Z_m and $\varepsilon_m(i, j, i', j')$ are independent of $X_n(l, k)$ for $n < m$ and any $l, i, j \in E_{r_1}$ and $k, i', j' \in E_{r_2}$.

Theorem 2.3. *Time series $\{X_n(Z_n, V_n), Z_n, V_n\}$ given by Definitions 2.1 and 2.2 is a Markov chain of the first order.*

Proof. To ease notation, let $\mathbf{Y}_n = (X_n(Z_n, V_n), Z_n, V_n)$ and $\mathbf{y}_n = (x_n, z_n, v_n)$. First, we see that it holds

$$\mathbf{Y}_n = \mathbf{y}_n \Leftrightarrow (Z_n = z_n, V_n = v_n, X_n(z_n, v_n) = x_n).$$

Let $A = \{\mathbf{Y}_s = \mathbf{y}_s, 0 \leq s < n-1\}$, $p_{n-1,n} = P(Z_n = z_n | Z_{n-1} = z_{n-1})$ and $p^*_{n-1,n} = P(V_n = v_n | V_{n-1} = v_{n-1})$.
Now, let's consider the conditional probability $P_{n-1,n} = P(\mathbf{Y}_n = \mathbf{y}_n | \mathbf{Y}_{n-1} = \mathbf{y}_{n-1}, A)$. Using the definition of the process, we have that

$$\mathbf{Y}_n = \mathbf{y}_n \Leftrightarrow \left(\sum_{i=1}^{X_{n-1}(Z_{n-1}, V_{n-1})} U_i + \varepsilon_n(Z_{n-1}, Z_n, V_{n-1}, V_n) = x_n \quad \wedge \quad Z_n = z_n \quad \wedge \quad V_n = v_n \right),$$

while it is

$$\mathbf{Y}_{n-1} = \mathbf{y}_{n-1} \Leftrightarrow (Z_{n-1} = z_{n-1} \quad \wedge \quad V_{n-1} = v_{n-1} \quad \wedge \quad X_{n-1}(z_{n-1}, v_{n-1}) = x_{n-1}).$$

From there, we conclude that

$$P_{n-1,n} = P\left(\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n) = x_n, Z_n = z_n, V_n = v_n | B \right),$$

where we have denoted the event $\{Z_{n-1} = z_{n-1}\} \cap \{V_{n-1} = v_{n-1}\} \cap \{X_{n-1}(z_{n-1}, v_{n-1}) = x_{n-1}\} \cap A$ by B. Due to the independence of the corresponding random variables, the probability above is equal to the product of the probabilities

$$P\left(\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n) = x_n | B\right) \cdot P(Z_n = z_n|B) \cdot P(V_n = v_n|B).$$

Since $\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n)$ does not depend on the random variables that appear in B, the first factor of the product is equal to $P\Big(\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n) = x_n\Big)$. Since $\{Z_n\}$ and $\{V_n\}$ are first order Markov chains, Z_n and V_n depend only on Z_{n-1} and V_{n-1}, respectively, so the last two factors in the product are equal to $P(Z_n = z_n|Z_{n-1} = z_{n-1}) = p_{n-1,n}$ and $P(V_n = v_n|V_{n-1} = v_{n-1}) = p *_{n-1,n}$. Finally, we have that

$$P_{n-1,n} = p_{n-1,n} \cdot p *_{n-1,n} \cdot P\left(\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n) = x_n\right). \tag{1}$$

We will apply the procedure similarly to $P(\mathbf{Y}_n = \mathbf{y}_n|\mathbf{Y}_{n-1} = \mathbf{y}_{n-1})$. Again, we can express this probability as the product of three probabilities. In the same manner as before, we conclude that.

$$P(\mathbf{Y}_n = \mathbf{y}_n|\mathbf{Y}_{n-1} = \mathbf{y}_{n-1}) = p_{n-1,n} \cdot p *_{n-1,n} \cdot P\left(\sum_{i=1}^{x_{n-1}} U_i + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n) = x_n\right). \tag{2}$$

From (1) and (2), it follows that

$$P(\mathbf{Y}_n = \mathbf{y}_n|\mathbf{Y}_{n-1} = \mathbf{y}_{n-1}) = P(\mathbf{Y}_n = \mathbf{y}_n|\mathbf{Y}_{n-1} = \mathbf{y}_{n-1}, A),$$

indicating that $\{X_n(Z_n, V_n), Z_n, V_n\}$ forms a first-order Markov chain. □

2.1. Construction of the process with geometric marginal distribution and a negative binomial thinning operator

Let's consider a model that is assumed to have a geometric marginal distribution due to the overdispersion observed in real-world data. Additionally, because of the dynamic nature of the systems being observed, we assume that the model is based on a negative binomial thinning operator.

Definition 2.4. *Let $\{z_n\}$ and $\{v_n\}$ be realizations of the random environment processes $\{Z_n\}$ and $\{V_n\}$, with r_1 and r_2 states, respectively.*

We define the sequence $\{X_n(z_n, v_n)\}$ for $n \in \mathbb{N}_0$ as a value-correlation controlled process with r_1 and r_2 states and geometric marginal distribution, based on the negative binomial thinning operator (V-CCNGINAR(1)). The random variable $X_n(z_n, v_n)$ at time n is defined as follows:

$$X_n(z_n, v_n) = \alpha_{v_n} * X_{n-1}(z_{n-1}, v_{n-1}) + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n), \quad n \in N,$$

where $X_n \sim Geom(\frac{\mu_{z_n}}{1+\mu_{z_n}})$, has a geometric distribution with expectation μ_{z_n}, that is

$$P(X_n(z_n, v_n) = x) = \frac{\mu_{z_n}^x}{(1 + \mu_{z_n})^{x+1}}, \quad x \in N_0,$$

with possible parameter values, $\mu_{z_n} \in \{\mu_1, \mu_2, \ldots, \mu_{r_1}\}, r_1 \in N$. Also, for $\alpha_{v_n} \in (0, 1)$, $v_n \in \{1, 2, \ldots, r_2\}$, counting sequence $\{U_i\}$, $i \in N$, incorporated in $\alpha_{v_n}*$ make a sequence of i.i.d. random variables with probability mass function (pmf) given as

$$P(U_i = u) = \frac{\alpha_{v_n}^u}{(1 + \alpha_{v_n})^{u+1}}, \quad u \in N_0, \quad \alpha_{v_n} \in \{\alpha_1, \alpha_2, \ldots, \alpha_{r_2}\}, \quad r_2 \in N,$$

where $\alpha_{v_n} = \sum_{v=1}^{r_2} \alpha_v I(V_n = v)$, *namely*

$$\alpha_{v_n} = \begin{cases} \alpha_1, & \text{with probability} \quad P(V_n = 1), \\ \alpha_2, & \text{with probability} \quad P(V_n = 2), \\ \vdots & \\ \alpha_{r_2}, & \text{with probability} \quad P(V_n = r_2). \end{cases}$$

A random variable $\varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n)$ *is defined as outlined in Definition* 2.2.

The previous definition clearly illustrates how the process $\{z_n\}$ influences the distribution of X_n. Conversely, the realized values of the process $\{V_n\}$ determine the distribution of the counting sequence, which affects the correlation structure of the process.

This model is one of the many possible special cases of the model given by Definition 2.2 that can be defined, which presents numerous opportunities for further research.

3. Properties of the process

In this section, we will derive several properties of the model described above. The next theorem identifies the distribution of the innovation random variable. Many steps and derivations in the proofs of the theorems in this section closely follow the techniques used in the corresponding results from [11]. However, since we are dealing here with two random states that control the processes, the proofs are presented in detail.

Theorem 3.1. *Let* $\{X_n(z_n, v_n)\}$, *for* $n \in N_0$, *be the time series model introduced by Definition 2.4. Assume that* $\mu_1 > 0, \mu_2 > 0, \ldots, \mu_{r_1} > 0$. *If* $0 \le \alpha_{v_n} \le \min\{\frac{\mu_l}{1+\mu_k}; k, l \in E_{r_1}\}$, *for any* $v_n \in E_{r_2}$, *then when* $z_n = j$ *and* $z_{n-1} = i$, *with* $i, j \in E_{r_1}$ *and* $v_n = j'$ *and* $v_{n-1} = i'$, *where* $j', i' \in E_{r_2}$, *the distributions of the innovation random variables can be expressed as mixtures of two geometric distributions, i.e. as*

$$\varepsilon_n(i, j, i', j') \stackrel{d}{=} \begin{cases} Geom(\frac{\mu_j}{1+\mu_j}), & w.p. \quad 1 - \frac{\alpha_{j'}\mu_i}{\mu_j - \alpha_{j'}}, \\ Geom(\frac{\alpha_{j'}}{1+\alpha_{j'}}), & w.p. \quad \frac{\alpha_{j'}\mu_i}{\mu_j - \alpha_{j'}}, \end{cases} \tag{3}$$

where the equality refers to equality in distribution.

Proof. The probability generating function is considered. Based on Definition 2.4 and the conditions $z_n = j$, $z_{n-1} = i$, $v_n = j'$, and $v_{n-1} = i'$, it follows that $X_n(j, j') = \alpha_{j'} * X_{n-1}(i, i') + \varepsilon_n(i, j, i', j')$. Therefore,

$$\Phi_{X_n(j,j')}(s) = E\Big(s^{X_n(j,j')}\Big) = E\Big(s^{\alpha_{j'} * X_{n-1}(i,i') + \varepsilon_n(i,j,i',j')}\Big).$$

Based on the independence of the random variables $\varepsilon_n(i, j, i', j')$ and $X_{n-1}(i, j)$, and considering the properties of expectation, we can conclude that

$$E\Big(s^{\alpha_{j'} * X_{n-1}(i,i') + \varepsilon_n(i,j,i',j')}\Big) = E\Big(s^{\alpha_{j'} * X_{n-1}(i,i')}\Big) E\Big(s^{\varepsilon_n(i,j,i',j')}\Big).$$

Furthermore, utilizing the properties of the negative binomial thinning operator, we find that

$$E\Big(s^{\alpha_{j'} * X_{n-1}(i,i')}\Big) = E\Big((Es^{U_1})^{X_{n-1}(i,i')}\Big) = \Phi_{X_{n-1}(i,i')}\Big(\Phi_{U_1}(s)\Big).$$

Thus, it follows that

$$\Phi_{X_n(j,j')}(s) = \Phi_{X_{n-1}(i,i')}\Big(\Phi_{U_1}(s)\Big)\Phi_{\varepsilon_n(i,j,i',j')}(s),$$

or equivalently,

$$\Phi_{\varepsilon_n(i,j,i',j')}(s) = \frac{\Phi_{X_n(j,j')}(s)}{\Phi_{X_{n-1}(i,i')}\Big(\Phi_{U_1}(s)\Big)}. \tag{4}$$

Given that the random variable $X_n(j, j')$ follows a geometric distribution with parameter μ_j, we find that

$$\Phi_{X_n(j,j')}(s) = \frac{1}{1+\mu_j-\mu_j s}, \tag{5}$$

which is derived from the properties of the geometric distribution. Similarly, we have

$$\Phi_{X_{n-1}(i,i')}(s) = \frac{1}{1+\mu_i-\mu_i s}. \tag{6}$$

Moreover, the random variable U_1 also has a geometric distribution with parameter $\alpha_{j'}$, thus

$$\Phi_{U_1}(s) = \frac{1}{1+\alpha_{j'}-\alpha_{j'}s}.$$

From the fact that the random variable $X_{n-1}(i, i')$ has a geometric distribution with expectation μ_i, we conclude that

$$\begin{aligned}\Phi_{X_{n-1}(i,i')}\Big(\Phi_{U_1}(s)\Big) &= E\Big(\Big(\frac{1}{1+\alpha_{j'}-\alpha_{j'}s}\Big)^{X_{n-1}(i,i')}\Big)\\ &= \frac{1}{1+\mu_i-\mu_i\frac{1}{1+\alpha_{j'}-\alpha_{j'}s}} = \frac{1+\alpha_{j'}-\alpha_{j'}s}{1+\alpha_{j'}(1+\mu_i)-\alpha_{j'}(1+\mu_i)s}.\end{aligned}$$

Now, replacing the previous equality and equations (5) and (6) into (4), we obtain:

$$\begin{aligned}\Phi_{\varepsilon_n(i,j,i',j')}(s) &= \frac{1+\alpha_{j'}(1+\mu_i)-\alpha_{j'}(1+\mu_i)s}{(1+\alpha_{j'}-\alpha_{j'}s)(1+\mu_j-\mu_j s)}\\ &= \frac{1+\alpha_{j'}-\alpha_{j'}s+\alpha_{j'}\mu_i+\alpha_{j'}\mu_i s}{(1+\alpha_{j'}-\alpha_{j'}s)(1+\mu_j-\mu_j s)}\\ &= \frac{1}{1+\mu_j-\mu_j s}+\frac{\alpha_{j'}\mu_i}{\mu_j\alpha_{j'}}\cdot\frac{\mu_j-\mu_j s-\alpha_{j'}+\alpha_{j'}s}{(1+\alpha_{j'}-\alpha_{j'}s)(1+\mu_j-\mu_j s)}\\ &= \frac{1}{1+\mu_j-\mu_j s}+\frac{\alpha_{j'}\mu_i}{\mu_j\alpha_{j'}}\cdot\left(\frac{1}{1+\alpha_{j'}-\alpha_{j'}s}-\frac{1}{1+\mu_j-\mu_j s}\right)\\ &= \left(1-\frac{\alpha_{j'}\mu_i}{\mu_j-\alpha_{j'}}\right)\cdot\frac{1}{1+\mu_j-\mu_j s}+\frac{\alpha_{j'}\mu_i}{\mu_j-\alpha_{j'}}\cdot\frac{1}{1+\alpha_{j'}-\alpha_{j'}s}.\end{aligned}$$

We find that the probability generating function of the random variable $\varepsilon_n(i, j, i', j')$ can be expressed in the form of (3), given the condition that $0 \le \alpha_{j'} \le \frac{\mu_j}{1+\mu_i}, \forall j' \in E_{r_2}$. This condition guarantees that the corresponding probabilities in (3) are non-negative. Since, $i, j \in E_{r_1}$ and $i', j' \in E_{r_2}$ are arbitrary, it follows that for any $i, j \in E_{r_1}$, $0 \le \alpha_{j'} \le \frac{\mu_j}{1+\mu_i}, \forall j' \in E_{r_2}$. This is achieved if $\alpha_{j'}$ satisfies

$$\alpha_{j'} \in \bigcap_{k,l\in E_{r_1}}\left[0, \frac{\mu_l}{1+\mu_k}\right], \quad \forall j' \in E_{r_2}.$$

In other words, we require that $0 \le \alpha_{j'} \le \min\left\{\frac{\mu_l}{1+\mu_k}, k, l \in E_{r_1}\right\}, \forall j' \in E_{r_2}$. □

In the following theorem, we will provide the expectation and variance of the innovation process.

Theorem 3.2. *Let $z_{n-1} = i$, $z_n = j$, $v_{n-1} = i'$ and $v_n = j'$, where $i, j \in E_{r_1}$ and $i', j' \in E_{r_2}$. The expectation and variance of the random variable $\varepsilon_n(i, j, i', j')$ are given by $E\Big(\varepsilon_n(i, j, i', j')\Big) = \mu_j - \alpha_{j'}\mu_i$ and $D\Big(\varepsilon_n(i, j, i', j')\Big) = \mu_j(1 + \mu_j) - \alpha_{j'}\mu_i(1 + 2\alpha_{j'} + \alpha_{j'}\mu_i)$.*

Proof. Let $\Phi_\varepsilon(s)$ be the probability generating function of the random variable $\varepsilon_n(i, j, i', j')$. Then, we have the expectation $E\Big(\varepsilon_n(i, j, i', j')\Big) = \Phi'_\varepsilon(1)$ and the variance is given by $D\Big(\varepsilon_n(i, j)\Big) = \Phi''_\varepsilon(1) + \Phi'_\varepsilon(1)\Big(1 - \Phi'_\varepsilon(1)\Big)$. Using the proof from the previous theorem, we can conclude that $\Phi'_\varepsilon(1) = \mu_j - \alpha_{j'}\mu_i$ and $\Phi''_\varepsilon(1) = 2(\mu_j^2 - \alpha_{j'}\mu_i\mu_j - \alpha_{j'}^2\mu_i)$. Thus, we have that

$$E\Big(\varepsilon_n(i, j, i', j')\Big) = \mu_j - \alpha_{j'}\mu_i$$

and

$$\begin{aligned} D\Big(\varepsilon_n(i, j, i', j')\Big) &= 2(\mu_j^2 - \alpha_{j'}\mu_i\mu_j - \alpha_{j'}^2\mu_i) + (\mu_j - \alpha_{j'}\mu_j)(1 - \mu_j + \alpha_{j'}\mu_i) \\ &= 2\mu_j^2 - 2\alpha_{j'}\mu_i\mu_j - 2\alpha_{j'}^2\mu_i + \mu_j - \mu_j^2 + \alpha_{j'}\mu_i\mu_j - \alpha_{j'}\mu_i + \alpha_{j'}\mu_i\mu_j - \alpha_{j'}^2\mu_i^2 \\ &= \mu_j^2 + \mu_j - \alpha_{j'}\mu_i - 2\alpha_{j'}^2\mu_i - \alpha_{j'}^2\mu_i^2 \\ &= \mu_j(1 + \mu_j) - \alpha_{j'}\mu_i(1 + 2\alpha_{j'} + \alpha_{j'}\mu_i). \end{aligned}$$

□

3.1. Correlation structure

The key feature of any autoregressive process is its covariance-correlation structure, which represents the interconnection and dependence of individual elements. In our model, this structure is governed by a Markov process that controls the value of a thinning parameter α, allowing the strength of dependence between observations to vary over time. This flexibility enables the model to capture realistic temporal behavior in systems where external conditions-such as seasonality, demand cycles, or environmental fluctuations - affect not only the marginal distribution, but also the persistence and variability of the process. Analyzing these functions under such stochastic control is essential for understanding and applying the model in non-stationary environments.

Theorem 3.3. *Let $\{X_n(z_n, v_n)\}$, where $n \in N_0$, be the process given by Definition 2.4. Assume that $\mu_1 > 0, \mu_2 > 0, \ldots, \mu_{r_1} > 0$. Then, we have the following:*

(i) the covariance function between the random variables $X_n(z_n, v_n)$ and $X_{n-k}(z_{n-k}, v_{n-k})$, for $k \in \{0, 1, \ldots, n\}$ is positive. It is given as

$$\gamma_n^{(k)} = Cov(X_n(z_n, v_n), X_{n-k}(z_{n-k}, v_{n-k})) = \alpha_{v_n}^k \cdot \mu_{z_{n-k}}(1 + \mu_{z_{n-k}}),$$

(ii) the correlation function between the random variables $X_n(z_n, v_n)$ and $X_{n-k}(z_{n-k}, v_{n-k})$, for $k \in \{1, 2, \ldots, n\}$ is positive, and always less than 1. It is given by

$$\rho_n^{(k)} = Corr(X_n(z_n, v_n), X_{n-k}(z_{n-k}, v_{n-k})) = \alpha_{v_n}^k \cdot \sqrt{\frac{\mu_{z_{n-k}}(1 + \mu_{z_{n-k}})}{\mu_{z_n}(1 + \mu_{z_n})}}.$$

Proof. (*i*) Model $\{X_n(z_n, v_n)\}$, with $n \in N_0$, satisfies the equation $X_n(z_n, v_n) = \alpha_{v_n} * X_{n-1}(z_{n-1}, v_{n-1}) + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n)$, so we have that

$$\begin{aligned} \gamma_n^{(k)} &= Cov(X_n(z_n, v_n), X_{n-k}(z_{n-k}, v_{n-k})) \\ &= Cov(\alpha_{v_n} * X_{n-1}(z_{n-1}, v_{n-1}) + \varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n), X_{n-k}(z_{n-k}, v_{n-k})). \end{aligned}$$

Due to the independence of random variables $\varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n)$ and $X_{n-k}(z_{n-k}, v_{n-k})$ the expression

$$\gamma_n^{(k)} = Cov(\alpha_{v_n} * X_{n-1}(z_{n-1}, v_{n-1}), X_{n-k}(z_{n-k}, v_{n-k}))$$

is valid. For random variables X and Y that are independent of the counting sequence included in the thinning operator, it holds that $Cov(\alpha * X, Y) = \alpha \cdot Cov(X, Y)$. Using this property, we have that

$$\gamma_n^{(k)} = \alpha_{v_n} \cdot Cov(X_{n-1}(z_{n-1}, v_{n-1}), X_{n-k}(z_{n-k}, v_{n-k})) = \alpha_{v_n} \cdot \gamma_{n-1}^{(k-1)}.$$

Applying this equation $k-1$ times, we find that the covariance function between the random variables $X_n(z_n, v_n)$ and $X_{n-k}(z_{n-k}, v_{n-k})$ is given by $\gamma_n^{(k)} = \alpha_{v_n}^k \cdot \gamma_{n-k}^{(0)}$. Here, $\gamma_{n-k}^{(0)}$ represents the variance of the random variable $X_{n-k}(z_{n-k}, v_{n-k})$ which follows a geometric distribution with mean parameter $\mu_{z_{n-k}}$. Thus we can express the covariance as $\gamma_n^{(k)} = \alpha_{v_n}^k \cdot \mu_{z_{n-k}}(1 + \mu_{z_{n-k}})$.

(*ii*) The correlation function between the random variables $X_n(z_n, v_n)$ and $X_{n-k}(z_{n-k}, v_{n-k})$ can be expressed in terms of the corresponding covariance function as follows:

$$\rho_n^{(k)} = \frac{\gamma_n^{(k)}}{\sqrt{\gamma_n^{(0)}\gamma_{n-k}^{(0)}}}.$$

Now, using result (i), we obtain that the correlation function between the random variables $X_n(z_n, v_n)$ and $X_{n-k}(z_{n-k}, v_{n-k})$ is given by

$$\rho_n^{(k)} = \alpha_{v_n}^k \cdot \sqrt{\frac{\mu_{z_{n-k}}(1 + \mu_{z_{n-k}})}{\mu_{z_n}(1 + \mu_{z_n})}}.$$

It is still necessary to show that the correlations $\rho_n^{(k)}$ are always less than 1. Based on Theorem 3.1 and Theorem 3.2, we can conclude that the expectation of the random variable $\varepsilon_n(i, j, i', j')$ for $i, j \in E_{r_1}, i', j' \in E_{r_2}$ is positive. Specifically, from Theorem 3.1 is

$$0 \leq \alpha_{j'} \leq \min\left\{\frac{\mu_l}{1 + \mu_k}; k, l \in E_r\right\} \leq \frac{\mu_j}{1 + \mu_i} < \frac{\mu_j}{\mu_i}, \forall j' \in E_{r_2}. \tag{7}$$

Now, from (7), for $i = z_{n-k}$, $j = z_n$, $i' = v_{n-k}$ and $j' = v_n$, we have that

$$\alpha_{v_n} < \frac{\mu_{z_n}}{1 + \mu_{z_{n-k}}} < \frac{\mu_{z_n}}{\mu_{z_{n-k}}} < \frac{1 + \mu_{z_n}}{\mu_{z_{n-k}}},$$

which implies that

$$\alpha_{v_n} < \sqrt{\frac{\mu_{z_n}(1 + \mu_{z_n})}{\mu_{z_{n-k}}(1 + \mu_{z_{n-k}})}}. \tag{8}$$

Since $\alpha_{v_n} < 1$, then we have that $\alpha_{v_n}^k < \alpha_{v_n}$, so from (8) it finally follows that

$$\rho_n^{(k)} < \alpha_{v_n} \cdot \sqrt{\frac{\mu_{z_{n-k}}(1 + \mu_{z_{n-k}})}{\mu_{z_n}(1 + \mu_{z_n})}} < 1.$$

□

As expected, it is clear that both $\gamma_n^{(k)}$ and $\rho_n^{(k)}$ converge to 0, as $k \to \infty$.

3.2. Conditional properties

Now, we will examine some regression properties of the process defined above, as illustrated by the following theorem. These conditional properties-including the conditional expectation and variance-are essential for understanding the short-term dynamics of the process. In particular, they reveal how the current state of the environment and the recent history of the process jointly influence future outcomes. This information is critical for forecasting and inference, especially in systems where external conditions evolve over time and affect not only the mean level but also the uncertainty of future values.

Theorem 3.4. *Let* $\{X_n(z_n, v_n)\}$, *where* $n \in N_0$, *be the process defined by Definition 2.4. Assume that* $\mu_1, \mu_2, \ldots, \mu_{r_1} > 0$. *Then, we have the following:*
(a) The conditional expectation of the random variable $X_{n+k}(z_{n+k}, v_{n+k})$ *on* $X_n(z_n, v_n)$ *is expressed as:*

$$E\Big(X_{n+k}(z_{n+k}, v_{n+k})|X_n(z_n, v_n)\Big) = \alpha_{v_{n+k}}\alpha_{v_{n+k-1}}\ldots\alpha_{v_{n+1}}X_n(z_n, v_n) + \mu_{z_{n+k}} - \alpha_{v_{n+k}}\alpha_{v_{n+k-1}}\ldots\alpha_{v_{n+1}}\mu_{z_n}.$$

(b) The conditional variance of the random variable $X_{n+k}(z_{n+k}, v_{n+k})$ *on* $X_n(z_n, v_n)$ *is expressed as:*

$$\begin{aligned}
D\Big(X_{n+k}(z_{n+k}, v_{n+k})|X_n(z_n, v_n)\Big) = &\sum_{l=0}^{k-2}\prod_{p=0}^{l-2}\alpha^2_{v_{n+k-p}}\alpha_{v_{n+k-l}}(1+\alpha_{v_{n+k-l}})\\
&\times\Big[\alpha_{v_{n+k-l-1}}\ldots\alpha_{v_{n+1}}X_n(z_n, v_n) + \mu_{z_{n+k-l-1}} - \alpha_{v_{n+k-l-1}}\ldots\alpha_{v_{n+1}}\mu_{z_n}\Big]\\
&+\alpha_{v_{n+k}}(1+\alpha_{v_n+k})\alpha_{v_{n+k-1}}\ldots\alpha_{v_{n+1}}X_n(z_n, v_n)\\
&-\alpha_{v_{n+k}}(1+\alpha_{v_n+k})\alpha_{v_{n+k-1}}\ldots\alpha_{v_{n+1}}\mu_{z_n}\\
&+\alpha_{v_{n+k}}(1+\alpha_{v_{n+k}})\mu_{z_{n+k-1}}(1-\alpha^2_{v_{n+k}}\mu_{z_{n+k-1}}(1+\mu_{z_{n+k-1}}))\\
&+\mu_{z_{n+k}}(1+\mu_{z_{n+k}})\\
&+\sum_{l=1}^{k-1}\prod_{p=0}^{l-1}\alpha^2_{v_{n+k-p}}\Big(\mu_{z_{n+k-l}}(1+\mu_{z_{n+k-l}})\\
&\qquad -\alpha_{v_{n+k-l}}\mu_{z_{n+k-l}}(1+2\alpha_{v_{n+k-l}}+\alpha_{v_{n+k-l}}\mu_{z_{n+k-l}})\Big).
\end{aligned}$$

Proof. (a) To simplify the notation, let

$$\mu_{n+k|n} = E\Big(X_{n+k}(z_{n+k}, v_{n+k})|X_n(z_n, v_n)\Big)$$

and

$$\mu_{\varepsilon_n} = E\Big(\varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n)\Big).$$

From the definition of time series model and the independence of the random variables $\varepsilon_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})$ and $X_{n+k-1}(z_{n+k-1}, v_{n+k-1})$, we obtain

$$\begin{aligned}
\mu_{n+k|n} &= E\Big(\alpha_{v_{n+k}} * X_{n+k-1}(z_{n+k-1}, v_{n+k-1}) + \varepsilon_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})|X_n(z_n, v_n)\Big)\\
&= E\Big(\alpha_{v_{n+k}} * X_{n+k-1}(z_{n+k-1}, v_{n+k-1})|X_n(z_n, v_n)\Big) + E\Big(\varepsilon_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})\Big).
\end{aligned}$$

Since $E(\alpha * X|Y) = \alpha E(X|Y)$ holds for random variables X and Y, which are independent of counting sequence within the operator, we have that $\mu_{n+k|n} = \alpha_{v_{n+k}}\mu_{n+k-1|n} + \mu_{\varepsilon_{n+k}}$. By applying the last equation $k-1$ times, we can derive further results.

$$\begin{aligned}
\mu_{n+k|n} &= \alpha_{v_{n+k}}(\alpha_{v_{n+k-1}}\mu_{n+k-2|n} + \mu_{\varepsilon_{n+k-1}}) + \mu_{\varepsilon_{n+k}}\alpha_{v_{n+k}}\alpha_{v_{n+k-1}}(\alpha_{v_{n+k-2}}\mu_{n+k-3|n} + \mu_{\varepsilon_{n+k-2}}) + \alpha_{v_{n+k}}\mu_{\varepsilon_{n+k-1}} + \mu_{\varepsilon_{n+k}}\\
&= \alpha_{v_{n+k}}\alpha_{v_{n+k-1}}\alpha_{v_{n+k-2}}\mu_{n+k-3|n} + \alpha_{v_{n+k}}\alpha_{v_{n+k-1}}\mu_{\varepsilon_{n+k-2}} + \alpha_{v_{n+k}}\mu_{\varepsilon_{n+k-1}} + \mu_{\varepsilon_{n+k}}\\
&= \alpha_{v_{n+k}}\alpha_{v_{n+k-1}}\ldots\alpha_{v_{n+1}}\mu_{n|n} + \sum_{l=0}^{k-1}\left(\prod_{p=0}^{l}\alpha_{v_{n+k-p}}\right)\mu_{\varepsilon_{n+k-l-1}}.
\end{aligned}$$

Now, we substitute $\mu_{n|n} = X_n(z_n, v_n)$ and use the obtained result for expectation of random variable $\varepsilon_{n+k-l}, l \in \{0, 1, \ldots, k-1\}$. This leads to

$$\begin{aligned}
\mu_{n+k|n} &= \alpha_{v_{n+k}}\alpha_{v_{n+k-1}} \ldots \alpha_{v_{n+1}} X_n(z_n, v_n) + \mu_{z_{n+k}} - \alpha_{v_{n+k}}\mu_{z_{n+k-1}} + \sum_{l=0}^{k-1}\left(\prod_{p=0}^{l}\alpha_{v_{n+k-p}}\right)\left(\mu_{z_{n+k-l-1}} - \alpha_{v_{n_k-l-1}}\mu_{z_{n+k-l-2}}\right) \\
&= \alpha_{v_{n+k}}\alpha_{v_{n+k-1}} \ldots \alpha_{v_{n+1}} X_n(z_n, v_n) + \mu_{z_{n+k}} - \alpha_{v_{n+k}}\mu_{z_{n+k-1}} + \alpha_{v_{n+k}}\mu_{z_{n+k-1}} - \alpha_{v_{n+k}}\alpha_{v_{n+k-1}}\mu_{z_{n+k-2}} \\
&\quad + \sum_{l=1}^{k-1}\left(\prod_{p=0}^{l}\alpha_{v_{n+k-p}}\right)\left(\mu_{z_{n+k-l-1}} - \alpha_{v_{n+k-l-1}}\mu_{z_{n+k-l-2}}\right) \\
&= \alpha_{v_{n+k}}\alpha_{v_{n+k-1}} \ldots \alpha_{v_{n+1}} X_n(z_n, v_n) + \mu_{z_{n+k}} - \alpha_{v_{n+k}}\alpha_{v_{n+k-1}} \ldots \alpha_{v_{n+1}}\mu_{z_n}. \qquad (9)
\end{aligned}$$

(b) Now, let us consider the conditional variance. As above, in order to ease the notation we introduce

$$\sigma^2_{n+k|n} = D\Big(X_{n+k}(z_{n+k}, v_{n+k})|X_n(z_n, v_n)\Big)$$

and

$$\sigma^2_{\varepsilon_n} = D\Big(\varepsilon_n(z_{n-1}, z_n, v_{n-1}, v_n)\Big).$$

Given the independence of the random variables $\varepsilon_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})$ from $X_n(z_n, v_n)$ and $X_{n+k-1}(z_{n+k-1}, v_{n+k-1})$, we conclude that

$$\begin{aligned}
E\Big(X^2_{n+k}(z_{n+k}, v_{n+k})|X_n(z_n, v_n)\Big) &= E\Big([\alpha_{v_{n+k}} * X_{n+k-1}(z_{n+k-1}, v_{n+k-1}) + \varepsilon_{n+k}(z_{n+k-1}, z_{n+k}, v_{n_k-1}, v_{n+k})]^2|X_n(z_n, v_n)\Big) \\
&= E\Big([\alpha_{v_{n+k}} * X_{n+k-1}(z_{n+k-1}, v_{n+k-1})]^2|X_n(z_n, v_n)\Big) + 2E\Big(\alpha_{v_{n+k}} * X_{n+k-1}(z_{n+k-1}, v_{n+k-1}) \\
&\quad + \varepsilon_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})|X_n(z_n, v_n)\Big) \\
&\quad + E(\varepsilon^2_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})|X_n(z_n, v_n)) \\
&= E\Big((\alpha_{v_{n+k}} * X_{n+k-1}(z_{n+k-1}, v_{n+k-1}))^2|X_n(z_n, v_n)\Big) \\
&\quad + 2E\Big(\alpha_{v_{n+k}} * X_{n+k-1}(z_{n+k-1}, v_{n+k-1})|X_n(z_n, v_n)\Big) + E(\varepsilon_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})) \\
&\quad + E(\varepsilon^2_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})).
\end{aligned}$$

Since $E(\alpha * X|Y) = \alpha E(X|Y)$, $E\Big((\alpha * X)^2|Y\Big) = \alpha^2 E(X^2|Y) + \alpha(1+\alpha)E(X|Y)$, $\sigma^2_{n+k|n} = E(X^2_{n+k}(z_{n+k}, v_{n+k})|X_n(z_n, v_n)) - \Big(E(X_{n+k}(z_{n+k}, v_{n+k})|X_n(z_n, v_n))\Big)^2$ and $\sigma^2_{\varepsilon_{n+k}} = E(\varepsilon^2_{n+k}|X_n) - \Big(E(\varepsilon_{n+k}|X_n)\Big)^2$, we obtain

$$\begin{aligned}
\sigma^2_{n+k|n} &= \alpha^2_{v_{n+k}} E\Big(X^2_{n+k-1}(z_{n+k-1}, v_{n+k-1})|X_n(z_n, v_n)\Big) + \alpha_{v_{n+k}}(1 + \alpha_{v_{n+k}})E(X_{n+k-1}(z_{n+k-1}, v_{n+k-1})|X_n(z_n, v_n)) \\
&\quad + 2\alpha_{v_{n+k}} E(X_{n+k-1}(z_{n+k-1}, v_{n+k-1})|X_n(z_n, v_n))E(\varepsilon_{n+k}(z_{n+k-1}, v_{n+k-1})) \\
&\quad + E(\varepsilon^2_{n+k}(z_{n+k-1}, z_{n+k}, v_{n+k-1}, v_{n+k})) - \mu^2_{n+k|n} \\
&= \alpha^2_{v_{n+k}}(\sigma^2_{n+k-1|n} + \mu^2_{n+k-1|n}) + \alpha_{v_{n+k}}(1 + \alpha_{v_{n+k}})\mu_{n+k-1|n} + 2\alpha_{v_{n+k}}\mu_{n+k-1|n}\mu_{\varepsilon_{n+k}} + \sigma^2_{\varepsilon_{n+k}} + \mu^2_{\varepsilon_{n+k}} - \mu_{n+k|n} \\
&= \alpha^2_{v_{n+k}}\sigma^2_{n+k-1|n} + \alpha_{v_{n+k}}(1 + \alpha_{v_{n+k}})\mu_{n+k-1|n} + \sigma^2_{\varepsilon_{n+k}} + (\alpha_{v_{n+k}}\mu_{n+k-1|n} + \mu_{\varepsilon_{n+k}})^2 - \mu^2_{n+k|n}.
\end{aligned}$$

Since $\mu_{n+k|n} = \alpha_{v_{n+k}}\mu_{n+k-1|n} + \mu_{\varepsilon_{n+k}}$, the following recurrence relation can be derived.

$$\begin{aligned}
\sigma^2_{n+k|n} &= \alpha^2_{v_{n+k}}\sigma^2_{n+k-1|n} + \alpha_{v_{n+k}}(1 + \alpha_{v_{n+k}})\mu_{n+k-1|n} + \sigma^2_{\varepsilon_{n+k}} + \mu^2_{n+k|n} - \mu^2_{n+k|n} \\
&= \alpha^2_{v_{n+k}}\sigma^2_{n+k-1|n} + \alpha_{v_{n+k}}(1 + \alpha_{v_{n+k}})\mu_{n+k-1|n} + \sigma^2_{\varepsilon_{n+k}}.
\end{aligned}$$

Applying the previous equation $k-1$ times, we obtain

$$\begin{aligned}
\sigma^2_{n+k|n} &= \alpha^2_{v_{n+k}}(\alpha^2_{v_{n+k-1}}\sigma^2_{n+k-2|n} + \alpha_{v_{n+k-1}}(1+\alpha_{v_{n+k-1}})\mu_{n+k-2} + \sigma^2_{\varepsilon_{n+k-1}}) + \alpha_{v_{n+k}}(1+\alpha_{v_{n+k}})\mu_{n+k-1|n} + \sigma^2_{\varepsilon_{n+k}} \\
&= \alpha^2_{v_{n+k}}\alpha^2_{v_{n+k-1}}(\alpha^2_{v_{n+k-2}}\sigma^2_{n+k-3|n} + \alpha_{v_{n+k-2}}(1+\alpha_{v_{n+k-2}})\mu_{n+k-3|n} + \sigma^2_{\varepsilon_{n+k-2}}) \\
&\quad + \alpha^2_{v_{n+k}}\alpha_{v_{n+k-1}}(1+\alpha_{v_{n+k-1}})\mu_{n+k-2|n} + \alpha^2_{v_{n+k}}\sigma^2_{\varepsilon_{n+k-1}} + \alpha_{v_{n+k}}(1+\alpha_{v_{n+k}})\mu_{n+k-1|n} + \sigma^2_{\varepsilon_{n+k}} \\
&= \alpha^2_{v_{n+k}}\alpha^2_{v_{n+k-1}}\dots\alpha^2_{v_{n+1}}\sigma^2_{n|n} + \sum_{l=0}^{k-2}\prod_{p=0}^{l-2}\alpha^2_{v_{n+k-p}}\alpha_{v_{n+k-l}}(1+\alpha_{v_{n+k-l}})\mu_{n+k-l-1|n} \\
&\quad + \alpha_{v_{n+k}}(1+\alpha_{v_{n+k}})\mu_{n+k-1|n} + \sigma^2_{\varepsilon_{n+k}} + \sum_{l=1}^{k-1}\prod_{p=0}^{l-1}\alpha^2_{v_{n+k-p}}\sigma^2_{\varepsilon_{n+k-l}}.
\end{aligned}$$

Given that $\sigma^2_{n|n} = 0$, and using the variance result for the random variable $\varepsilon_{n+k-l}, l \in \{0, 1, where \dots, k-1\}$, from equation (9), it follows that

$$\begin{aligned}
\sigma^2_{n+k|n} &= \sum_{l=0}^{k-2}\prod_{p=0}^{l-2}\alpha^2_{v_{n+k-p}}\alpha_{v_{n+k-l}}(1+\alpha_{v_{n+k-l}})\mu_{n+k-l-1|n} + \alpha_{v_{n+k}}(1+\alpha_{v_{n+k}})\mu_{n+k-1|n} + \mu_{z_{n+k}}(1+\mu_{z_{n+k}}) \\
&\quad - \alpha_{v_{n+k}}\mu_{z_{n+k-1}}(1+2\alpha_{v_{n+k}} + \alpha_{v_{n+k}}\mu_{z_{n+k-1}}) + \sum_{l=1}^{k-1}\prod_{p=0}^{l-1}\alpha^2_{v_{n+k-p}}\Big(\mu_{z_{n+k-l}}(1+\mu_{z_{n+k-l}}) \\
&\quad - \alpha_{v_{n+k-l}}\mu_{z_{n+k-l}}(1+2\alpha_{v_{n+k-l}} + \alpha_{v_{n+k-l}}\mu_{z_{n+k-l-1}})\Big) \\
&= \sum_{l=0}^{k-2}\prod_{p=0}^{l-2}\alpha^2_{v_{n+k-p}}\alpha_{v_{n+k-l}}(1+\alpha_{v_{n+k-l}})\Big[\alpha_{v_{n+k-l-1}}\dots\alpha_{v_{n+1}}X_n(z_n, v_n) + \mu_{z_{n+k-l-1}} - \alpha_{v_{n+k-l-1}}\dots\alpha_{v_{n+1}}\mu_{z_n}\Big] \\
&\quad + \alpha_{v_{n+k}}(1+\alpha_{v_{n+k}})\Big[\alpha_{v_{n+k-1}}\dots\alpha_{v_{n+1}}X_n(z_n, v_n) + \mu_{z_{n+k-1}} - \alpha_{v_{n+k-1}}\dots\alpha_{v_{n+1}}\mu_{z_n}\Big] + \mu_{z_{n+k}}(1+\mu_{z_{n+k}}) \\
&\quad - \alpha_{v_{n+k}}\mu_{z_{n+k-1}}(1+2\alpha_{v_{n+k}} + \alpha_{v_{n+k}}\mu_{z_{n+k-1}}) \\
&\quad + \sum_{l=1}^{k-1}\prod_{p=0}^{l-1}\alpha^2_{v_{n+k-p}}\Big(\mu_{z_{n+k-l}}(1+\mu_{z_{n+k-l}}) - \alpha_{v_{n+k-l}}\mu_{z_{n+k-l}}(1+2\alpha_{v_{n+k-l}} + \alpha_{v_{n+k-l}}\mu_{z_{n+k-l}})\Big) \\
&= \sum_{l=0}^{k-2}\prod_{p=0}^{l-2}\alpha^2_{v_{n+k-p}}\alpha_{v_{n+k-l}}(1+\alpha_{v_{n+k-l}})\Big[\alpha_{v_{n+k-l-1}}\dots\alpha_{v_{n+1}}X_n(z_n, v_n) + \mu_{z_{n+k-l-1}} - \alpha_{v_{n+k-l-1}}\dots\alpha_{v_{n+1}}\mu_{z_n}\Big] \\
&\quad + \alpha_{v_{n+k}}(1+\alpha_{v_n+k})\alpha_{v_{n+k-1}}\dots\alpha_{v_{n+1}}X_n(z_n, v_n) - \alpha_{v_{n+k}}(1+\alpha_{v_n+k})\alpha_{v_{n+k-1}}\dots\alpha_{v_{n+1}}\mu_{z_n} \\
&\quad + \alpha_{v_{n+k}}(1+\alpha_{v_{n+k}})\mu_{z_{n+k-1}}(1-\alpha^2_{v_{n+k}}\mu_{z_{n+k-1}}(1+\mu_{z_{n+k-1}})) + \mu_{z_{n+k}}(1+\mu_{z_{n+k}}) \\
&\quad + \sum_{l=1}^{k-1}\prod_{p=0}^{l-1}\alpha^2_{v_{n+k-p}}\Big(\mu_{z_{n+k-l}}(1+\mu_{z_{n+k-l}}) - \alpha_{v_{n+k-l}}\mu_{z_{n+k-l}}(1+2\alpha_{v_{n+k-l}} + \alpha_{v_{n+k-l}}\mu_{z_{n+k-l}})\Big). \qquad (10)
\end{aligned}$$

□

For all i, it holds that $|\alpha_i| < 1$. Let $\alpha_n = \max\{\alpha_{v_{n+k}}, \dots, \alpha_{v_{n+1}}\}$. This implies that

$$0 \le \alpha_{v_{n+k}}\dots\alpha_{v_{n+1}} \le (\alpha_n)^k \to 0, \quad k \to \infty,$$

Thus, for sufficiently large k, we have $E(X_{n+k}|X_n) \approx \mu_{z_{n+k}} = E(X_{n+k})$. Finnaly, since $\mu_{z_{n+k-1}} = E(X_{n+k-1}) = E(\alpha_{v_{n+k-1}} * X_{n+k-2} + \varepsilon_{n+k-1}) = \alpha_{v_{n+k-1}}\mu_{z_{n+k-2}} + \mu_{\varepsilon_{n+k-1}}$, and by using recurrence relations based on (10), it follows that $D(X_{n+k}|X_n) \approx \mu_{z_{n+k}}(1+\mu_{z_{n+k}}) = D(X_{n+k})$.

4. Understanding model control through two distinct random environment processes

In introducing our model based on the RrNGINAR(1) framework presented by [11], we propose a new random environmental process denoted as $\{V_n\}$. This process is independent of $\{Z_n\}$ and impacts the thinning operator α, thereby illustrating the dependence among the elements of the process $\{X_n\}$.

This modification of the RrNGINAR(1) model enhances its adaptability to data, making it more applicable in real-world situations. Initially, it is crucial to determine the realizations $\{z_n\}$ and $\{v_n\}$. For this purpose, let's consider a dataset $x_1, x_2, \ldots, x_N$ that we aim to analyze using the introduced model. The first question that arises is how to establish these realizations of the controlling processes based on this data. Since the value of z_n dictates the distribution of the random variable X_n, it is anticipated that values x_i, where $i \in \{1, 2, \ldots, N\}$, that are close to one another correspond to the same state z_i. Accordingly, we cluster the data $x_1, x_2, \ldots, x_N$ into r_1 different clusters, each corresponding to a specific state, as outlined in the work of [11]. We use K-Means Clustering Algorithm, as it is explained in [6].

The values from the set E_{r_1} are crucial in linking $X_n(z_n, v_n)$ to the corresponding distribution parameter μ_{z_n}. Consequently, the specific numbering of the clusters is irrelevant.

A new question emerges regarding how to determine the realizations $\{v_n\}$. Our approach is to calculate the corresponding autocorrelations $\rho^{(i)}$, where $i \in \{1, 2, \ldots, N-p\}$, for subsamples of size p derived from the whole sample of size N. The subsamples are drawn in a sequential manner, starting with the first p elements. The second subsample consists of elements $2, 3, \ldots, p+1$, and this process continues accordingly. Using this method, we have that

$$\begin{aligned}
\rho^{(1)} &= \frac{\sum_{i=1}^{p}(x_i - \overline{x}_p)(x_{i+1} - \overline{x}_p)}{\sum_{i=1}^{p}(x_i - \overline{x}_p)^2}, \\
\rho^{(2)} &= \frac{\sum_{i=2}^{p+1}(x_i - \overline{x}_p)(x_{i+1} - \overline{x}_p)}{\sum_{i=2}^{p+1}(x_i - \overline{x}_p)^2}, \\
&\ldots \\
\rho^{(N-p)} &= \frac{\sum_{i=N-p}^{N-1}(x_i - \overline{x}_p)(x_{i+1} - \overline{x}_p)}{\sum_{i=N-p}^{N-1}(x_i - \overline{x}_p)^2}.
\end{aligned}$$

By using the clustering procedure, we consider the values $\rho^{(1)}, \rho^{(2)}, \ldots, \rho^{(N-p)}$ that are close to each other to correspond to the same state, denoted as v_i. Furthermore, these values can be approximated by a steady curve, indicating that they are unlikely to be distributed in a "chaotic" manner on the plot. This behavior arises because we select subsamples successively, effectively "sliding" through the observed data step by step. As a result, every two consecutive subsamples contain the maximum possible number of overlapping elements, specifically having $p-1$ common elements. For instance, we might find that $\rho^{(1)} \approx \rho^{(2)}$, while it is possible for $\rho^{(1)}$ and $\rho^{(50)}$ to differ significantly. Now, we can group the obtained values $\rho^{(1)}, \rho^{(2)}, \ldots, \rho^{(N-p)}$ into r_2 distinct clusters. Values belonging to the same cluster correspond to the same state. Importantly, the specific numbering of the clusters is not significant in this context.

Now, the question arises of how to determine the value of p. Intuitively, it is evident that as p increases, we have fewer disjoint subsamples and, consequently, fewer distinctly different values of $\rho^{(i)}$ for $i \in \{1, 2, \ldots, N-p\}$. Conversely, if p is small, we may have a larger number of different $\rho^{(i)}$ values. Therefore, the sample size N influences the size of the subsample p, and it is clear that the number of states r_2 depends on p.

The size of the subsample p can be determined using the formula $p = 5 \cdot \log N$. This approach is commonly used in Sample Theory when working with large datasets to establish an appropriate subsample size for further analysis and data processing. From the properties of the logarithmic function, it is clear that it grows more rapidly when the values of N are small and then tapers off as N increases. Consequently, the logarithmic function has the property of "compressing" large values of N into smaller values of p while mapping smaller values of N onto larger values of p. This characteristic is useful for representing a wide range of numbers.

It should be noted that we will later derive the corresponding $\alpha^{(i)}$ values for $i \in \{1, 2, \ldots, N-p\}$ based on the second part of Theorem 3.3.

However, it is important to emphasize that due to the negative binomial thinning operator, the following condition for α is mandatory:

$$0 \leq \alpha \leq \min\left\{\frac{\mu_l}{1+\mu_k},\, l,k \in E_{r_1}\right\}, \quad \alpha \in \{\alpha_1, \alpha_2, \ldots, \alpha_{r_2}\}.$$

5. Parameter estimation

In this section, we will concentrate on estimating the model's parameters, which is a crucial step in model development. This process typically involves using mathematical and statistical methods to optimize and estimate parameters, ensuring that the model aligns well with the data. Accurate parameter estimates enable us to customize the model according to the data and draw meaningful conclusions.

In what follows, we explore two approaches to parameter estimation for the proposed model: the Yule-Walker (YW) method and the Conditional Maximum Likelihood (CML) method. The YW method is particularly well-suited for scenarios where rapid estimation is required and the underlying dependence structure can be effectively captured through autocorrelations. It is computationally simple and performs well when the latent environment remains relatively stable over time. However, its applicability may be limited in more complex, non-stationary settings. On the other hand, the CML method is more flexible and robust in the presence of time-varying dynamics and latent state processes, as it explicitly incorporates the influence of the random environment. While CML is computationally more intensive, it often yields more accurate and consistent parameter estimates in models with stochastic environmental control. In practice, the choice between these two methods depends on the trade-off between computational efficiency and the complexity of the data-generating process.

5.1. Yule-Walker estimation

Let's examine a segment of the process where there is no change in state. For integers $k, n \in \mathbb{N}$ such that $z_k \neq i$, $z_{k+1} = z_{k+2} = \cdots = z_n = i$, and $z_{n+1} \neq i$, we can observe the subsample:

$$X_{k+1}(i, v_{k+1}), X_{k+2}(i, v_{k+2}), \ldots, X_n(i, v_n).$$

Since these elements all correspond to the same state i, this subsample can be considered a sample from the NGINAR(1) process introduced in [13], with an expectation denoted as μ_i. Because the NGINAR(1) process is stationary, the corresponding sample covariance is also strictly stationary, which allows for accurate estimation. For further details, we refer to the work of [11], where estimators based on the subsample S_k corresponding to k are defined. They are given by the following expressions:

$$\widehat{\mu_k} = \frac{1}{n_k}\sum_{i\in I_k} X_i(k, v_i), \quad \widehat{\gamma_0}^{(k)} = \frac{1}{n_k}\sum_{i\in I_k}(X_i(k, v_i) - \widehat{\mu_k})^2,$$

$$\widehat{\gamma_1}^{(k)} = \frac{1}{n_k}\sum_{i,i+1\subseteq I_k}(X_{i+1}(k, v_{i+1}) - \widehat{\mu_k})(X_i(k, v_i) - \widehat{\mu_k}),$$

where

$$I_k = \{i \in 1, 2, \ldots, N | z_i = k\},$$

represent the index set of process elements corresponding to value k of VC-process. Also, we have that

$$\bigcup_{k=1}^{r_1} I_k = \{1, 2, \ldots, N\}, \quad |I_k| = n_k, \quad n_1 + n_2 + \cdots + n_{r_1} = N,$$

$$S_k = \Big(X_{k_1}(k, v_{k_1}), X_{k_2}(k, v_{k_2}), \ldots, X_{k_{n_k}}(k, v_{k_{n_k}})\Big), \quad k_i \in I_k, k_i < k_{i+1}, \forall i \in \{1, 2, \ldots, n_k - 1\}.$$

In our discussion, let S_k represent the subsample derived from the initial sample. This subsample includes all elements associated with a specific circumstance (a possible value of the random state) k and excludes elements linked to any other circumstances.

In the work of [11], the strong consistency of the estimators was established (see Theorem 5). This property is essential for ensuring the reliability of the estimators, which is particularly important when making significant decisions and drawing conclusions.

In this section, we will focus on the parameters of the thinning operator. Therefore, employing the estimating approach described in the previous section, we have

$$\widehat{\rho}^{(j)} = \frac{\sum_{i=j}^{p+j-1}(X_i - \overline{X_p})(X_{i+1} - \overline{X_p})}{\sum_{i=j}^{p+j-1}(X_i - \overline{X_p})^2},$$

where $\overline{X_p} = \frac{1}{p}\sum_{i=j}^{p+j-1} X_i$. For an arbitrary $k \in \{1, 2, \ldots, r_2\}$, let

$$J_k = \{i \in \{1, 2, \ldots, N-p\} | v_i = k\},$$

denote the index set of all the process elements that are in the same state, i.e. which are controlled by the same value of CC-process. It follows that

$$\bigcup_{k=1}^{r_2} J_k = \{1, 2, \ldots, N-p\}, \quad |J_k| = m_k, \quad m_1 + m_2 + \cdots + m_{r_2} = N - p.$$

Next define

$$F_k = \left(\alpha^{(k_1)}(k), \alpha^{(k_2)}(k), \ldots, \alpha^{(k_{n_k})}(k)\right),$$

where $k_i \in J_k, k_i < k_{i+1}, \forall i \in \{1, 2, \ldots, m_k - 1\}$. Thus, F_k contains all $\alpha^{(j)}$ corresponding to state k and none from other states, formed by the transformation of the correlations associated with the k-th cluster. It can be noticed that F_k consists of all maximal subsets of $\alpha^{(j)}$ corresponding to state k. Let these subsets be denoted as $F_{k,1}, F_{k,2}, \ldots, F_{k,i_k}$, where i_k is the total number of such subsets. Furthermore, let $J_{k,l} = \{i \in \{1, 2, \ldots, N-p\} | \alpha^{(i)}(v_i) \in F_{k,l}\}, |J_{k,l}| = m_{k,l}$ and note that $m_{k,1} + m_{k,2} + \cdots + m_{k,i_k} = m_k$. The estimates obtained from $F_{k,l}$ are given by

$$\widehat{\alpha}_{k,l} = \frac{1}{m_{k,l}} \sum_{i \in J_{k,l}} \alpha^{(i)}(k),$$

which are strongly consistent. Therefore, we have $P(\lim_{m_{k,l}\to\infty} \widehat{\alpha}_{k,l} = \alpha_k) = 1$. This implies that, $\lim_{m_{k,l}\to\infty} \widehat{\alpha}_{k,l} = \alpha_k$ holds almost everywhere except on the set $\Omega_{k,l}$, where $P(\Omega_{k,l}) = 0$. Consequently, we conclude that $\widehat{\alpha}_{k,l} = \alpha_k + o(m_{k,l}), m_{k,l} \to \infty$ almost everywhere except on the set $\Omega_{k,l}$.

The estimate based on the entire F_k is represented by the following expression.

$$\begin{aligned} \widehat{\alpha}_k &= \frac{1}{m_k} \sum_{i \in J_k} \alpha^{(i)}(k) = \frac{1}{m_k} \sum_{l=1}^{i_k} \sum_{i \in J_{k,l}} \alpha^{(i)}(k) \\ &= \sum_{l=1}^{i_k} \frac{m_{k,l}}{m_k} \frac{1}{m_{k,l}} \sum_{i \in J_{k,l}} \alpha^{(i)}(k) = \sum_{l=1}^{i_k} \frac{m_{k,l}}{m_k} \widehat{\alpha}_{k,l}. \end{aligned} \tag{11}$$

Since $\frac{\sum_{l=1}^{i_k} m_{k,l}}{m_k} = 1$ and for all $l \in \{1, 2, \ldots, i_k\}$, $\lim_{m_{k,l}\to\infty} \frac{m_{k,l}}{m_k} < \infty$, it follows that $\widehat{\alpha}_k \to \alpha_k$ and $m_{k,l} \to \infty$ for all $i \in \{1, 2, \ldots, i_k\}$ everywhere except on the set $\Omega_k = \bigcup_{l=1}^{i_k} \Omega_{k,l}$.

Based on the inequality $P(\Omega_k) = P\left(\bigcup_{l=1}^{i_k} \Omega_{k,l}\right) \leq \sum_{l=1}^{i_k} P(\Omega_{k,l})$, and the property of non-negativity of probability, we can conclude that $P(\Omega_k) = 0$. Therefore, we have:

$$P(\widehat{\alpha}_k \to \alpha_k, m_{k,i} \to \infty, \forall i \in \{1, 2, \ldots, i_k\}) = 1.$$

We need to demonstrate strong consistency as m_k approaches infinity, as strong consistency requires convergence when the sample size tends to infinity. Accordingly, we will formulate the following theorem.

Theorem 5.1. *The estimator $\widehat{\alpha_k}$ defined by (11) is strongly consistent.*

Proof. We need to demonstrate that $P(\widehat{\alpha_k} \to \alpha_k, m_k \to \infty) = 1$. We have that $m_k = m_{k,1} + m_{k,2} + \cdots + m_{k,i_k}$. As m_k approaches infinity, if all the sizes of the subsamples $m_{k,j}$ were finite, then m_k would also be finite, which leads to a contradiction. Therefore, for m_k to approach infinity, at least one of the $m_{k,j}$ must also approach infinity.

Let us assume that for some $d \in \{1, 2, \ldots, k\}$, the following holds:

$$\begin{aligned} &m_{k,l} \to \infty, \quad \forall l \in \{1, 2, \ldots, d\}, \\ &m_{k,j} \to c_j < \infty, \quad \forall j \in \{d+1, d+2, \ldots, i_k\}. \end{aligned} \tag{12}$$

If this is not the case, we can prenumerate them. Now,

$$\widehat{\alpha_k} = \sum_{l=1}^{i_k} \frac{m_{k,l}}{m_k} \widehat{\alpha}_{k,l} = \sum_{l=1}^{d} \frac{m_{k,l}}{m_k} \widehat{\alpha}_{k,l} + \sum_{l=d+1}^{i_k} \frac{m_{k,l}}{m_k} \widehat{\alpha}_{k,l}.$$

Based on (12), we have that $\lim_{m_k \to \infty} \frac{m_{l,j}}{m_k} = 0$, for $j \in \{d+1, \ldots, i_k\}$. Therefore, it follows that

$$\lim_{m_k \to \infty} \widehat{\alpha_k} = \lim_{m_k \to \infty} \sum_{l=1}^{d} \frac{m_{k,l}}{m_k} \widehat{\alpha}_{k,l},$$

Here m_k is defined as $m_k = m_{k,1} + m_{k,2} + \cdots + m_{k,i_k} = \sum_{l=1}^{d} m_{k,l} + \sum_{j=d+1}^{i_k} m_{k,j} \to \sum_{l=1}^{d} m_{k,l} + \sum_{j=d+1}^{i_k} c_j$, when $m_k \to \infty$. Since the sum $\sum_{j=d+1}^{i_k} c_j$ is finite, it becomes negligible compared to $m_{k,l}$, for $l \in \{1, 2, \ldots, d\}$. This allows us to simplify m_k to $m_k = m_{k,1} + \cdots + m_{k,d}$ when considering the limit as $m_k \to \infty$.

Based on (12), we find that

$$(m_k \to \infty) \Leftrightarrow (m_{k,l} \to \infty, \forall l \in \{1, 2, \ldots, d\}).$$

Thus we have, $\lim_{m_k \to \infty} \widehat{\alpha_k} = \lim_{m_{k,l} \to \infty, \forall l \in \{1,2,\ldots,d\}} \widehat{\alpha_k} = \alpha_k$. Therefore, it follows that $\widehat{\alpha_k} \to \alpha_k, \quad m_k \to \infty$ everywhere, except on the set $\Omega_k = \bigcup_{l=1}^{i_k} \Omega_{k,l}$. This can be expressed as $P(\widehat{\alpha_k} \to \alpha_k, m_k \to \infty) = 1$, indicating its strong consistency. □

5.2. Conditional Maximum Likelihood Estimation

Based on a random sample of size N from the V-CCNGINAR(1) process, denoted as $X_1(z_1, v_1), X_2(z_2, v_2), \ldots, X_N(z_N, v_N)$, we define a log-likelihood function that can be utilized for the parametric estimation of unknown model parameters $\mu_1, \mu_2, \ldots, \mu_{r_1}, \alpha_1, \alpha_2, \ldots, \alpha_{r_2}$. We assume that we have already obtained the realized values x_0, z_0, and v_0. These values will enhance the clarity of the joint density without affecting the quality of the estimation procedure. Utilizing Theorem 2.3, we derive the joint log-likelihood function as follows.

$$\begin{aligned} logL &= logL(x_1, z_1, \ldots, x_N, z_N | \mu_1, \ldots, \mu_{r_1}, \alpha_1, \ldots, \alpha_{r_2}) \\ &= \sum_{i=2}^{N} logP(X_i(z_i, v_i) = x_i | X_{i-1}(z_{i-1}, v_{i-1}) = x_{i-1}) p_{i-1,i} p^*_{i-1,i}, \end{aligned}$$

where $p_{i-1,i} = P(Z_i = z_i|Z_{i-1} = z_{i-1})$ and $p*_{i-1,i} = P(V_i = v_i|V_{i-1} = v_{i-1})$. Thus, we have

$$\begin{aligned} logL &= \sum_{i=2}^{N} logp_{i-1,i}p*_{i-1,i}\, P(X_i(z_i, v_i) = x_i|X_{i-1}(z_{i-1}, v_{i-1}) = x_{i-1}) \\ &= \sum_{i=2}^{N} logp_{i-1,i}p*_{i-1,i}\, P(\alpha_{v_i} * X_{i-1}(z_{i-1}, v_{i-1}) \\ &\qquad +\varepsilon_i(z_{i-1}, z_i, v_{i-1}, v_i) = x_i|X_{i-1}(z_{i-1}, v_{i-1}) = x_{i-1}) \\ &= \sum_{i=2}^{N} logp_{i-1,i}p*_{i-1,i}\Big\{\sum_{k=0}^{x_i} P(\alpha_{v_i} * X_{i-1}(z_{i-1}, v_{i-1}) = k)P(\varepsilon_i(z_{i-1}, z_i, v_{i-1}, v_i) = x_i - k)\cdot I_{\{x_{i-1}\neq 0\}} \\ &\qquad +P(\varepsilon_i(z_{i-1}, z_i, v_{i-1}, v_i) = x_i)\cdot I_{\{x_{i-1}=0\}}\Big\} \end{aligned}$$

Based on the distribution of the random variable $\varepsilon_i(z_{i-1}, z_i, v_{i-1}, v_i)$,

$$P(\varepsilon_i(z_{i-1}, z_i, v_{i-1}, v_i) = x) = \Big(1 - \frac{\alpha_{v_i}\mu_{z_{i-1}}}{\mu_{z_i} - \alpha_{v_i}}\Big)\cdot\frac{\mu_{z_i}^x}{(1+\mu_{z_i})^{x+1}} + \frac{\alpha_{v_i}\mu_{z_{i-1}}}{\mu_{z_i} - \alpha_{v_i}}\cdot\frac{\alpha_{v_i}^x}{(1+\alpha_{v_i})^{x+1}},$$

it follows that

$$\begin{aligned} logL &= \sum_{i=2}^{N} logp_{i-1,i}p*_{i-1,i}\Big\{\sum_{k=0}^{x_i} D_i[(1-C_i)C(x_i-k) + C_iC(x_i-k)]I_{\{x_{i-1}\neq 0\}} \\ &\qquad +[(1-C_i)C(x_i) + C_iC(x_i)]I_{\{x_{i-1}=0\}}\Big\}, \end{aligned}$$

where we used the following notation $D_i = \binom{x_{i-1}+k-1}{x_{i-1}-1}\frac{\alpha_{v_i}^k}{(1+\alpha_{v_i})^{k+x_{i-1}}}$, $C_i = \frac{\alpha_{v_i}\mu_{z_{i-1}}}{\mu_{z_i}-\alpha_{v_i}}$, $C(x) = \frac{\mu_{z_i}^x}{(1+\mu_{z_i})^{x+1}}$.

Now, using numerical algorithms, as implemented in known statistical software packages, we can obtain the conditional maximum likelihood estimations by maximizing this function.

5.3. Model simulations

We will now evaluate the quality of the Yule-Walker (YW) estimates and the conditional maximum likelihood (CML) estimates, as defined in the previous subsections. Our goal is to demonstrate that the estimates of the unknown parameters converge to their true values as the sample size increases. To accomplish this, we simulated samples of various sizes from the processes under study and calculated the resulting estimates for these samples. Specifically, we simulated a sample of size 10,000 across 100 replications simultaneously. To construct the observed model, we first simulated two random environment processes and then used these to generate the corresponding INAR process. We have chosen what we believe to be the most likely practical case for this analysis, assuming that the random environment processes $\{Z_n\}$ and $\{V_n\}$ have state spaces $E_3 = \{1, 2, 3\}$ and $E_2 = \{1, 2\}$, respectively. To explore how the parameter values influence the behavior of the estimates, we will consider the following cases.

In case (a), the true mean vector is given by $\mu = (8, 9, 10)'$. According to Theorem 3.1, the largest possible value of the parameter α_j for $j \in \{1, 2\}$ is 0.73. Therefore, we have chosen $\alpha = (0.5, 0.7)'$. We assume that the choice of the initial random state is nearly fair, so we selected $p_{vec}^Z = (0.35, 0.35, 0.3)'$ and $p_{vec}^V = (0.45, 0.55)'$. The transition probability matrices are highly significant as they determine the structure of the random environment processes, which in turn affects the structure of the V-CCINAR(1) process. We define the transition probability matrices as follows:

$p_{mat}^Z = \begin{bmatrix} 0.8 & 0.1 & 0.1 \\ 0.05 & 0.8 & 0.15 \\ 0.1 & 0.05 & 0.85 \end{bmatrix}$ and $p_{mat}^V = \begin{bmatrix} 0.98 & 0.02 \\ 0.01 & 0.99 \end{bmatrix}$. In these matrices, the values on the diagonal represent the probabilities of remaining in the same state. We observe that these values are significantly higher than

the probabilities of transitioning to one of the other states in the first process, and to the second state in the second process. The transition probability matrices set the dynamic nature of the process itself, influencing how frequently it transitions between passive and active states. Notably, the main diagonal of the matrix p^V_{mat} contains large values, while the off-diagonal elements are relatively small. It is important to note that the process $\{V_n\}$ controls the correlations. In cases where the INAR process is non-stationary, there can be periods during which the correlation remains consistent over an extended time interval. When a regime change occurs, the correlation shifts to a new level and stabilizes as long as that regime continues.

In scenario (b), we will use the same expectation vector μ and the same α vector as in the previous case. However, we will assume that at the initial moment, the second state of the process $\{Z_n\}$ is more likely than the others, according to the vector $p^Z_{vec} = (0.2, 0.6, 0.2)'$. The chosen transition matrix is $p^Z_{mat} = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.4 & 0.5 \end{bmatrix}$. In this case, the highest probabilities are associated with remaining in the same state, while the probabilities of transitioning from the first state to the third state, or from the third state to the first state, are the lowest. For the second control process $\{V_n\}$, the initial vector and transition matrix remain the same as in the previous case.

Table 1: YW estimates for case (a)

n_1	$\hat{\mu}^1_{YW}$	$\hat{\mu}^2_{YW}$	$\hat{\mu}^3_{YW}$	$\hat{\alpha}^1_{YW}$	$\hat{\alpha}^2_{YW}$
100	7.832959	9.403668	10.254088	0.5096555	0.6825996
St. errors	3.192264	3.63779	2.853453	0.1201141	0.1140894
500	7.728912	9.129296	9.871219	0.4901601	0.7273919
St. errors	1.339548	1.588186	1.548956	0.06870043	0.05598686
1000	7.893795	9.128044	9.989333	0.497689	0.7191186
St. errors	0.9330041	1.1431919	1.0660006	0.05798667	0.03603559
5000	7.992915	9.072429	10.003724	0.4972187	0.7226275
St. errors	0.3682673	0.4248365	0.4607977	0.02133199	0.01441725
10000	7.995422	8.986868	9.997408	0.5012686	0.7272925
St. errors	0.3136503	0.2864798	0.3466683	0.01503872	0.01044569

Table 2: CML estimates for case (a)

n_1	$\hat{\mu}^1_{CML}$	$\hat{\mu}^2_{CML}$	$\hat{\mu}^3_{CML}$	$\hat{\alpha}^1_{CML}$	$\hat{\alpha}^2_{CML}$
100	7.444193	8.990206	10.333766	0.5024442	0.7024345
St. errors	1.75748	1.840264	1.64653	0.1252357	0.1134753
500	7.758841	9.099439	9.972805	0.5049372	0.7009727
St. errors	0.2992385	0.2773392	0.3179715	0.0221322	0.01317968
1000	7.854522	9.04394	9.997843	0.5000597	0.7009984
St. errors	0.1535008	0.1425151	0.1565295	0.009022036	0.006835159
5000	7.956966	9.013465	10.00108	0.503569	0.7006194
St. errors	0.02642322	0.01961257	0.02201731	0.001646149	0.00116224
10000	7.997491	9.004679	9.991952	0.5015705	0.6999096
St. errors	0.011842871	0.009217488	0.010822276	0.000710686	0.000656608

Table 3: YW estimates for case (b)

n_1	$\hat{\mu}^1_{YW}$	$\hat{\mu}^2_{YW}$	$\hat{\mu}^3_{YW}$	$\hat{\alpha}^1_{YW}$	$\hat{\alpha}^2_{YW}$
100	8.05715	9.158191	9.972508	0.5232052	0.6989305
St. errors	2.172533	2.167735	2.820093	0.1328363	0.1249328
500	8.154316	9.235489	10.235489	0.505303	0.7293339
St. errors	1.077184	1.114418	1.324761	0.08849577	0.05508949
1000	8.027569	9.015926	10.077965	0.4983323	0.724834
St. errors	0.8090873	0.7161218	0.9964789	0.05778862	0.04084331
5000	8.014484	9.005021	10.013848	0.4986456	0.7238331
St. errors	0.3477684	0.3192316	0.4098634	0.02230396	0.01492933
10000	8.010784	9.024727	10.017258	0.4996013	0.7234465
St. errors	0.2560424	0.2465305	0.3204341	0.01527528	0.01118749

Table 4: CML estimates for case (b)

n_1	$\hat{\mu}^1_{CML}$	$\hat{\mu}^2_{CML}$	$\hat{\mu}^3_{CML}$	$\hat{\alpha}^1_{CML}$	$\hat{\alpha}^2_{CML}$
100	7.596924	9.109163	10.197828	0.5077181	0.7138195
St. errors	1.133419	1.126538	1.196808	0.1387191	0.1154847
500	7.901334	9.018258	10.10529	0.5112666	0.7034844
St. errors	0.1809004	0.1876676	0.1609393	0.02390348	0.01353522
1000	7.959451	8.966761	10.054064	0.4995878	0.7018725
St. errors	0.09340222	0.08344782	0.08557829	0.008417337	0.005909572
5000	7.99624	8.986343	10.01677	0.5022544	0.6995132
St. errors	0.01780595	0.01670531	0.01406533	0.001910923	0.00117593
10000	8.002048	8.999809	10.001591	0.5020619	0.6999207
St. errors	0.008826434	0.008250537	0.007599185	0.000853276	0.000594153

In the previous two cases, we estimated the unknown parameters of the INAR model, specifically α and μ, using Yule-Walker and CML statistics. The parameter estimates were obtained for subsamples of sizes 100, 500, 1000, 5000, and 10000. In this analysis, we utilized 100 realized simulated time series and calculated the corresponding standard errors. As the sample size increases, all estimates converge, and the standard errors decrease toward zero. These results are detailed in the previous tables. However, we did not estimate the transition probabilities, as they are not parameters of the INAR process itself. In the next section, which focuses on modeling a real-life counting process, we will attempt to estimate this transition probability matrix as well, making en effort of capturing a real-time sequence of random counting circumstances.

6. Real data application

To demonstrate the effectiveness and competitiveness of our model in representing real-life situations, we utilized data from the database hosted on the *Forecasting Principles* website (www.forecastingprinciples.com). The time series selected for analysis consisted of monthly fraud offense counts reported in the 12th police car beat in Pittsburgh, covering the period from January 1990 to December 2001. Based on the autocorrelation and partial autocorrelation structures of the series, we determined that employing INAR(1) models was appropriate. The trajectory of the time series is shown in Figure 1. A cluster analysis of the time

series values revealed two distinct states, represented as black and white points in this figure. This finding provided a solid foundation for comparing our model against the RrNGINAR(1) model, where $r = 2$, which was introduced in [11] and served as the inspiration for our development in this article. Furthermore, the cluster analysis of the sub-sample correlations confirmed the existence of two distinct states for our new V-CCNGINAR(1) model. These two different states resulted in two noticeably different α parameters for our model. We used the Maximum Likelihood Estimation method to estimate the unknown parameters for both models. The estimation and prediction results are summarized in Table 6, indicating that our model significantly outperformed the R2NGINAR(1) model in this case.

To further validate our new model, we applied it to a second dataset originally presented in the work of [11], where the RrGINAR(1) model was first introduced. This dataset includes monthly drug offense counts reported in the 27th police car beat in Pittsburgh, covering the same period from January 1990 to December 2001. As illustrated in Figure 2, the time series exhibits three distinct clusters. Both models produced similar estimations of the value states, confirming their ability to identify the underlying structures. To apply our model, we conducted a cluster analysis of sub-samples of realized correlations. The model fitting results indicate that the two correlation parameters in our new model led to slightly better prediction performance compared to the single correlation parameter in the R3NGINAR(1) model, as measured by RMSE statistics. The parameter estimation and forecasting results are summarized in Table 6. In this case, there was no significant difference between the α parameters in the V-CCNGINAR(1) model, similar to what we observed with the fraud data analyzed earlier. This likely explains why our model did not outperform the R3NGINAR(1) model such significantly in this instance; the changes in correlation structure across corresponding correlation random states over time in the drug data were not as pronounced as they were in the first case.

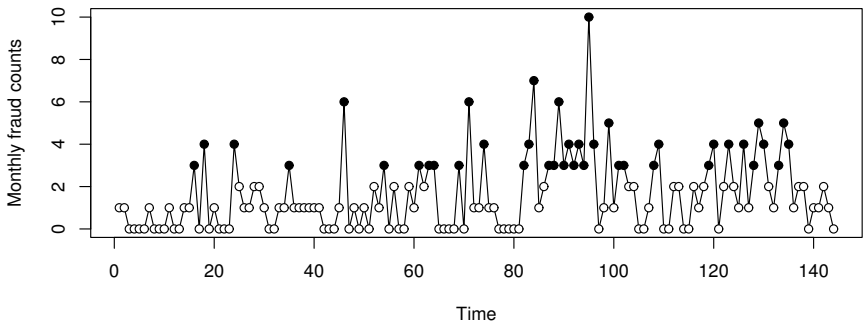

Figure 1: Monthly fraud offense counts reported in the 12th police car beat in Pittsburgh in a period from January 1990 to December 2001

Table 5: Model comparison

	Model	RrNGINAR(1)		V-CCNGINAR(1)	
Observation	Parameter	MLE	RMSE	MLE	RMSE
FRAUDS	$\hat{\mu}_1$	0.655	1.8503	0.343	0.9968
	$\hat{\mu}_2$	3.973		10.016	
	$\hat{\alpha}_1$	0.1317		0.01	
	$\hat{\alpha}_2$			0.1224	
DRUGS	$\hat{\mu}_1$	9.5906	1.6628	9.4956	1.6574
	$\hat{\mu}_2$	0.821		0.8219	
	$\hat{\mu}_3$	23.249		23.3993	
	$\hat{\alpha}_1$	0.028		0.0243	
	$\hat{\alpha}_2$			0.0337	

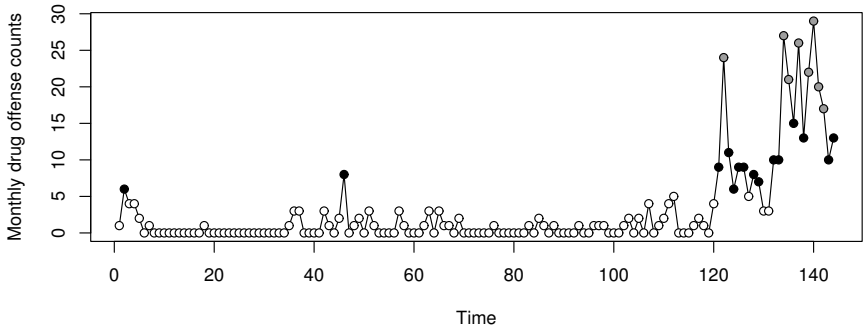

Figure 2: Monthly drug offense counts reported in the 27th police car beat in Pittsburgh in a period from January 1990 to December 2001

7. Concluding remarks

In this manuscript, we introduced a new integer-valued first-order autoregressive model with two independent Markov processes, one controlling the marginal distribution and the other controlling the autocorrelation structure. This structure allows the model to better capture dynamic and non-stationary environments. The fundamental theoretical properties were derived, and the parameters were estimated using the Yule-Walker method and the Conditional Maximum Likelihood approach. The effectiveness of the estimators was confirmed through simulation studies. The model was applied to real-world data and demonstrated strong performance compared to existing approaches, confirming its practical value.

Acknowledgement: The first author acknowledges the grant of MNTRI 451-03-136/2025-03/200124, and the second and third authors acknowledge the grant of MNTRI 451-03-137/2025-03/200124 for carrying out this research.

References

[1] M. A. Al-Osh, E. E. A. A. Aly, *First order autoregressive time series with negative binomial and geometric marginals*, Communication in Statistic- Theory and Methods **21** (1992) 2483–2492.
[2] M. A. Al-Osh, A. A. Alzaid, *First-order integer-valued autoregressive (INAR(1)) process*, Journal of Time Series Analysis **8** (1987) 261–275.
[3] E. E. A. A. Aly, N. Bouzar,*On Some Integer-Valued Autoregressive Moving Average Models*, Journal of Multivariate Analysis **50** (1994) 132–151.
[4] A. A. Alzaid, M. A. Al-Osh, *First-order integer-valued autoregressive (INAR(1)) process:distributional and regression properties*, Statistica Neerlandica **42** (1988) 53–61.
[5] A. A. Alzaid, M. A. Al-Osh, *Some autoregressive moving average processes with generalized Poisson marginal distributions*, Annals of the Institute of Statistical Mathematics **45** (1993) 223–232.
[6] J. A. Hartigan, M. A. Wong, *Algorithm AS 136: A K-Means Clustering Algorithm*, Journal of the Royal Statistical Society, Series C **28** (1979) 100–108.
[7] P.N. Laketa, A. S. Nastić, M. M. Ristić, *Generalized Random Environment INAR Models of Higher Order*, Mediterranean Journal of Mathematics **15** (2018) 9–30.
[8] E. McKenzie, *Some simple models for discrete variate time series*, Water Resources Bulletin **21** (1985) 645–650.
[9] E. McKenzie, *Autoregressive moving-average processes with negative binomial and geometric distributions*, Advances in Applied Probability **18** (1986) 679–705.
[10] A. S. Nastić, M. M. Ristić, H. S. Bakouch, *A combined geometric INAR(p) model based on negative binomial thinning*, Mathematical and Computer Modeling **55** (2012) 1665–1672.
[11] A. S. Nastić, P. N. Laketa, M. M. Ristić, *Random Environment Integer Valued Autoregressive process*, Journal of Time Series Analysis **37** (2016) 267–287.
[12] A. S. Nastić, P. N. Laketa, M. M. Ristić, *Random Environment INAR models of higher order*, RevStat:Statistical Journal **17** (2017) 35–65.
[13] M. M. Ristić, H. S. Bakouch, A. S. Nastić, *A new geometric first-order integer-valued autoregressive (NGINAR(1)) process*, Journal of Statistical Planning and Infesence **139** (2009) 2218–2226.

[14] H. Zheng, I. V. Basawa, S. Datta, *Inference for pth-order random coefficient integer-valued autoregressive processes*, Journal of Time Series Analysis **27** (2006) 411–440.
[15] H. Zheng, I. V. Basawa, S. Datta, *First-order random coefficient integer-valued autoregressive processes*, Journal of Statistical Planning and Infesence **137** (2007) 212–229.