axioms

MDPI

Article

Minification Integer-Valued Split-BREAK Process with Power Series Innovations and Application in Fire Safety Dynamics

Vladica S. Stojanović [1,*], Nikola Mitrović [2], Kristina Tomović [3], Hassan S. Bakouch [4,5] and Shuhrah Alghamdi [6]

[1] Department of Informatics & Computer Sciences, University of Criminal Investigation and Police Studies, 11000 Belgrade, Serbia
[2] Department of Criminalistics, University of Criminal Investigation and Police Studies, 11000 Belgrade, Serbia; nikola.mitrovic@kpu.edu.rs
[3] Department of Mathematics, Faculty of Sciences, University in Kosovska Mitrovica, Lole Ribara 29, 38220 Kosovska Mitrovica, Serbia; kristina.tomovic@pr.ac.rs
[4] Department of Mathematics, College of Science, Qassim University, Buraydah 51452, Saudi Arabia; h.bakouch@qu.edu.sa
[5] Department of Mathematics, Faculty of Science, Tanta University, Tanta 31111, Egypt
[6] Department of Mathematical Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11564, Saudi Arabia; shaalghamdi@pnu.edu.sa
* Correspondence: vladica.stojanovic@kpu.edu.rs

Abstract

This manuscript introduces a new class of count time series models, referred to as the minification integer-valued Split-BREAK (MIN–SB) process. The proposed framework extends the Split-BREAK modeling philosophy to the integer-valued setting and provides a flexible mechanism for capturing rare events, zero inflation, and structural regime changes frequently observed in safety-related data. The main stochastic properties of the MIN–SB process are derived, including stationarity conditions, explicit moment structure, and correlation dynamics. A key theoretical result reveals an implicit hidden Markov structure underlying the observable process, providing a structural explanation for zero clustering observed in rare-event count processes. Parameter estimation is developed using a simulated method of moments (SMM) approach based on zero-related statistics, and the asymptotic properties of the resulting estimators are established. A Monte Carlo simulation study demonstrates favorable finite-sample performance of the proposed estimation procedure. The practical usefulness of the model is illustrated through an empirical application to time series of injuries and fatalities caused by fire accidents in Serbia. The results show that the MIN–SB specification provides a flexible and accurate framework for modeling zero-inflated count processes arising in fire safety dynamics.

Keywords: count time series; integer-valued processes; power series distributions; zero inflation; simulated method of moments; hidden Markov structure; fire safety data

MSC: 62M10; 60G10; 62M20

Check for updates

Academic Editor: Jiujun Zhang

Received: 23 March 2026
Revised: 15 May 2026
Accepted: 21 May 2026
Published: 22 May 2026

Correction Statement: This article has been republished with a minor change. The change does not affect the scientific content of the article and further details are available within the backmatter of the website version of this article.

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Integer-valued time series models play a central role in the analysis of count data arising in numerous applied fields. Among the most widely studied approaches are integer-valued autoregressive (INAR) models, for which an extensive literature has been developed (for some recent ones, see, e.g., [1–7]). Despite their popularity, standard INAR-type

benchmarks often fail to adequately capture empirical features such as rare events, abrupt structural changes, and pronounced zero inflation, which frequently occur in real-world count data. These limitations have motivated the development of alternative modeling frameworks capable of incorporating richer stochastic dynamics. One important direction in this context is represented by time series minification models, originally introduced by Tavares [8,9], and later discussed in [10,11] for the case of processes with continuous, exponential and Weibull marginals. Discrete minification processes were subsequently investigated in [12], followed by several extensions focusing mainly on minification INAR models with geometric marginals [13–15]. More recently, minification INAR processes with Poisson–Lindley marginals have also been considered [16], which highlights the growing interest in stochastic mechanisms based on minimization and reduction effects in count-based time series.

Motivated by this line of research, a new stochastic framework termed the minification integer-valued Split-BREAK (MIN–SB) process is introduced here. The proposed model is conceptually related to the General Split-BREAK (GSB) process introduced by Jovanović et al. [17] and further developed by Stojanović et al. [18] and Ljajko et al. [19], where continuous-valued dynamics driven by persistent fluctuations were considered. In contrast to the additive structure underlying the GSB and the recently proposed integer-valued Split-BREAK (INSB) process [20], the MIN–SB model is based on the minimization principle, which fundamentally changes the resulting stochastic dynamics. This mechanism allows the model to naturally capture abrupt reductions, excess zeros, and complex dependence patterns commonly observed in count data. Furthermore, the innovation series of the MIN–SB process is assumed to follow power series (PS) distributions, a broad and flexible class including classical count models such as binomial, geometric, negative-binomial and Poisson distributions (see, e.g., [21–23]). Within this framework, the proposed process can be interpreted as a branching-type count model with randomly gated immigration. A key theoretical finding of this paper also shows that the observable process admits an implicit hidden Markov structure, providing a structural explanation of its dependence dynamics and offering a new perspective on minification-based count models.

Therefore, the following three aspects can be highlighted as the main contributions of this paper. First, we establish the fundamental probabilistic properties of the MIN–SB process, including stationarity conditions, moment structure, and dependence characteristics. Second, we analyze its marginal distributional behavior, demonstrating its ability to accommodate over-dispersion and zero inflation within a unified stochastic framework. Third, parameter estimation is developed using the simulated method of moments (SMM), and the asymptotic properties of the resulting estimators are derived and examined through Monte Carlo (MC) experiments. The remainder of the paper is organized as follows. Section 2 introduces the preliminary assumptions and definitions, including the negative binomial thinning (NBT) operator and PS distributions. Next, Section 3 investigates the main stochastic properties of the proposed process, including moments, stationarity conditions, overdispersion, and zero-inflation behavior. Parameter estimation procedures based on the SMM estimation methodology are developed in Section 4, while Section 5 presents an MC simulation study illustrating the finite-sample performance of the proposed estimators. In Section 6, the practical relevance of the MIN–SB model is demonstrated through applications to real-world time series of fire-related injuries and fatalities. Finally, Section 7 contains some concluding remarks.

2. Preliminaries

Here we describe some basic concepts and terms, which are of particular importance for further consideration.

2.1. Negative Binomial Thinning

In this subsection, we introduce the negative binomial thinning (NBT) operator and summarize its main probabilistic properties, which will be repeatedly used in the sequel.

Definition 1. *Let* $0 < \alpha < 1$ *and* (W_j), $j = 1, 2, \ldots$ *be a sequence of independent and identically distributed (IID) random variables (RVs) with geometric distribution* $\text{Geom}(\alpha/(1+\alpha))$, *whose probability mass function (PMF) is given by:*

$$p_W(x;\alpha) := \mathbb{P}(W_j = x) = \frac{\alpha^x}{(1+\alpha)^{x+1}}, \qquad x = 0, 1, 2, \ldots. \tag{1}$$

For an integer-valued random variable X, *independent of* (W_j), *the negative binomial thinning operator "*$\circ$*" is defined by:*

$$\alpha \circ X := \begin{cases} \sum_{j=1}^{X} W_j, & X > 0, \\ 0, & X = 0. \end{cases} \tag{2}$$

It should be noted that the NBT operator was introduced by Ristić et al. [24] and has since been widely applied and extended in the modeling of integer-valued time series (see, e.g., [25–28] for some recent contributions). Also, note that conditional on $X = x > 0$, the RV $\alpha \circ X$ represents a sum of x independent geometrically distributed RVs. Consequently, by a classical result (see, e.g., Pitman [29]), $\alpha \circ X$ follows a negative binomial distribution with "number of successes" x and success probability $p = 1/(1+\alpha)$. Hence, the conditional PMF of $\alpha \circ X$ is given by:

$$\mathbb{P}(\alpha \circ X = y \mid X = x\} = \binom{x+y-1}{y} \left(\frac{1}{1+\alpha}\right)^x \left(\frac{\alpha}{1+\alpha}\right)^y, \qquad y = 0, 1, 2, \ldots, \tag{3}$$

where according to Equation (2), we adopt the convention $\alpha \circ 0 = 0$. Using Equations (1)–(3), the conditional probability generating function (PGF) of $\alpha \circ X$ is obtained as:

$$\begin{aligned} \Psi_{\alpha\circ X|X}(u;\alpha) = \mathbb{E}\left[u^{\alpha\circ X} \mid X\right] &= \left(\mathbb{E}\left[u^{W_1}\right]\right)^X = \left[\frac{1}{1+\alpha} \sum_{x=0}^{\infty} \left(\frac{\alpha u}{1+\alpha}\right)^x\right]^X \\ &= (1+\alpha-\alpha u)^{-X}, \end{aligned} \tag{4}$$

for $|u| < (1+\alpha)/\alpha$. Based on properties of PGFs and the NBT operator (see, e.g., [24]), the conditional moments of $\alpha \circ X$ follow directly. In particular, the conditional mean is:

$$\mathbb{E}\left[\alpha \circ X \mid X\right] = \frac{\partial}{\partial u} \Psi_{\alpha\circ X|X}(u;\alpha)\Big|_{u=1} = \alpha X, \tag{5}$$

while the second factorial moment satisfies:

$$\mathbb{E}\left[(\alpha \circ X)(\alpha \circ X - 1) \mid X\right] = \frac{\partial^2}{\partial u^2} \Psi_{\alpha\circ X|X}(u;\alpha)\Big|_{u=1} = \alpha^2 X(X+1).$$

Consequently, the conditional second moment and variance are given by:

$$\mathbb{E}\left[(\alpha \circ X)^2 \mid X\right] = \alpha^2 X^2 + \alpha(1+\alpha)X, \qquad \text{Var}\left[\alpha \circ X \mid X\right] = \alpha(1+\alpha)X. \tag{6}$$

In the following, the NBT operator will play a key role in generating the latent-state dynamics underlying the MIN–SB process.

2.2. Power Series Distributions

Following [21–23], we next introduce a general class of integer-valued innovations based on power series (PS) distributions.

Definition 2. *Let $(\varepsilon_t), t \in \mathbb{Z}$ be a series of IID random variables. We say that (ε_t) follows a power series (PS) distribution if its PMF is of the form*

$$p_\varepsilon(x) := \mathbb{P}(\varepsilon_t = x) = \frac{a(x)\,\theta^x}{f(\theta)}, \qquad x \in \mathcal{S}, \tag{7}$$

where $\mathcal{S} \subseteq \mathbb{Z}^+ = 0, 1, 2, \ldots$ is a discrete support, $a(x) \geq 0$ is a weight function, $\theta > 0$ is a parameter, and $f(\theta) := \sum_{x\in\mathcal{S}} a(x)\,\theta^x$ is a power series, convergent for $\theta \in (0, R)$, with $R > 0$.

For suitable choices of $a(x)$ and $f(\theta)$, Equation (7) yields many classical discrete distributions, as illustrated in Table 1. Straightforward calculations, similar to the previous section (see, e.g., [21]), show that the mean, variance, and PGF of (ε_t) are given by the following proposition.

Proposition 1. *Let (ε_t) be a PS-distributed series with the PMF given by Equation (7). Then*

$$\mu_\varepsilon := \mathbb{E}[\varepsilon_t] = \theta g'(\theta), \quad \sigma_\varepsilon^2 := \mathrm{Var}[\varepsilon_t] = \mu_\varepsilon + \theta^2 g''(\theta), \quad \Psi_\varepsilon(u) := \mathbb{E}[u^{\varepsilon_t}] = \frac{f(\theta u)}{f(\theta)}, \tag{8}$$

where $g(\theta) = \log f(\theta)$.

Since f and g are continuously differentiable on $(0, R)$, the PGF $\Psi_\varepsilon(u;\theta)$ is well defined on $(0, R/\theta)$. Closed-form expressions for the mean, variance, and PGF corresponding to several well-known PS distributions are also reported in Table 1. Further, we define the dispersion index $D_\varepsilon(\theta) := \sigma_\varepsilon^2 - \mu_\varepsilon = \theta^2 g''(\theta)$. Thus, over-dispersion is directly characterized by the convexity of $g(\theta)$, because the inequality $D_\varepsilon(\theta) > 0$ holds if and only if $g''(\theta) > 0$ for all $\theta \in (0, R)$.

Table 1. Some specific PS-distributions, along with their over-dispersion indices and PGFs.

Distributions	$\mathcal{S}$	$a(x)$	$(0, R)$	$f(\theta)$	μ_ε	$D_\varepsilon(\theta)$	$\Psi_\varepsilon(u;\theta)$	R
1. Bernoulli	$\{0,1\}$	1	$(0,\infty)$	$1+\theta$	$\frac{\theta}{1+\theta}$	$-\frac{\theta^2}{(1+\theta)^2}$	$\frac{1+\theta u}{1+\theta}$	∞
2. Binomial	$\{0,\ldots,n\}$	$\binom{n}{x}$	$(0,\infty)$	$(1+\theta)^n$	$\frac{n\theta}{1+\theta}$	$-\frac{n\theta^2}{(1+\theta)^2}$	$\left(\frac{1+\theta u}{1+\theta}\right)^n$	∞
3. Poisson	$\{0,\ldots,\infty\}$	$\frac{1}{x!}$	$(0,\infty)$	$\exp(\theta)$	θ	0	$\exp(\theta(u-1))$	∞
4. Geometric	$\{0,\ldots,\infty\}$	1	$(0,1)$	$\frac{1}{1-\theta}$	$\frac{\theta}{1-\theta}$	$\frac{\theta^2}{(1-\theta)^2}$	$\frac{1-\theta}{1-\theta u}$	1
5. Negative Binomial	$\{0,\ldots,\infty\}$	$\binom{x+n-1}{x}$	$(0,1)$	$\frac{1}{(1-\theta)^n}$	$\frac{n\theta}{1-\theta}$	$\frac{n\theta^2}{(1-\theta)^2}$	$\left(\frac{1-\theta}{1-\theta u}\right)^n$	1
6. Logarithmic	$\{1,\ldots,\infty\}$	$\frac{1}{x}$	$(0,1)$	$-\ell(\theta)^*$	$\frac{-\theta}{(1-\theta)\ell(\theta)}$	$\frac{-\theta^2\,[1+\ell(\theta)]}{(1-\theta)^2\ell^2(\theta)}$	$\frac{\ell(\theta u)}{\ell(\theta)}$	1

$^*\ell(\theta) = \log(1-\theta)$.

Note that the above constructions provide the main probabilistic building blocks of the MIN–SB process. In particular, the NBT operator governs the random propagation mechanism, while PS distributions allow flexible modeling of the innovation structure. These components will be combined in the next section to define the MIN–SB dynamics and derive its stochastic properties.

3. Definition and Key Properties of the Model

In this section, we introduce the MIN–SB process with PS innovations and discuss its main structural and stochastic properties.

Definition 3. *Let $(\Omega, \mathcal{F}, P)$ be the probability space and $t \in \mathbb{Z}$. The MIN–SB process is represented by the following time series:*

(i) (ε_t) is the IID series following the PS-distribution given by Equation (7).

(ii) (X_t) is an autoregressive (AR) process defined by:

$$X_t = \alpha \circ (X_{t-1} + q_{t-1}\, \varepsilon_{t-1}), \tag{9}$$

where $0 < \alpha < 1$, "$\circ$" denotes the NBT operator introduced in Equation (2), and:

$$q_t = I(\varepsilon_{t-1} \geq c) = \begin{cases} 1, & \varepsilon_{t-1} \geq c, \\ 0, & \varepsilon_{t-1} < c \end{cases} \tag{10}$$

is the indicator variable with the integer-valued threshold parameter $c \in \mathbb{N}_0$, governs the activation of the innovation component.

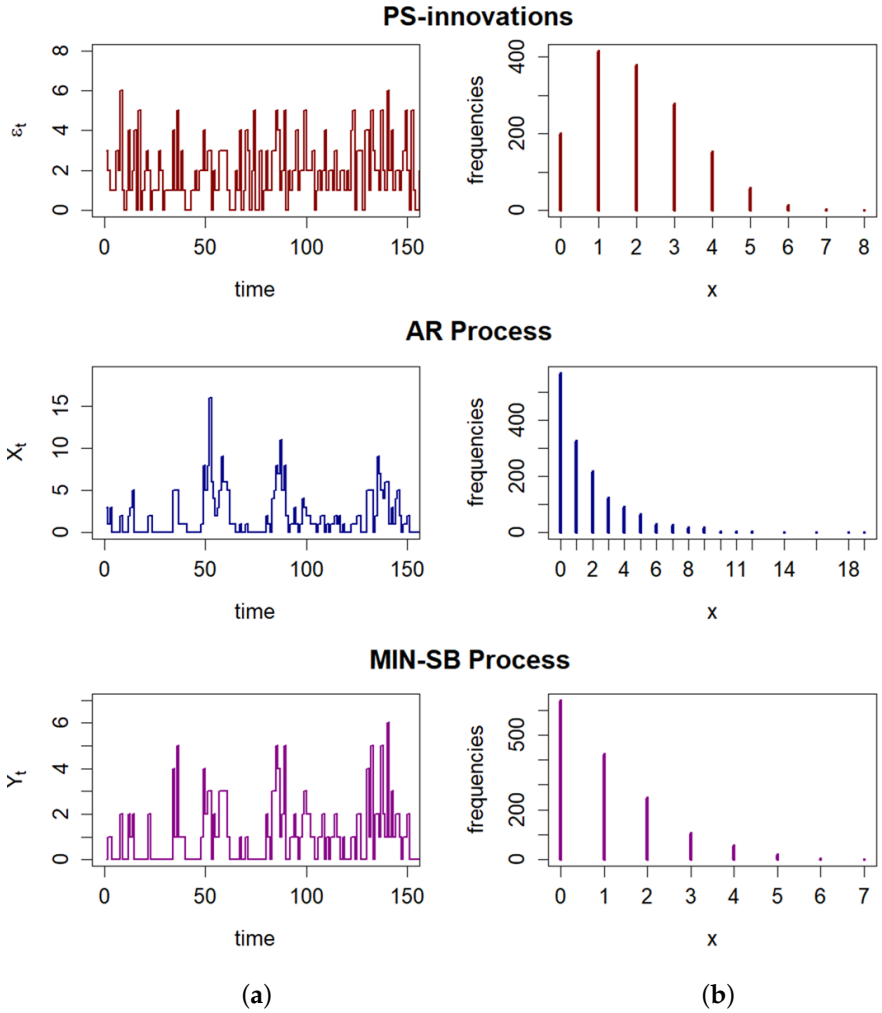
(iii) The observable MIN–SB series (Y_t) is defined by:

$$Y_t = \min(X_t, \varepsilon_t). \tag{11}$$

In the following, we assume that the probability space is expanded by filtration $F = (\mathcal{F}_t)$, where $\mathcal{F}_t = \sigma\{\varepsilon_t \mid s \leq t\}$, $t \in \mathbb{Z}$. Then, (ε_t) are $\mathcal{F}_t$ adapted and constitute the innovation series of the MIN–SB process. In contrast, the process (X_t) is $\mathcal{F}_{t-1}$ adapted and has an AR structure that differs from the usual INAR framework. Namely, when $q_{t-1} = 0$, the value of X_t is equal to $\alpha \circ X_{t-1}$, i.e., the process (X_t) evolves only through the thinning mechanism. Otherwise, the case $q_{t-1} = 1$ indicates a pronounced fluctuation of (X_t). As an illustration, Figure 1 shows the realizations of all the mentioned series, where PS-distributed innovations with Poisson distribution and parameter $\theta = 2$ are taken.

From a dynamical perspective, the MIN–SB process interpolates between two limiting regimes. Larger values of c imply $q_t \overset{as}{=} 0$ and therefore the equality $X_t = \alpha \circ X_{t-1}$. Hence, when the zero has occurred, i.e., $X_{t-1} = 0$, for some $t \in \mathbb{Z}$, it follows that $X_t = \alpha \circ X_{t-1} = 0$, that is, $X_{t+j} = Y_{t+j} = 0$, for all $j = 0, 1, 2, \ldots$. Thus, the thinning property $\alpha \circ 0 = 0$ implies absorption of the both series (X_t) and (Y_t), once a zero value occurs. For these reasons, the inclusion of the innovation term $(q_t \varepsilon_t)$ prevents this collapse, ensuring that zero is not an absorbing state and allowing the process to retain non-degenerate stochastic dynamics (as shown in Figure 1). In addition, the case $c = 0$ corresponds to $q_t \overset{as}{=} 1$, i.e., to the boundary regime in which the gating mechanism is effectively inactive. Then, the latent process (X_t) coincides with a standard INAR(1)-type process, while the observed series (Y_t) reduces to the common minification-type process without threshold filtering.

In this way, the MIN–SB process can be interpreted as a threshold-driven modification of a branching-type mechanism. The latent process (X_t) evolves according to a thinning operation, while the observed process (Y_t) is obtained through a minimum operator with the innovation series (ε_t). Finally, the threshold parameter c acts as a gating mechanism, determining when past innovations contribute to the current state. This structure naturally induces zero inflation and clustering, which are key features of the model. Moreover, the indicator process (q_t) acts as a latent state variable, implying that the observed process (Y_t) inherits an implicit hidden Markov structure. This property, as well as some other stochastic characteristics of the MIN–SB process (conditional probabilities and strict stationarity conditions), is given by the following statement.



(**a**) (**b**)

Figure 1. (**a**) Realizations of the PS-series (ε_t) with Poisson distribution, AR series (X_t), and MIN–SB series (Y_t). (**b**) The empirical distributions of observed series. (Parameters are: $\alpha = 0.5$, $c = 1$, $\theta = 2$).

Theorem 1. *Let the MIN–SB model be defined by Equations (9)–(11) and define the augmented process* $\mathbf{Z}_t = (X_t, \varepsilon_{t-1})$, $t \in \mathbb{Z}$. *Then:*

(i) The process $(\mathbf{Z}_t)$ *forms a bivariate time-homogeneous Markov chain.*

(ii) The process (X_t) *is a latent component of* $(\mathbf{Z}_t)$ *with the conditional probabilities:*

$$\begin{aligned} p_{i,j}^{(X)} &:= \mathbb{P}(X_t = j \mid X_{t-1} = i) \\ &= \sum_{k=0}^{\infty} p_\varepsilon(k;\theta) \Bigg[(1 - S_\varepsilon(c)) \binom{i+j-1}{j} \Big(\frac{1}{1+\alpha}\Big)^i \Big(\frac{\alpha}{1+\alpha}\Big)^j \\ &\qquad + S_\varepsilon(c) \binom{i+j+k-1}{j} \Big(\frac{1}{1+\alpha}\Big)^{i+k} \Big(\frac{\alpha}{1+\alpha}\Big)^j \Bigg], \end{aligned} \tag{12}$$

where $S_\varepsilon(c) = \mathbb{P}(\varepsilon_t \geq c)$ *is the survival function of the series* (ε_t) *and, by convention,*

$$\binom{j-1}{j} = \begin{cases} 1, & j = 0, \\ 0, & j = 1, 2, \ldots \end{cases}$$

(iii) The observed process (Y_t), $t \in \mathbb{Z}$ *is a measurable transformation of* $(\mathbf{Z}_t, \varepsilon_t)$, *and therefore represents a hidden-state process, with the conditional probabilities:*

$$\begin{aligned} p_{i,j}^{(Y)} &:= \mathbb{P}(Y_t = j \mid Y_{t-1} = i) \\ &= p_\varepsilon(j) \sum_{k=j+1}^{\infty} \mathbb{P}(X_t = k \mid Y_{t-1} = i) + S_\varepsilon(j)\, \mathbb{P}(X_t = j \mid Y_{t-1} = i). \end{aligned} \tag{13}$$

(iv) If $0 < \alpha < 1$, *the augmented Markov chain* $(\mathbf{Z}_t)$ *is positive recurrent and admits a unique stationary distribution. Consequently, the latent process* (X_t) *and the observed process* (Y_t) *are stationary with finite moments.*

Proof. (i) Let $\mathcal{G}_t = \sigma\{\mathbf{Z}_s : s \le t\}$ denote the natural filtration generated by the augmented process $\mathbf{Z}_t = (X_t, \varepsilon_{t-1})$. According to Equation (9), the latent variable X_{t+1} satisfies:

$$X_{t+1} = \alpha \circ \big(X_t + q_t \varepsilon_t\big),$$

where $q_t = I(\varepsilon_{t-1} \ge c)$. Since ε_t is independent of $\mathcal{G}_t$ and the randomness of the NBT operator is also independent of the past, the conditional distribution of $\mathbf{Z}_{t+1} = (X_{t+1}, \varepsilon_t)$ depends only on the current state $\mathbf{Z}_t = (X_t, \varepsilon_{t-1})$. Hence, for any measurable set $A \subseteq \mathbb{N}_0^2$,

$$\mathbb{P}(\mathbf{Z}_{t+1} \in A \mid \mathcal{G}_t) = \mathbb{P}(\mathbf{Z}_{t+1} \in A \mid \mathbf{Z}_t),$$

which proves that $(\mathbf{Z}_t)$ is a Markov chain. In addition, the transition mechanism does not depend on t, so the chain $(\mathbf{Z}_t)$ is time-homogeneous.

(ii) The conditional probabilities of (X_t) can be derived by the thinning operation and the threshold-based innovation component. Using Equation (9), the law of total probability, and the property of the NBT operator given by Equation (3), conditioning on $X_{t-1} = i$ yields:

$$\begin{aligned} p_{i,j}^{(X)} &:= \mathbb{P}(X_t = j \mid X_{t-1} = i) = \mathbb{P}(\alpha \circ (X_{t-1} + q_{t-1}\varepsilon_{t-1}) = j \mid X_{t-1} = i) \\ &= \mathbb{P}(\alpha \circ X_{t-1} = j \mid X_{t-1} = i) P(q_{t-1} = 0) + \mathbb{P}(\alpha \circ (X_{t-1} + \varepsilon_{t-1}) = j \mid X_{t-1} = i) P(q_{t-1} = 1) \\ &= [1 - S_\varepsilon(c)] \binom{i+j-1}{j} \left(\frac{1}{1+\alpha}\right)^i \left(\frac{\alpha}{1+\alpha}\right)^j + S_\varepsilon(c) \sum_{k=0}^{\infty} \mathbb{P}(\varepsilon_{t-1} = k)\, P\big(\alpha \circ (i+k) = j\big) \\ &= [1 - S_\varepsilon(c)] \binom{i+j-1}{j} \left(\frac{1}{1+\alpha}\right)^i \left(\frac{\alpha}{1+\alpha}\right)^j + S_\varepsilon(c) \sum_{k=0}^{\infty} p_\varepsilon(k) \binom{i+j+k-1}{i+k} \left(\frac{1}{1+\alpha}\right)^{i+k} \left(\frac{\alpha}{1+\alpha}\right)^j. \end{aligned}$$

Here, according to Equation (10), $S_\varepsilon(c) = \mathbb{P}(q_{t-1} = 1) = 1 - \mathbb{P}(q_{t-1} = 0)$ is the survival function of (ε_t). Accordingly, using the equality $\sum_{k=0}^{\infty} p_\varepsilon(k) = 1$, the stated formula in Equation (12) is easily obtained.

(iii) First, note that the process (X_t) represents the latent component of the vector series $(\mathbf{Z}_t)$, so the observed process (Y_t) can be written as:

$$Y_t = h(\mathbf{Z}_t, \varepsilon_t) = \min(X_t, \varepsilon_t), \qquad t \in \mathbb{Z}.$$

Hence, (Y_t) is a measurable function of the augmented system $(\mathbf{Z}_t, \varepsilon_t)$. Now, fix an arbitrary $i, j \in \mathbb{N}_0$ and note that when $X_t = k$, we have:

$$\mathbb{P}(Y_t = j \mid X_t = k) = \left\{ \begin{array}{ll} p_\varepsilon(j), & j < k, \\ S_\varepsilon(k), & j = k, \\ 0, & j > k, \end{array} \right\} = p_\varepsilon(j) I(j < k) + S_\varepsilon(k) I(j = k), \tag{14}$$

where $p_\varepsilon(j) = \mathbb{P}(\varepsilon_t = j)$ and $S_\varepsilon(k) = \mathbb{P}(\varepsilon_t \geq k)$. Thus, for fixed $X_t = k$, the event $\{Y_t = j\}$ depends only on the innovation ε_t. Moreover, by model assumptions, ε_t is independent of $\mathcal{F}_{t-1}$ and hence independent of Y_{t-1}. Consequently, using again the law of total probability and Equation (14), conditioning on X_t, one obtains:

$$\begin{aligned} p_{i,j}^{(Y)} = \mathbb{P}(Y_t = j \mid Y_{t-1} = i) &= \sum_{k=0}^{\infty} \mathbb{P}(Y_t = j \mid X_t = k, Y_{t-1} = i)\, \mathbb{P}(X_t = k \mid Y_{t-1} = i) \\ &= \sum_{k=0}^{\infty} \mathbb{P}(Y_t = j \mid X_t = k)\, \mathbb{P}(X_t = k \mid Y_{t-1} = i) \\ &= \sum_{k=0}^{\infty} \left[p_\varepsilon(j) I(j < k) + S_\varepsilon(k) I(j = k) \right] \mathbb{P}(X_t = k \mid Y_{t-1} = i) \\ &= \sum_{k=0}^{\infty} p_\varepsilon(j) I(j < k)\, \mathbb{P}(X_t = k \mid Y_{t-1} = i) \\ &\qquad + S_\varepsilon(j)\, \mathbb{P}(X_t = j \mid Y_{t-1} = i), \end{aligned}$$

which coincides with the formula stated in Equation (13).

(iv) Consider the Lyapunov function $V(x, e) = x + \mu_\varepsilon$, where $\mu_\varepsilon = \mathbb{E}[\varepsilon_t] > 0$ is the finite mean of the IID series (ε_t). Using the definition of the NBT operator and the inequality $0 \leq q_{t-1} \leq 1$, it follows that:

$$\mathbb{E}[X_t \mid \mathbf{Z}_{t-1}] = \alpha\big(X_{t-1} + q_{t-1}\mu_\varepsilon\big) \leq \alpha(X_{t-1} + \mu_\varepsilon) = \alpha V(\mathbf{Z}_{t-1}),$$

i.e.,

$$\mathbb{E}[V(\mathbf{Z}_t) \mid \mathbf{Z}_{t-1}] = \mathbb{E}[X_t + \mu_\varepsilon \mid \mathbf{Z}_{t-1}] \leq \alpha V(\mathbf{Z}_{t-1}) + \mu_\varepsilon. \tag{15}$$

Hence, under the assumption $0 < \alpha < 1$, for the conditional drift function one obtains:

$$\Delta(z) := \mathbb{E}[V(\mathbf{Z}_t) - V(\mathbf{Z}_{t-1}) \mid \mathbf{Z}_{t-1} = z] \leq (\alpha - 1)V(z) + \mu_\varepsilon \leq -\delta < 0,$$

when $V(z) \geq (\delta + \mu_\varepsilon)/(1 - \alpha) > 0$. Thus, the Markov chain $(\mathbf{Z}_t)$ satisfies a Foster–Lyapunov drift condition (see, e.g., [30]), implying its positive recurrence and existence of a unique stationary distribution. Consequently, stationarity of (X_t) follows since it is a projection of Markov chain $(\mathbf{Z}_t)$, while stationarity of (Y_t) follows because it is a measurable transformation of stationary vector process $(\mathbf{Z}_t, \varepsilon_t)$. Finally, standard drift arguments imply finiteness of moments for (X_t), and since $0 \leq Y_t \leq X_t$, the same holds for (Y_t). □

Remark 1 (Geometric ergodicity and SLLN)**.** *Note that the Foster–Lyapunov drift condition established in part (iv) of Theorem 1 is linear in the Lyapunov function $V(x, e)$. Moreover, the inequality $0 < \alpha < 1$ implies the geometric drift condition given by Equation (15). Therefore, the augmented process $(\mathbf{Z}_t)$ is geometrically ergodic and according to the classical theory of Markov chains (see, e.g., Meyn and Tweedie [31] or Gallegos-Herrada et al. [32]), it is positively recurrent and converges to its stationary distribution at a geometric rate. In addition, by the ergodic theorem for Markov chains (see, e.g., Douc et al. [33]), a strong law of large numbers (SLLN) holds for any*

integrable measurable function of $(\mathbf{Z}_t)$. *In particular, since* $Y_t = h(\mathbf{Z}_t, \varepsilon_t) = \min(X_t, \varepsilon_t)$ *and* $\mathbb{E}|h(\mathbf{Z}_1, \varepsilon_1)| = \mathbb{E}|\min(X_1, \varepsilon_1)| < \infty$, *the following almost surely (a.s.) convergence holds:*

$$\frac{1}{T}\sum_{t=1}^{T} Y_t \xrightarrow{a.s.} \mathbb{E}[Y_1], \quad T \to +\infty.$$

Remark 2 (Distributional properties and PGFs)**.** *According to the conditional probabilities given by Equations (12) and (13), the marginal PMFs of the MIN–SB series* (X_t) *and* (Y_t) *can be obtained by recurrence relations:*

$$p_X(x) := \mathbb{P}\{X_t = x\} = \sum_{k=0}^{\infty} P\{X_t = x \mid X_{t-1} = k\} P\{X_{t-1} = k\} = \sum_{k=0}^{\infty} p_{k,x}^{(X)} p_X(k),$$
$$p_Y(y) := P\{Y_t = y\} = \sum_{k=0}^{\infty} P\{Y_t = y \mid Y_{t-1} = k\} P\{Y_{t-1} = k\} = \sum_{k=0}^{\infty} p_{k,y}^{(Y)} p_Y(k).$$

To obtain more explicit distributional forms of these processes, we now consider the PGFs of the series (X_t) *and* (Y_t). *This approach also allows us to summarize the combined effect of thinning and thresholding in a compact analytical form.*

First, note that by Theorem 1, the augmented chain $(\mathbf{Z}_t)$ *admits a unique stationary distribution, so the process* (X_t) *is also stationary, and its PGF is given with* $\Psi_X(u) = \mathbb{E}[u^{X_t}]$, $|u| \le 1$. *Next, denote with* $\eta_t = q_t \varepsilon_t$ *threshold innovations, which are mutually uncorrelated (see, e.g., [20]) and also uncorrelated with* X_t, *for each* $t \in \mathbb{Z}$. *Then, according to Equations (7), (8) and (10), as well as by applying the law of total probability, the PGF of the RVs* (η_t) *can be obtained as follows:*

$$\Psi_\eta(u) := \mathbb{E}[u^{\eta_t}] = \mathbb{E}[u^0]\,\mathbb{P}(q_t = 0) + \mathbb{E}[u^{\varepsilon_t}]\,\mathbb{P}(q_t = 1) = 1 - S_\varepsilon(c) + S_\varepsilon(c)\Psi_\varepsilon(u)$$
$$= 1 + S_\varepsilon(c)\left(\frac{f(\theta u)}{f(\theta)} - 1\right).$$

Further, note that Equations (4) and (9) imply a recursive relation:

$$\Psi_X(u) = \mathbb{E}[u^{X_t}] = \mathbb{E}[u^{\alpha \circ (X_{t-1} + \eta_{t-1})}] = \mathbb{E}\left(\phi(u)^{X_{t-1} + \eta_{t-1}}\right) = \Psi_\eta\left(\phi(u)\right)\,\Psi_X\left(\phi(u)\right),$$

where $\phi(u) = (1 + \alpha - \alpha u)^{-1}$. *Thus, by iteratively repeating this relation and using the previously obtained PGF* $\Psi_\eta(u)$, *it follows:*

$$\Psi_X(u) = \lim_{k\to\infty} \prod_{j=1}^{k} \Psi_\eta\left(\phi^j(u)\right) \Psi_X\left(\phi^k(u)\right) = \prod_{k=1}^{\infty} \Psi_\eta\left(\phi^k(u)\right)$$
$$= \prod_{k=1}^{\infty} \left[1 + S_\varepsilon(c)\left(\frac{f\left(\theta\,\phi^k(u)\right)}{f(\theta)} - 1\right)\right]. \tag{16}$$

Here, for the function $\phi^k(u)$, *using the induction method and after simple, but somewhat lengthy algebraic manipulations, one obtains:*

$$\phi^k(u) = \frac{1 - \alpha u - \alpha^k(1-u)}{1 - \alpha u - \alpha^{k+1}(1-u)},$$

and the product in Equation (16) converges when $0 < \alpha < 1$, *because* $\phi^k(u) \to 1$ *when* $k \to \infty$.

Now, let us consider the PGF $\Psi_Y(u) := \mathbb{E}[u^{Y_t}]$, $|u| \leq 1$, *of the observed process* (Y_t). *Conditioning on* $X_t = k$ *and using the independence of* ε_t, *one obtains:*

$$Y_t = \begin{cases} \varepsilon_t, & \varepsilon_t < k, \\ k, & \varepsilon_t \geq k. \end{cases}$$

Hence, the conditional PGF of (Y_t) *is given by:*

$$\Psi_{Y|X}(u) = \mathbb{E}\big(u^{Y_t} \mid X_t = k\big) = \sum_{j=0}^{k-1} u^j \, p_\varepsilon(j) + u^k \, S_\varepsilon(k),$$

and taking expectation with respect to X_t *yields the PGF representation:*

$$\Psi_Y(u) = \mathbb{E}\Big[\mathbb{E}\big(u^{Y_t} \mid X_t\big)\Big] = \mathbb{E}\left[\sum_{j=0}^{X_t-1} u^j \, p_\varepsilon(j) + u^{X_t} S_\varepsilon(X_t)\right]. \tag{17}$$

Note that the representation given by Equation (17) *is not in closed form, but it provides a complete characterization of the distribution for* (Y_t) *and thus allows further derivations of some of its properties. The PGFs given by Equations* (16) *and* (17) *also clearly show how the distribution of innovations* (ε_t) *and the thinning mechanism determine the marginal behavior of the MIN–SB process. Thus, their structure reflects the compounded effect of thinning and thresholding, and provides a convenient analytical representation for deriving distributional properties. In particular, the PMFs of the series* (X_t) *and* (Y_t) *can be expressed in alternative forms:*

$$p_X(k) = \frac{1}{k!}\Psi_X^{(k)}(0), \qquad p_Y(k) = \frac{1}{k!}\Psi_Y^{(k)}(0). \tag{18}$$

Based on Theorem 1, the following statement characterizes the moment structure and dependence decay of both the latent and observable components of the MIN–SB model.

Theorem 2. *Let* (X_t) *and* (Y_t) *be the stochastic processes defined by Equations* (9) *and* (11), *respectively, and assume that* $0 < \alpha < 1$. *Then the following statements hold:*

(i) The process (X_t) *admits a unique strictly stationary distribution with:*

$$\mathbb{E}[X_t] = \frac{\alpha S_\varepsilon(c)\mu_\varepsilon}{1-\alpha}, \quad \mathrm{Var}[X_t] = \frac{\alpha \, S_\varepsilon(c) \, \mu_\varepsilon}{(1-\alpha)^2} + \frac{\alpha^2}{1-\alpha^2}\Big[S_\varepsilon(c)\sigma_\varepsilon^2 + S_\varepsilon(c)(1-S_\varepsilon(c))\mu_\varepsilon^2\Big]. \tag{19}$$

where μ_ε *and* σ_ε^2 *are given by Equation* (8). *In addition, the autocorrelation function (ACF) of* (X_t) *is* $\rho_X(h) = \alpha^h$, $h \geq 0$.

(ii) The stationary mean and variance of the MIN–SB process (Y_t) *satisfy:*

$$\mathbb{E}[Y_t] = \mathbb{E}\left[\sum_{j=1}^{X_t} S_\varepsilon(j)\right], \quad \mathrm{Var}[Y_t] = \mathbb{E}\left[\sum_{j=1}^{X_t}(2j-1)\, S_\varepsilon(j)\right] - (\mathbb{E}[Y_t])^2. \tag{20}$$

In addition, the ACF of (Y_t) *satisfies* $|\rho_Y(h)| \leq C\,\alpha^h$, *for some finite constant* $C > 0$.

Proof. (i) Under the condition $0 < \alpha < 1$ and the stationary regime established in item (iv) of Theorem 1, the augmented chain $(\mathbf{Z}_t)$ admits a unique stationary distribution, so the latent component (X_t) is also stationary. Taking expectations in the recursion of (X_t), given by Equation (9), and using the moment properties of the NBT operator (cf. Equation (5)), one obtains:

$$\mathbb{E}(X_t \mid X_{t-1} + q_{t-1}\varepsilon_{t-1}) = \alpha(X_{t-1} + q_{t-1}\varepsilon_{t-1}). \tag{21}$$

After taking expectations again, Equation (21) yields:

$$\mathbb{E}[X_t] = \alpha \mathbb{E}[X_{t-1}] + \alpha S_\varepsilon(c)\mu_\varepsilon,$$

and solving this equation for the stationary mean $\mathbb{E}[X_t]$ gives the first equality in Equations (19). To compute the variance, we apply the law of total variance and (again) the properties of the NBT operator (cf. Equation (6)):

$$\begin{aligned}\operatorname{Var}(X_t) &= \mathbb{E}\big[\operatorname{Var}(X_t \mid X_{t-1} + q_{t-1}\varepsilon_{t-1})\big] + \operatorname{Var}\big[\mathbb{E}(X_t \mid X_{t-1} + q_{t-1}\varepsilon_{t-1})\big] \\ &= \alpha(1+\alpha)\,\mathbb{E}[X_{t-1} + \eta_{t-1}] + \alpha^2\,\operatorname{Var}[X_{t-1} + \eta_{t-1}] \\ &= \alpha(1+\alpha)\,\big(\mathbb{E}[X_t] + \mathbb{E}[\eta_t]\big) + \alpha^2\,\big(\operatorname{Var}[X_t] + \operatorname{Var}[\eta_t]\big),\end{aligned}$$

which implies:

$$\begin{aligned}(1-\alpha^2)\operatorname{Var}[X_t] &= \alpha(1+\alpha)\left(\frac{\alpha}{1-\alpha}\mathbb{E}[\eta_t] + \mathbb{E}[\eta_t]\right) + \alpha^2 \operatorname{Var}[\eta_t] \\ &= \frac{\alpha(1+\alpha)}{1-\alpha}\,\mathbb{E}[\eta_t] + \alpha^2 \operatorname{Var}[\eta_t].\end{aligned}$$

Here, as stated in Remark 2, the RVs $\eta_t = q_t\varepsilon_t$ are mutually uncorrelated, and

$$\mathbb{E}[\eta_t] = S_\varepsilon(c)\mu_\varepsilon, \qquad \operatorname{Var}(\eta_t) = \mathbb{E}[\eta_t^2] - (\mathbb{E}[\eta_t])^2 = S_\varepsilon(c)\sigma_\varepsilon^2 + S_\varepsilon(c)(1 - S_\varepsilon(c))\mu_\varepsilon^2,$$

from which the second formula in Equation (19) is easily obtained.

Now, starting from the recursive representation of the latent process (X_t), given by Equation (9), we iterate its conditional expectations. Since the RVs η_t are uncorrelated with X_t, for each $t \in \mathbb{Z}$, we have:

$$\mathbb{E}(X_t \mid X_{t-1}) = \mathbb{E}(\alpha \circ (X_{t-1} + \eta_{t-1}) \mid X_{t-1}) = \alpha(X_{t-1} + \mathbb{E}[\eta_{t-1}]),$$

and applying the same argument recursively yields:

$$\mathbb{E}(X_t \mid X_{t-2}) = \alpha(\mathbb{E}(X_{t-1} \mid X_{t-2}) + \mathbb{E}[\eta_{t-1}]) = \alpha^2 X_{t-2} + \alpha^2 \mathbb{E}[\eta_{t-2}] + \alpha \mathbb{E}[\eta_{t-1}].$$

Continuing this procedure for general $h = 1, 2, \ldots$, and using the stationarity of (η_t), i.e., the equality $\mathbb{E}[\eta_{t-j}] = \mathbb{E}[\eta_t]$, for all $j = 1, 2, \ldots$, one obtains:

$$\mathbb{E}(X_t \mid X_{t-h}) = \alpha^h X_{t-h} + \sum_{j=1}^{h} \alpha^j\,\mathbb{E}[\eta_{t-j}] = \alpha^h X_{t-h} + \frac{\alpha(1-\alpha^h)}{1-\alpha}\,\mathbb{E}[\eta_t]. \tag{22}$$

Further, by using the law of iterated expectations and the conditional expectation obtained in Equation (22), we get:

$$\begin{aligned}\mathbb{E}[X_t X_{t-h}] &= \mathbb{E}\big[\,X_{t-h}\,\mathbb{E}(X_t \mid X_{t-h})\big] = \mathbb{E}\left[X_{t-h}\left(\alpha^h X_{t-h} + \frac{\alpha(1-\alpha^h)}{1-\alpha}\,\mathbb{E}[\eta_t]\right)\right] \\ &= \alpha^h \mathbb{E}[X_{t-h}^2] + \frac{\alpha(1-\alpha^h)}{1-\alpha}\,\mathbb{E}[\eta_t]\,\mathbb{E}[X_{t-h}] \\ &= \alpha^h \mathbb{E}[X_t^2] + \frac{\alpha(1-\alpha^h)}{1-\alpha}\,\mathbb{E}[\eta_t]\,\mathbb{E}[X_t],\end{aligned}$$

where, due to the stationarity of the process (X_t), the equalities $\mathbb{E}[X_{t-h}] = \mathbb{E}[X_t]$ and $\mathbb{E}[X_{t-h}^2] = \mathbb{E}[X_t^2]$ hold. On the other hand, since $\mathbb{E}[X_t] = \alpha/(1-\alpha)\,\mathbb{E}[\eta_t]$, substituting this equality into the previous expression yields:

$$\mathbb{E}[X_t X_{t-h}] = \alpha^h \mathbb{E}[X_t^2] + (1-\alpha^h)\left(\mathbb{E}[X_t]\right)^2.$$

Therefore, the covariance of the process (X_t) is obtained as follows:

$$\mathrm{Cov}(X_t, X_{t-h}) = E[X_t X_{t-h}] - (\mathbb{E}[X_t])^2 = \alpha^h\Big(\mathbb{E}[X_t^2] - \left(\mathbb{E}[X_t]\right)^2\Big) = \alpha^h \,\mathrm{Var}[X_t],$$

which implies the equality $\rho_X(h) = \mathrm{Corr}(X_t, X_{t-h}) = \alpha^h$.

(ii) We now derive the first and second moments of the process $Y_t = \min(X_t, \varepsilon_t)$. To this end, according to Equation (14), for each $X_t = k \in \mathbb{N}_0$ one obtains:

$$\mathbb{E}(Y_t \mid X_t = k) = \sum_{j=0}^{k-1} j\, p_\varepsilon(j) + k\, S_\varepsilon(k) = \sum_{j=1}^{k-1} j\left(S_\varepsilon(j) - S_\varepsilon(j+1)\right) + k\, S_\varepsilon(k) = \sum_{j=1}^{k} S_\varepsilon(j), \quad (23)$$

with the convention that empty sums equal zero. Then, taking the expectation with respect to X_t gives the stated expression for $\mathbb{E}[Y_t]$ in Equation (20). Similarly as above, for the second conditional moment we get:

$$\mathbb{E}(Y_t^2 \mid X_t = k) = \sum_{j=0}^{k-1} j^2 p_\varepsilon(j) + k^2 S_\varepsilon(k).$$

By applying discrete summation by parts for j^2, and taking into account the well-known equality $j^2 = \sum_{i=1}^{j}(2i-1)$, it follows:

$$\sum_{j=0}^{k-1} j^2\, p_\varepsilon(j) = \sum_{j=0}^{k-1}\sum_{i=1}^{j}(2i-1)\, p_\varepsilon(j) = \sum_{i=1}^{k-1}(2i-1)\sum_{j=i}^{k-1} p_\varepsilon(j) = \sum_{i=1}^{k-1}(2i-1)\big(S_\varepsilon(i) - S_\varepsilon(k)\big).$$

Therefore,

$$\mathbb{E}(Y_t^2 \mid X_t = k) = \sum_{i=1}^{k-1}(2i-1)\big(S_\varepsilon(i) - S_\varepsilon(k)\big) + \sum_{i=1}^{k}(2i-1)S_\varepsilon(k) = \sum_{i=1}^{k}(2i-1)\, S_\varepsilon(i),$$

and taking expectation with respect to X_t yields:

$$\mathbb{E}(Y_t^2) = \mathbb{E}\left[\sum_{j=1}^{X_t}(2j-1)\, S_\varepsilon(j)\right].$$

Combining this term with the expression for $\mathbb{E}(Y_t)$ gives the second equality in Equation (20).

Further, let $h \geq 0$. By the law of total covariance:

$$\mathrm{Cov}(Y_t, Y_{t-h}) = \mathrm{Cov}\left(\mathbb{E}(Y_t \mid X_{t-h}), \mathbb{E}(Y_{t-h} \mid X_{t-h})\right) + \mathbb{E}[\mathrm{Cov}(Y_t, Y_{t-h} \mid X_{t-h})]. \quad (24)$$

and we first analyze the second term in Equation (24). Note that conditionally on X_{t-h}, the RV Y_{t-h} depends only on ε_{t-h}, whereas Y_t depends exclusively on future innovations $\varepsilon_{t-h+1}, \ldots, \varepsilon_t$. Since the innovations (ε_t) are IID and independent of the past, conditionally on X_{t-h}, the RVs Y_{t-h} and Y_t depend on disjointed sets of innovations. Hence, it follows:

$$\mathrm{Cov}(Y_t, Y_{t-h} \mid X_{t-h}) = 0,$$

and therefore:

$$\mathbb{E}\left[\mathrm{Cov}(Y_t, Y_{t-h} \mid X_{t-h})\right] = 0. \tag{25}$$

Thus, by applying Equations (24) and (25), one obtains:

$$\mathrm{Cov}(Y_t, Y_{t-h}) = \mathrm{Cov}\left(\mathbb{E}(Y_t \mid X_{t-h}), \mathbb{E}(Y_{t-h} \mid X_{t-h})\right). \tag{26}$$

On the other hand, recall that, according to Equation (23), the conditional mean of the observed process (Y_t) satisfies:

$$\mathbb{E}(Y_s \mid X_s) = G(X_s),$$

where $G(k) = \sum_{j=1}^{k} S_\varepsilon(j)$ is a deterministic, non-decreasing function satisfying:

$$0 \le G(k) \le k, \qquad k \in \mathbb{N}_0. \tag{27}$$

Hence, the condition given by Equation (27) implies that the dependence structure of $Y_t = \min(X_t, \varepsilon_t)$ is primarily driven by the latent thinning mechanism, i.e.,

$$\mathbb{E}(Y_{t-h} \mid X_{t-h}) = G(X_{t-h}),$$

and by the law of iterated expectations:

$$\mathbb{E}(Y_t \mid X_{t-h}) = \mathbb{E}\left[\mathbb{E}(Y_t \mid X_t) \mid X_{t-h}\right] = \mathbb{E}\left[G(X_t) \mid X_{t-h}\right].$$

Substituting both of the above expressions into Equation (26) gives:

$$\mathrm{Cov}(Y_t, Y_{t-h}) = \mathrm{Cov}\left(\mathbb{E}\left(G(X_t) \mid X_{t-h}\right), G(X_{t-h})\right), \tag{28}$$

where according to the sub-linearity property in Equation (27), it holds:

$$0 \le \mathbb{E}\left(G(X_t) \mid X_{t-h}\right) \le \mathbb{E}(X_t \mid X_{t-h}). \tag{29}$$

Since the augmented process $(\mathbf{Z}_t)$ is stationary and ergodic, the latent component (X_t) is stationary and its conditional mean satisfies (cf. Equation (22)):

$$\mathbb{E}(X_t \mid X_{t-h}) = \alpha^h X_{t-h} + C_h, \tag{30}$$

where $C_h = \alpha(1-\alpha^h)/(1-\alpha)\,\mathbb{E}[\eta_t]$. In addition, the function $G(k)$ is non-decreasing and satisfies the inequality (27), hence the following:

$$|G(k_1) - G(k_2)| \le |k_1 - k_2|, \qquad k_1, k_2 \in \mathbb{N}_0. \tag{31}$$

Thus, the function $G(k)$ is 1-Lipschitz on $\mathbb{N}_0$.

Further, let $(X_s^{(x)})$ and $(X_s^{(y)})$, $s \ge t-h$, denote two versions of the process (X_t) started from $X_{t-h} = x$ and $X_{t-h} = y$, respectively. Additionally, if we define the function:

$$\varphi(x) := \mathbb{E}[G(X_t) \mid X_{t-h} = x] = \mathbb{E}\left[G(X_t^{(x)})\right],$$

then according to Equations (29)–(31), it follows:

$$\begin{aligned} |\varphi(x) - \varphi(y)| &= \left|\mathbb{E}(G(X_t) \mid X_{t-h} = x) - \mathbb{E}(G(X_t) \mid X_{t-h} = y)\right| \\ &\le \mathbb{E}\left[\left|G(X_t^{(x)}) - G(X_t^{(y)})\right|\right] \le \mathbb{E}\left[\left|X_t^{(x)} - X_t^{(y)}\right|\right] \end{aligned}$$

$$\leq \alpha^h \, |x - y|.$$

Hence, $\varphi(x)$ is α^h-Lipschitz on $\mathbb{N}_0$.

Now, applying the Cauchy–Schwarz inequality and Lemma A1 (see Appendix A) to the Lipschitz functions $\varphi(X)$ and $G(X)$, when $X := X_{t-h}$, one obtains:

$$\left|\operatorname{Cov}(\varphi(X), G(X))\right| \leq \sqrt{\operatorname{Var}(\varphi(X))\operatorname{Var}(G(X))} \leq \alpha^h \operatorname{Var}(X) = \operatorname{Cov}(\alpha^h X + C_h, X),$$

i.e.,

$$\left|\operatorname{Cov}\left(\mathbb{E}\left[G(X_t) \mid X_{t-h}\right], G(X_{t-h})\right)\right| \leq \operatorname{Cov}\left(\alpha^h X_{t-h} + C_h, X_{t-h}\right).$$

Due to stationarity, the right-hand side simplifies to $\alpha^h \operatorname{Var}(X_{t-h}) = \alpha^h \operatorname{Var}(X_t)$, so the above inequality and Equation (28) give:

$$\left|\operatorname{Cov}(Y_t, Y_{t-h})\right| \leq \alpha^h \operatorname{Var}(X_t). \tag{32}$$

Finally, dividing Equation (32) by $\operatorname{Var}(Y_t) > 0$ yields:

$$|\rho_Y(h)| = \frac{\left|\operatorname{Cov}(Y_t, Y_{t-h})\right|}{\operatorname{Var}(Y_t)} \leq C\,\alpha^h,$$

for finite constant $C = \operatorname{Var}(X_t)/\operatorname{Var}(Y_t) > 0$. This proves that the observable process (Y_t) exhibits geometric dependence decay. □

Remark 3 (Over-dispersion and correlation properties)**.** *Recall that the over-dispersion index for innovations $D(\varepsilon_t) = \sigma_\varepsilon^2 - \mu_\varepsilon = \theta^2 g''(\theta)$ is positive if and only if $g(\theta) = \log f(\theta)$ is convex on $(0, R)$. Hence, over-dispersion of (ε_t) is fully characterized by the power-series generating function $f(\theta)$. Further, for the latent MIN–SB process (X_t), using Equation (19), one obtains:*

$$D(X_t) = \operatorname{Var}(X_t) - \mathbb{E}(X_t) = \alpha^2 S_\varepsilon(c)\left[\frac{\mu_\varepsilon}{(1-\alpha)^2} + \frac{\sigma_\varepsilon^2 + (1 - S_\varepsilon(c))\mu_\varepsilon^2}{1-\alpha^2}\right] > 0,$$

for all $\alpha \in (0, 1)$. Thus, the series (X_t) is over-dispersed regardless of the dispersion properties of innovations (ε_t). Otherwise, for the observed process (Y_t), according to Equation (20), we get:

$$D(Y_t) = \mathbb{E}\left[\sum_{j=1}^{X_t} (2j-1)\, S_\varepsilon(j)\right] - \left(\mathbb{E}\left[\sum_{j=1}^{X_t} S_\varepsilon(j)\right]\right)^2 - \mathbb{E}\left[\sum_{j=1}^{X_t} S_\varepsilon(j)\right].$$

Since $0 \leq S_\varepsilon(j) \leq 1$, for all $j \geq 1$, it follows:

$$\sum_{j=1}^{X_t} S_\varepsilon(j) \leq X_t, \qquad \sum_{j=1}^{X_t} (2j-1)\, S_\varepsilon(j) \leq X_t^2,$$

and therefore $\mathbb{E}[Y_t] \leq \mathbb{E}[X_t]$, $\mathbb{E}[Y_t^2] \leq \mathbb{E}[X_t^2]$. This indicates that the min-operator attenuates the over-dispersion generated by the latent process, typically leading to $D(Y_t) < D(X_t)$, although the exact sign and magnitude of $D(Y_t)$ depend on the full parameter configuration and the distribution of innovations (ε_t).

A similar attenuation effect appears in the dependence structure. As proven in Theorem 2, process (X_t) has a geometric autocorrelation dependence identical to that in standard INAR(1) models. On the other hand, process (Y_t) is a measurable transformation of the latent geometrically ergodic Markov chain $(\mathbf{Z}_t)$, so its temporal dependence is inherited from the latent dynamics. Consequently, the autocorrelation decay is governed by the thinning parameter α, while the nonlinear min-operator reduces the overall correlation strength. Figure 2 illustrates these effects. The upper panels display

empirical estimates of the dispersion indices for (X_t) *and* (Y_t)*, showing that over-dispersion is substantially reduced at the observable level. The lower panels present the corresponding ACFs, confirming that the observable MIN–SB series* (Y_t) *exhibits weaker dependence while preserving the geometric decay rate predicted by Theorem 2.*

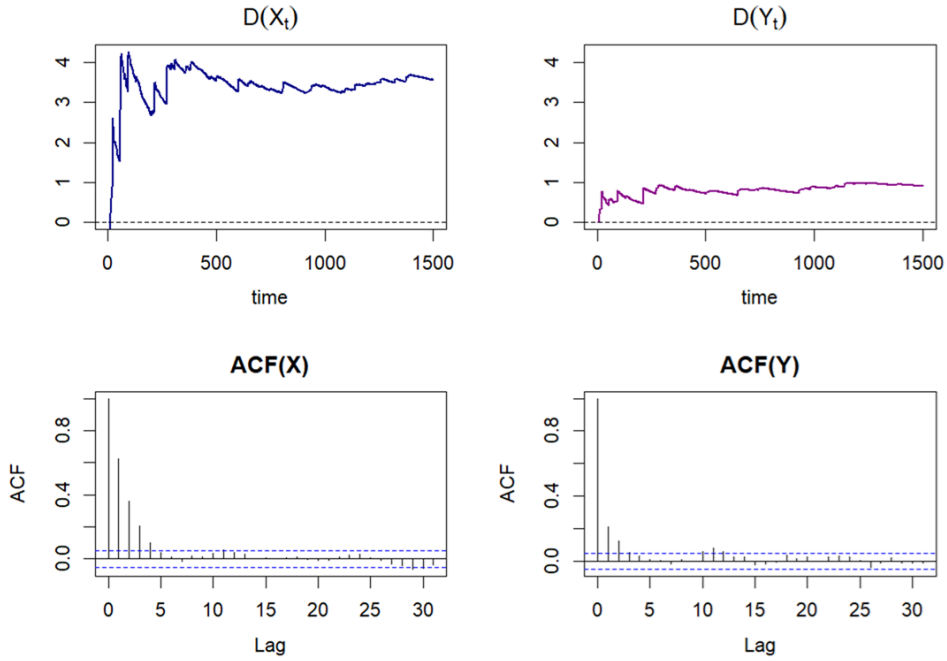


Figure 2. Plots above: Estimated values of the over-dispersion index for the series (X_t) and (Y_t). Plots below: The ACF values of the same series. (Innovations (ε_t) are geometrically distributed and parameters are: $\alpha = \theta = 0.6$, $c = 1$).

In analogy with zero-inflated count-type models (see, e.g., [34,35]), we next examine the distribution of zero-run lengths in the MIN–SB framework. Since the observable process is generated through a latent-state mechanism, the zero-run behavior provides additional insight into how the hidden dynamics shape the empirical zero-inflation patterns. To this end, assume that the innovations distribution satisfies $0 \in \mathcal{S}$ (cf. Table 1) and define a zero run as the number of consecutive zeros occurring between two non-zero observations. The following theorem gives explicit expressions for the expected zero-run lengths of the latent and observable components of the MIN–SB model.

Theorem 3. *Let* (X_t, Y_t) *denote the stationary latent and observable components of the MIN–SB process, defined by Equations (9) and (11), respectively. Then, the expected lengths of the zeros of these two series are:*

$$\mathbb{E}\left[L_0^{(X)}\right] = \frac{f(\theta)}{S_\varepsilon(c)[f(\theta) - f(\theta/(1+\alpha))]}, \tag{33}$$

$$\mathbb{E}\left[L_0^{(Y)}\right] = \frac{f(\theta)[f(\theta)p_X(0) + a(0)(1-p_X(0))]}{(f(\theta)-a(0))\left[p_X(0)\,S_\varepsilon(c)(f(\theta) - f(\theta/(1+\alpha))) + a(0)\left(1-\Psi_X\left((1+\alpha)^{-1}\right)\right)\right]}. \tag{34}$$

where $L_0^{(X)}$ *and* $L_0^{(Y)}$ *are the lengths of zero runs of* (X_t) *and* (Y_t)*, respectively,* $p_X(0) = \mathbb{P}(X_t = 0)$*, and* $f(\theta)$ *is given as in Definition 2.*

Proof. First, we use Equation (12) for the conditional probabilities of (X_t), as well as Equations (7) and (8) for the PMF and PGF of PS-distributed innovations (ε_t), respectively.

According to them, the zero-to-zero and zero-to-nonzero conditional probabilities of the series (X_t) are as follows:

$$\begin{aligned} p_{00}^{(X)} &:= P\{X_t = 0 \mid X_{t-1} = 0\} = 1 - S_\varepsilon(c) + S_\varepsilon(c) \sum_{k=0}^{\infty} p_\varepsilon(k) \Big(\frac{1}{1+\alpha}\Big)^k \\ &= 1 + S_\varepsilon(c) \left[\frac{1}{f(\theta)} \sum_{k=0}^{\infty} a(k) \left(\frac{\theta}{1+\alpha} \right)^k - 1 \right] = 1 + S_\varepsilon(c) \left[\frac{f(\theta/(1+\alpha))}{f(\theta)} - 1 \right] \\ &= 1 + S_\varepsilon(c) \Big[\Psi_\varepsilon\big((1+\alpha)^{-1}\big) - 1 \Big], \\ 1 - p_{00}^{(X)} &:= \mathbb{P}(X_t > 0 \mid X_{t-1} = 0) = S_\varepsilon(c) \Big[1 - \Psi_\varepsilon\big((1+\alpha)^{-1}\big) \Big], \end{aligned} \tag{35}$$

where $\Psi_\varepsilon(u) = f(\theta u)/f(\theta)$ is the PGF and $S_\varepsilon(c) = \mathbb{P}(\varepsilon_t \geq c)$ is the survival function of (ε_t). Successive zeros therefore form runs whose lengths follow a geometric distribution:

$$\mathbb{P}(L_0^{(X)} = k) = (p_{00}^{(X)})^{k-1}(1 - p_{00}^{(X)}), \qquad k = 1, 2, \ldots,$$

and the expected length of zero is:

$$\mathbb{E}\Big(L_0^{(X)}\Big) = \frac{1}{1 - p_{00}^{(X)}}.$$

By substituting Equation (35) here, Equation (33) immediately follows.

Further, using Equation (13) and a similar procedure as the previous one, for the one-step zero-to-zero conditional probability of the series (Y_t) one obtains:

$$p_{00}^{(Y)} := \mathbb{P}(Y_t = 0 \mid Y_{t-1} = 0) = p_\varepsilon(0) + (1 - p_\varepsilon(0))\, \mathbb{P}(X_t = 0 \mid Y_{t-1} = 0),$$

where the expression for the probability $\mathbb{P}(X_t = 0 \mid Y_{t-1} = 0)$ is given by Equation (A1) in the Appendix A. By substitution (A1) in the last expression, we get:

$$p_{00}^{(Y)} = p_\varepsilon(0) + (1 - p_\varepsilon(0)) \frac{p_X(0) p_{00}^{(X)} + p_\varepsilon(0) \Big(\Psi_X\big((1+\alpha)^{-1}\big) - p_X(0) \Big)}{p_X(0) + (1 - p_X(0)) p_\varepsilon(0)}, \tag{36}$$

which, after some algebraic simplification, yields the explicit form of zero-to-nonzero probability:

$$1 - p_{00}^{(Y)} = \frac{(1 - p_\varepsilon(0)) \left[p_X(0) \left(1 - p_{00}^{(X)}\right) + p_\varepsilon(0) \left(1 - \Psi_X\big((1+\alpha)^{-1}\big)\right) \right]}{p_X(0) + (1 - p_X(0))\, p_\varepsilon(0)}. \tag{37}$$

Here, for PS innovations the equality $p_\varepsilon(0) = \mathbb{P}(\varepsilon_t = 0) = a(0)/f(\theta)$ holds, while the expression for $1 - p_{00}^{(X)}$ is given by Equation (35). Finally, in the same way as before, the expected value of the length of zero runs is as follows:

$$\mathbb{E}\Big[L_0^{(Y)}\Big] = \frac{1}{1 - p_{00}^{(Y)}}, \tag{38}$$

and substituting Equation (37) into (38) immediately yields Equation (34). □

Remark 4 (Zero-inflation and clustering of zeros)**.** *Note that the MIN–SB construction generates zero-inflation through two distinct but interacting mechanisms: (i) marginal amplification of the probability mass at zero, and (ii) temporal clustering of zeros induced by the latent Markov dynamics. These effects can be explicitly separated and quantified as follows.*

(i) Probability mass at zero. *Based on the previous results, the stationary mass of zeros for all MIN–SB series can be easily computed. Thus, for the series* (ε_t) *and* (η_t)*, we have:*

$$p_\varepsilon(0) := \mathbb{P}(\varepsilon_t = 0) = a(0)/f(\theta),$$
$$p_\eta(0) := \mathbb{P}(q_t\varepsilon_t = 0) = \mathbb{P}(q_t = 0) + \mathbb{P}(q_t = 1)\mathbb{P}(\varepsilon_t = 0) = 1 + S_\varepsilon(c)[p_\varepsilon(0) - 1].$$

From here, we get:

$$p_\eta(0) - p_\varepsilon(0) = (1 - S_\varepsilon(c))[1 - p_\varepsilon(0)] > 0,$$

whenever $0 < p_\varepsilon(0) < 1$*. Hence, RVs* (η_t) *have a more pronounced zero-inflation than* (ε_t)*. On the other hand, according to Equations (16) and (18), the stationary mass at zero for the latent process* (X_t) *is given by:*

$$p_X(0) = \mathbb{P}(X_t = 0) = \Psi_X(0) = \prod_{k=1}^{\infty} \Psi_\eta\left(\frac{1-\alpha^k}{1-\alpha^{k+1}}\right) = \prod_{k=1}^{\infty}\left[1 + S_\varepsilon(c)\left(\frac{f\left(\frac{\theta\,(1-\alpha^k)}{1-\alpha^{k+1}}\right)}{f(\theta)} - 1\right)\right],$$

where $\Psi_\eta(u) = 1 + S_\varepsilon(c)(f(\theta u)/f(\theta) - 1)$ *denotes the PGF of the gated innovations* $\eta_t = q_t\varepsilon_t$*. Moreover, using the independence of* ε_t *from* X_t*, for the series* (Y_t) *one obtains:*

$$p_Y(0) := \mathbb{P}(Y_t = 0) = \mathbb{P}(X_t = 0) + \mathbb{P}(X_t > 0, \varepsilon_t = 0) = p_X(0) + (1 - p_X(0))\, p_\varepsilon(0). \quad (39)$$

Thus, $p_Y(0) > p_X(0)$ *and* $p_Y(0) > p_\varepsilon(0)$ *hold, i.e., the MIN–SB model generates a hierarchical amplification of zeros along the mechanism: innovations* $\to$ *latent* $\to$ *observed process.*

(ii) Clustering of zeros. *Beyond marginal inflation, both processes* (X_t) *and* (Y_t) *exhibit temporal clustering of zeros. As is shown in Theorem 3, the zero-run lengths follow geometric distributions with parameters* $p_{00}^{(X)}$ *and* $p_{00}^{(Y)}$*, respectively. In particular, according to Equations (35), (37), and (39), as well as the results form Lemma A2 (see Appendix A), the transition probabilities between consecutive zeros of the series* (X_t) *and* (Y_t) *satisfy:*

$$p_{00}^{(X)} = 1 + S_\varepsilon(c)\left(\frac{c_1(\alpha,\theta)}{f(\theta)} - 1\right) > \prod_{k=1}^{\infty}\left[1 + S_\varepsilon(c)\left(\frac{c_k(\alpha,\theta)}{f(\theta)} - 1\right)\right] = p_X(0),$$

$$p_{00}^{(Y)} = p_\varepsilon(0) + (1 - p_\varepsilon(0))P(X_t = 0 \mid Y_{t-1} = 0) > p_\varepsilon(0) + (1 - p_\varepsilon(0))p_X(0) = p_Y(0),$$

where $c_k(\alpha,\theta) = f\left(\theta\,(1-\alpha^k)/(1-\alpha^{k+1})\right)$*. Therefore, the occurrence of a zero at the previous time* $(t-1)$ *increases the probability of the occurrence of another zero at the current time* (t)*.*

(iii) Quantification of zero-lengths. *According to the above results, the zero clustering arises from latent Markovian dynamics, while the minification mechanism primarily controls marginal inflation at zero. Since the zero-run lengths of both processes* (X_t) *and* (Y_t) *are geometrically distributed, zero clusters exhibit a memoryless structure despite the time dependence of the underlying process. In particular, note that from Equation (33) it follows that* $\mathbb{E}[L_0^{(X)}] = J(\alpha,\theta)/S_\varepsilon(c)$*, where*

$$J(\alpha,\theta) = \frac{f(\theta)}{f(\theta) - f(\theta/(1+\alpha))} > 0.$$

Since the survival function $S_\varepsilon(c) = \mathbb{P}(\varepsilon_t \geq c)$ *is non-increasing in* c*, the expected zero-run length* $\mathbb{E}[L_0^{(X)}]$ *is non-decreasing in the threshold parameter* c*. Consequently, different values of* c *generate different expected zero-run lengths, implying that the threshold parameter is identifiable from the zero-run statistics of the process (see Section 4). Finally, the quantity:*

$$K(\alpha,\theta) := f(\theta) - f(\theta/(1+\alpha))$$

which naturally appears in Equations (33) and (34), is an effective innovation intensity to survive the thinning mechanism and terminate zero runs. As $\alpha \to 0$, one has $K(\alpha, \theta) \to 0$, implying large expected zero-run lengths, and vice versa. Thus, the MIN–SB model contains a balance between the intensity of innovation (given by the parameter θ) and the strength of thinning (given by the parameter α), so that the interaction between them is given by $K(\alpha, \theta)$. Hence, zero inflation in the MIN–SB model arises structurally from the latent thinning and gating dynamics, which will be used in the parameter estimation procedure below.

4. Parameters Estimation Procedure

The estimation of the MIN–SB model requires special treatment due to the presence of a latent component and the non-linear minification mechanism. In addition, the ordinary moments and conditional probability distribution of the observed process (Y_t) are not given in closed form, so common estimation procedures, such as the method of moments or maximum likelihood-based procedures, cannot be used here. For these reasons, instead of relying on likelihood-based procedures, we adopt the simulated method of moments (SMM) framework that directly exploits the structural null characteristics of the model. In other words, the SMM is based on statistics that explicitly characterize the null structure of the process, such that the resulting estimators include marginal inflation, persistence, and null clustering, as explained in Remark 4. This allows for the identification of the structural parameters of the model and a natural bridge between the structural properties of the model and its practical inference. In particular, the (simulated) moments used for estimation are chosen to reflect the two fundamental mechanisms identified in Section 3, that is, marginal zero inflation and temporal clustering of zeros. This zero-based moment selection provides a natural and structurally consistent identification strategy for the vector of unknown parameters $\vartheta = (\alpha, \theta, c)^\top$. Also, it is worth noting that the SMM framework has recently been used in Markov chains and econometric theory (see, e.g., [36,37]).

4.1. Estimation via Zero-Based Simulated Moments

Let $\vartheta = (\alpha, \theta, c)^\top$ denote the parameter vector of the MIN–SB model and consider the following vector of zero-related population quantities:

$$\mathbf{m}(\vartheta) = (m_1, m_2, m_3)^\top = \left(p_Y(0),\, p_{00}^{(Y)},\, \mathbb{E}[L_0^{(Y)}]\right)^\top, \tag{40}$$

where expressions of $p_{00}^{(Y)}$, $\mathbb{E}[L_0^{(Y)}]$ and $p_Y(0)$ are given by Equations (36), (38) and (39), respectively. Note that the SMM estimator relies on the mapping:

$$(\alpha, \theta, c) \mapsto (p_Y(0), p_{00}^{(Y)}, E[L_0^{(Y)}]),$$

which is required to be injective. This is supported by the structural decomposition of the MIN–SB model. Specifically, the identification of the parameter vector (α, θ, c) is based on the fact that the selected zero-based moments capture distinct structural features of the MIN–SB process.

First, the innovation parameter θ determines the marginal distribution of the innovations (ε_t) and directly affects the overall zero-inflation level through $p_Y(0)$, as given in Equation (39). Moreover, according to Equation (7) and Remark 4 (item (i)), it follows that, for fixed (α, c), the probability $p_Y(0)$ is a monotone function of θ. Hence, the mapping $\theta \mapsto p_Y(0)$ is injective, ensuring identification of the innovation parameter θ. Second, the NB thinning parameter α governs the temporal dependence structure of the latent process, and thus primarily affects the zero-to-zero transition probability $p_{00}^{(Y)}$, given by Equation (36). Changes in α alter the clustering of zeros over time, which cannot be repli-

cated by adjustments in either c or θ, as these parameters influence marginal and threshold behavior rather than temporal dependence. Finally, as shown in Theorem 3 and Remark 4 (item (iii)), the expected zero-run length $E[L_0^{(Y)}]$ is a monotone function of the survival probability $S_\varepsilon(c)$, implying that different values of c induce distinct zero-run behavior and thus ensure identification of the threshold parameter. Taken jointly, these arguments imply that the mapping $\mathbf{m}(\vartheta)$, defined by Equation (40), is locally injective under regularity conditions. Hence, these properties ensure local identifiability of the parameter vector ϑ, which justifies the identification assumption used in the following (see Theorem 4 below). In particular, each parameter influences the selected moments in a structurally distinct way, preventing observational equivalence between different parameter configurations.

Further, for the given observation $Y_1, \ldots, Y_T$ of the MIN–SB process, the empirical moment vector is given by:

$$\widehat{\mathbf{m}}_T = \left(\widehat{p}_Y(0), \widehat{p}_{00}^{(Y)}, \widehat{L}_0^{(Y)}\right)^\top,$$

where:

$$\widehat{p}_Y(0) = \frac{1}{T}\sum_{t=1}^{T} I(Y_t = 0), \qquad \widehat{p}_{00}^{(Y)} = \frac{\sum_{t=2}^{T} I(Y_{t-1} = 0, Y_t = 0)}{\sum_{t=2}^{T} I(Y_{t-1} = 0)},$$

and

$$\widehat{L}_0^{(Y)} = \frac{\sum_{t=1}^{T} I(Y_t = 0)}{\sum_{t=2}^{T} I(Y_{t-1} > 0,\ Y_t = 0)}$$

is the estimator equivalent to the sample mean of all identified zero-run lengths, since the total number of zeros equals the sum of zero-run lengths. Hence, the objective function is defined as:

$$Q_T(\vartheta) = \big(\widehat{\mathbf{m}}_T - \mathbf{m}(\vartheta)\big)^\top W \big(\widehat{\mathbf{m}}_T - \mathbf{m}(\vartheta)\big),$$

where W is a symmetric positive definite weighting matrix. Thus, the appropriate SMM estimator $\widehat{\vartheta}_T = (\widehat{\alpha}, \widehat{\theta}, \widehat{c})^\top$ is obtained via a profiled minimization procedure:

$$\begin{cases} \widehat{c} & = \arg\min\limits_{c \in \mathcal{C}} \inf\limits_{(\alpha,\theta)\in\Theta_{\alpha,\theta}} Q_T(\alpha, \theta, c), \\ (\hat{\alpha}, \hat{\theta})^\top & = \arg\min\limits_{(\alpha,\theta)\in\Theta_{\alpha,\theta}} Q_T(\alpha, \theta, \widehat{c}), \end{cases} \tag{41}$$

where $\Theta_{\alpha,\theta} \subset (0,1) \times (0,\infty)$ and the threshold parameter c is selected over a finite grid $\mathcal{C} \subset \mathbb{N}_0$. Under certain regularity conditions, the consistency of the proposed estimators, as well as the asymptotic normality (AN) of continuous parameters, can be established as follows:

Theorem 4. *Let (Y_t), $t = 1, \ldots, T$ be observations of a stationary MIN–SB process, with true parameter vector $\vartheta_0 = (\alpha_0, \theta_0, c_0)^\top$, and let the SMM estimator $\widehat{\vartheta}_T = (\hat{\alpha}, \hat{\theta}, \hat{c})$ be defined as in Equation (41). In addition, assume that the following conditions hold:*

(A1) The parameter space satisfies $\Theta = \Theta_{\alpha,\theta} \times \mathcal{C}$, where $\Theta_{\alpha,\theta}$ is a compact set and (α_0, θ_0) lies in its interior, while $c_0 \in \mathcal{C}$.

(A2) The vector $\mathbf{m}(\alpha, \theta, c)$, defined in Equation (40), is continuously differentiable with respect to (α, θ) in a neighborhood of (α_0, θ_0), for each fixed $c \in \mathcal{C}$.

(A3) The mapping $\vartheta \mapsto \mathbf{m}(\vartheta)$ is locally injective in a neighborhood of ϑ_0.

(A4) The profiled population criterion function:

$$Q(\vartheta) = \big(\mathbf{m}(\vartheta_0) - \mathbf{m}(\vartheta)\big)^\top W \big(\mathbf{m}(\vartheta_0) - \mathbf{m}(\vartheta)\big).$$

is separated at the true threshold value c_0, i.e., $\inf_{(\alpha,\theta)\in\Theta_{\alpha,\theta}} Q(\vartheta) > 0$, for any $c \neq c_0$.

(A5) *The Jacobian matrix* $D(\vartheta) = \partial \mathbf{m}(\vartheta)/\partial(\alpha, \theta)^\top$ *has full column rank at the true parameter value* $\vartheta = \vartheta_0$.

Then the following holds:

(i) *The estimator* $\hat{c}$ *is consistent in the sense that* $\mathbb{P}(\hat{c} = c_0) \to 1$, *as* $T \to \infty$.

(ii) *Conditional on correct threshold selection* ($\hat{c} = c_0$), *the estimator* $(\hat{\alpha}, \hat{\theta})^\top$ *is strongly consistent and AN for the parameter* $(\alpha_0, \theta_0)^\top$.

Proof. The observed process (Y_t) is strongly stationary and ergodic, being a measurable transformation of the geometrically ergodic Markov chain $(\mathbf{Z}_t)$, established in Theorem 1. Hence, each component of the empirical moment vector $\widehat{\mathbf{m}}_T$ is an additive functional of a stationary ergodic process. Therefore, by the ergodic theorem (see, e.g., Jones [38]),

$$\widehat{\mathbf{m}}_T \xrightarrow{a.s.} \mathbf{m}(\vartheta_0), \qquad T \to \infty. \tag{42}$$

Further, by assumptions (A1)–(A2), for each fixed $c \in \mathcal{C}$ the moment function $\mathbf{m}(\vartheta)$ is continuous in (α, θ) on the compact set $\Theta_{\alpha,\theta}$. Along with the almost sure convergence established in Equation (42), this implies that $Q_T(\alpha, \theta, c)$ converges to $Q(\alpha, \theta, c)$ uniformly in (α, θ), for each fixed $c \in \mathcal{C}$. Thus, since $\mathcal{C}$ is finite, the convergence is uniform on $\Theta = \Theta_{\alpha,\theta} \times \mathcal{C}$. Moreover, by assumption (A3), for the true threshold value $c = c_0$, the population criterion $Q(\vartheta)$ attains a unique local minimum $Q(\vartheta_0) = 0$, in a neighborhood of $(\alpha_0, \theta_0) \in \Theta_{\alpha,\theta}$. Additionally, the separation condition (A4) ensures that:

$$\inf_{(\alpha,\theta)\in\Theta_{\alpha,\theta}} Q(\alpha, \theta, c) > Q(\alpha_0, \theta_0, c_0) = 0, \qquad c \neq c_0,$$

so the profiled population criterion admits a unique minimum at the true value c_0. By applying the argmin theorem for M-estimators under uniform convergence (see, e.g., [39], Theorem 2.1), and noting that $\mathcal{C}$ is finite so that the profiled problem reduces to a finite collection of standard M-estimation problems indexed by $c \in \mathcal{C}$, it follows that:

$$\mathbb{P}(\hat{c} = c_0) \to 1, \qquad T \to \infty,$$

and, conditional on $\hat{c} = c_0$,

$$(\hat{\alpha}, \hat{\theta})^\top \xrightarrow{a.s.} (\alpha_0, \theta_0)^\top, \qquad T \to \infty.$$

This proves strong consistency for the SMM estimator $\widehat{\vartheta}_T$. In addition, with probability tending to one, the estimation problem reduces to a standard SMM problem for the continuous parameter vector (α, θ) evaluated at the true threshold parameter c_0.

We now establish the AN property for $(\hat{\alpha}, \hat{\theta})^\top$. Since $(\mathbf{Z}_t)$ is geometrically ergodic, standard central limit theorems for additive functionals of geometrically ergodic Markov chains (see, e.g., [38]) imply:

$$\sqrt{T}\big(\widehat{\mathbf{m}}_T - \mathbf{m}(\vartheta_0)\big) \xrightarrow{d} \mathcal{N}(0, \Sigma), \qquad T \to \infty, \tag{43}$$

where Σ denotes the long-run covariance matrix of the moment process. Further, according to assumption (A2), the moment function $\mathbf{m}(\alpha, \theta, c_0)$ admits a first-order Taylor expansion:

$$\mathbf{m}(\hat{\alpha}, \hat{\theta}, c_0) = \mathbf{m}(\alpha_0, \theta_0, c_0) + D(\alpha_0, \theta_0, c_0)\begin{pmatrix} \hat{\alpha} - \alpha_0 \\ \hat{\theta} - \theta_0 \end{pmatrix} + o_p(\|(\hat{\alpha}, \hat{\theta}) - (\alpha_0, \theta_0)\|), \tag{44}$$

where $D(\alpha,\theta,c) = \partial \mathbf{m}(\alpha,\theta,c)/\partial(\alpha,\theta)^\top$. In addition, for fixed $c = c_0$, differentiating the criterion function $Q_T(\vartheta)$ with respect to (α,θ) gives:

$$\frac{\partial Q_T(\alpha,\theta,c_0)}{\partial(\alpha,\theta)^\top} = -2D(\alpha,\theta,c_0)^\top W\big(\widehat{\mathbf{m}}_T - \mathbf{m}(\alpha,\theta,c_0)\big).$$

Since $(\widehat{\alpha},\widehat{\theta})$ minimizes $Q_T(\alpha,\theta,\widehat{c})$ and $\mathbb{P}(\widehat{c} = c_0) \to 1$, it follows that, on an event with probability tending to one,

$$0 = D(\widehat{\alpha},\widehat{\theta},c_0)^\top W\big(\widehat{\mathbf{m}}_T - \mathbf{m}(\widehat{\alpha},\widehat{\theta},c_0)\big). \tag{45}$$

Moreover, by consistency of $(\widehat{\alpha},\widehat{\theta})$ and continuity of $D(\alpha,\theta,c)$, we have:

$$D(\widehat{\alpha},\widehat{\theta},c_0) = D(\alpha_0,\theta_0,c_0) + o_p(1),$$

so that, asymptotically, $D(\alpha_0,\theta_0,c_0)$ may be used in place of $D(\widehat{\alpha},\widehat{\theta},c_0)$. Thus, we further denote $D := D(\alpha_0,\theta_0,c_0)$, and substituting Equation (44) into (45) yields:

$$0 = D^\top W\left[\widehat{\mathbf{m}}_T - \mathbf{m}(\vartheta_0) - D\begin{pmatrix}\widehat{\alpha}-\alpha_0\\ \widehat{\theta}-\theta_0\end{pmatrix}\right] + o_p(T^{-1/2}).$$

By rearranging some terms above, one obtains:

$$D^\top WD\begin{pmatrix}\widehat{\alpha}-\alpha_0\\ \widehat{\theta}-\theta_0\end{pmatrix} = D^\top W\big(\widehat{\mathbf{m}}_T - \mathbf{m}(\vartheta_0)\big) + o_p(T^{-1/2}).$$

where, by assumption (A5), the matrix $D^\top WD$ is nonsingular. Hence, multiplying both sides by $\sqrt{T}$ gives:

$$\sqrt{T}\begin{pmatrix}\widehat{\alpha}-\alpha_0\\ \widehat{\theta}-\theta_0\end{pmatrix} = (D^\top WD)^{-1}D^\top W\sqrt{T}\big(\widehat{\mathbf{m}}_T - \mathbf{m}(\vartheta_0)\big) + o_p(1).$$

Finally, applying the central limit theorem from Equation (43) and Slutsky's theorem (see, e.g., [40], Ch. 6), yields:

$$\sqrt{T}\begin{pmatrix}\widehat{\alpha}-\alpha_0\\ \widehat{\theta}-\theta_0\end{pmatrix} \xrightarrow{d} \mathcal{N}\Big(0,(D^\top WD)^{-1}D^\top W\Sigma WD(D^\top WD)^{-1}\Big), \qquad T\to\infty.$$

Since $\mathbb{P}(\widehat{c} = c_0) \to 1$, the above asymptotic distribution holds on an event whose probability tends to one, which completes the proof. □

Remark 5 (Interpretation and implementation of the SMM estimator)**.** *The weighting matrix W plays a central role in the efficiency of the SMM estimator. In the simplest implementation, W may be chosen as the identity matrix, corresponding to equally weighted matching of the zero-related moment conditions. More generally, alternative choices allow different emphasis to be placed on marginal zero inflation versus dynamic clustering features. In addition, the asymptotic covariance matrix Σ, defined in Equation (43), admits the standard long-run representation associated with additive functionals of geometrically ergodic Markov chains. Although its analytical form is generally cumbersome, Σ can be consistently estimated using HAC (heteroskedasticity and autocorrelation consistent) estimators, which remain valid under serial dependence and conditional heteroskedasticity. The asymptotically efficient choice of the weighting matrix is $W = \Sigma^{-1}$, yielding the efficient SMM estimator with asymptotic variance $(D^\top\Sigma^{-1}D)^{-1}$. In practice, Σ is replaced by a consistent estimator $\widehat{\Sigma}_T$, leading to a two-step efficient SMM procedure. This optimal weighting scheme is used*

throughout the MC study in Section 5. Finally, by choosing moments based on zero frequencies and zero-run behavior, the SMM estimator avoids the instability associated with higher-order cumulants. In particular, the zero-based moments remain numerically stable even under heavy-tailed innovation distributions, increasing the robustness of the procedure.

4.2. Estimation Based on Some Specific PS Innovations

In this section, we consider two important special cases of the PS innovations used in the MIN–SB model, namely the geometric and Poisson distributions. The goal is to provide explicit expressions for the probability generating functions (PGFs), first and second moments, and zero-related quantities relevant for parameter estimation. In the following, using these specific PS distributions, the simulation study and practical application of the MIN–SB model are carried out. To this end, recall that the PS innovations (ε_t) follow a distribution with the PGF $\Psi_\varepsilon(u) = f(\theta u)/f(\theta)$, and that the gated innovations (η_t) are defined as $\eta_t = q_t\varepsilon_t$, where $q_t = I(\varepsilon_{t-1} \geq c)$. Here, the threshold parameter $c \in \mathbb{N}_0$ controls the probability $S_\varepsilon(c) = \mathbb{P}(\varepsilon_t \geq c)$, which plays a key role in the zero dynamics of the MIN–SB model. Finally, recall that the zero statistics used in estimation are given within the vector $\mathbf{m}(\vartheta)$, defined by Equation (40), where the general relations are defined by Equations (36)–(39).

(i) Geometric innovations. Let us assume that (ε_t) follow a geometric distribution on $\mathbb{N}_0$, which implies (see Table 1) $\mathbb{P}(\varepsilon_t = k) = (1-\theta)\theta^k$, where $0 < \theta < 1$, and $k = 0, 1, 2, \ldots$ Then, $f(\theta) = (1-\theta)^{-1}$, $\Psi_\varepsilon(u) = (1-\theta)/(1-\theta u)$, and hence $\mu_\varepsilon = \mathbb{E}[\varepsilon_t] = \theta/(1-\theta)$, $\sigma_\varepsilon^2 = \mathrm{Var}[\varepsilon_t] = \theta/(1-\theta)^2$. Further, $p_\varepsilon(0) = 1-\theta$, and the survival probability entering the MIN–SB dynamics is $S_\varepsilon(c) = \theta^c$, $c = 0, 1, 2, \ldots$ Substituting them into Equation (35) yields:

$$p_{00}^{(X)} = 1 + \theta^c\left[\frac{1-\theta}{1-\theta/(1+\alpha)} - 1\right] = 1 - \frac{\alpha\theta^{c+1}}{1+\alpha-\theta}. \tag{46}$$

Consequently, according to Remark 4, the zero probability of the gated innovation becomes

$$p_\eta(0) = 1 + S_\varepsilon(c)\big(p_\varepsilon(0) - 1\big) = 1 - \theta^{c+1},$$

and after some computations, the explicit form of $p_X(0)$ and $p_Y(0)$ is as follows:

$$p_X(0) = \Psi_X(0) = \prod_{k=1}^{\infty}\Psi_\eta\left(\frac{1-\alpha^k}{1-\alpha^{k+1}}\right) = \prod_{k=1}^{\infty}\left[1 + S_\varepsilon(c)\left(\Psi_\varepsilon\left(\frac{1-\alpha^k}{1-\alpha^{k+1}}\right) - 1\right)\right]$$
$$= \prod_{k=1}^{\infty}\left[1 - \frac{\alpha^k(1-\alpha)\theta^{c+1}}{1-\alpha^{k+1}-\theta(1-\alpha^k)}\right], \qquad p_Y(0) = (1-\theta) + \theta\, p_X(0).$$

Moreover, by using Equations (36) and (46), for the zero-to-zero transition probability of the observed process (Y_t), one obtains:

$$p_{00}^{(Y)} = \frac{(1-\theta)\Big[1-\theta+\theta\,\Psi_X\big((1+\alpha)^{-1}\big)\Big] + p_X(0)\big(\theta - \alpha\theta^{c+2}/(1+\alpha-\theta)\big)}{1-\theta+\theta p_X(0)}.$$

as the expected length of zeros is:

$$E[L_0^{(Y)}] = \frac{1-\theta+\theta p_X(0)}{(1-\theta)\theta\big(1-\Psi_X\big((1+\alpha)^{-1}\big)\big) + \alpha\theta^{c+2}/(1+\alpha-\theta)\, p_X(0)}.$$

Finally, using Equation (20) and PGF $\Psi_X(u)$, the mean of the observed process (Y_t) can be obtained as follows:

$$\mathbb{E}[Y_t] = \mathbb{E}\left[\sum_{j=1}^{X_t} S_\varepsilon(j)\right] = \mathbb{E}\left[\sum_{j=1}^{X_t} \theta^j\right] = \frac{\theta}{1-\theta}\mathbb{E}\left[1-\theta^{X_t}\right] = \frac{\theta}{1-\theta}\left[1-\Psi_X(\theta)\right]. \tag{47}$$

This representation shows that the mean of the observed process is fully driven by the latent PGF evaluated at the innovation parameter, highlighting the structural role of the thinning mechanism. Similarly, for the second moment, one obtains:

$$\mathbb{E}[Y_t^2] = \mathbb{E}\left[\sum_{j=1}^{X_t}(2j-1)\theta^j\right] = 2A - B,$$

where $A = \mathbb{E}\left[\sum_{j=1}^{X_t} j\theta^j\right]$, $B = \mathbb{E}\left[\sum_{j=1}^{X_t}\theta^j\right] = \mathbb{E}[Y_t]$. Hence, the term B is directly obtained from (47), while applying Lemma A3 from Appendix A (cf. Equation (A2)) gives:

$$A = \mathbb{E}\left[\sum_{j=1}^{\infty} j\,\theta^j\, I(X_t \geq j)\right] = \sum_{j=1}^{\infty} j\,\theta^j\,\mathbb{P}(X_t \geq j) = \frac{\theta}{(1-\theta)^2}\left[1-\Psi_X(\theta)\right] - \frac{\theta^2}{1-\theta}\,\Psi'_X(\theta).$$

Thus,

$$\mathbb{E}[Y_t^2] = \frac{\theta(1+\theta)}{(1-\theta)^2}\left[1-\Psi_X(\theta)\right] - \frac{2\theta^2}{1-\theta}\Psi'_X(\theta), \tag{48}$$

and by using Equations (47) and (48), the variance of (Y_t) is obtained as follows:

$$\mathrm{Var}[Y_t] = \mathbb{E}[Y_t^2] - \left[\mathbb{E}[Y_t]\right]^2 = \frac{\theta}{(1-\theta)^2}\left(1-\Psi_X(\theta)\right)\left(1+\theta\,\Psi_X(\theta)\right) - \frac{2\theta^2}{1-\theta}\Psi'_X(\theta).$$

(ii) Poisson innovations. Let us assume now that $\varepsilon_t \sim \mathrm{Pois}(\theta)$. Thus, the corresponding PGF is $\Psi_\varepsilon(u) = \exp(\theta(u-1))$, with $\mathbb{E}[\varepsilon_t] = \mathrm{Var}(\varepsilon_t) = \theta$. The zero probability and survival function are, respectively, $p_\varepsilon(0) = \mathrm{e}^{-\theta}$, $S_\varepsilon(c) = 1 - \mathrm{e}^{-\theta}\sum_{k=0}^{c-1}\theta^k/k!$, and substituting into Equation (35) gives:

$$p_{00}^{(X)} = 1 + \left(1 - e^{-\theta}\sum_{k=0}^{c-1}\frac{\theta^k}{k!}\right)\left(\mathrm{e}^{-\frac{\alpha\theta}{1+\alpha}} - 1\right).$$

Using the same procedure as above, for the key zero statistics used in the estimation procedure, one obtains:

$$p_Y(0) = \mathrm{e}^{-\theta} + (1-\mathrm{e}^{-\theta})\,p_X(0),$$

$$p_{00}^{(Y)} = \frac{\mathrm{e}^{-2\theta} + (1-\mathrm{e}^{-\theta})\left[p_X(0)\left(1-S_\varepsilon(c)\right)\left(1-\mathrm{e}^{-\frac{\alpha\theta}{1+\alpha}}\right) + \mathrm{e}^{-\theta}\,\Psi_X\left((1+\alpha)^{-1}\right)\right]}{\mathrm{e}^{-\theta} + (1-\mathrm{e}^{-\theta})p_X(0)},$$

$$\mathbb{E}[L_0^{(Y)}] = \frac{\mathrm{e}^{-\theta} + (1-\mathrm{e}^{-\theta})p_X(0)}{(1-\mathrm{e}^{-\theta})\left[p_X(0)S_\varepsilon(c)\left(1-\mathrm{e}^{-\frac{\alpha\theta}{1+\alpha}}\right) + \mathrm{e}^{-\theta}\,\Psi_X\left((1+\alpha)^{-1}\right)\right]}.$$

The mean of the series (Y_t) with Poisson innovations can be obtained as follows:

$$\mathbb{E}(Y_t) = \mathbb{E}[\min(X_t, \varepsilon_t)] = \sum_{j=1}^{\infty}\mathbb{P}(X_t \geq j)\mathbb{P}(\varepsilon_t \geq j) = \sum_{j=1}^{\infty} S_\varepsilon(j)\,S_X(j),$$

which reveals that the observed mean results from the interaction between the survival functions of the latent and innovation components. Finally, the variance is obtained by the same procedure as in the proof of Theorem 2. Note that in contrast to the geometric case, no closed-form PGF reduction is available, and numerical evaluation is performed via truncation of the tail sums.

Remark 6 (Estimation insight for the threshold parameter)**.** *Recall again that the threshold parameter c governs the frequency of gated immigration through the survival probability $S_\varepsilon(c)$. Since c is integer-valued, estimation is naturally implemented via a grid search over admissible values of c, while the remaining parameters are estimated conditionally. This separation substantially simplifies the numerical implementation of the SMM procedure. For instance, in the geometric case, the survival function is $S_\varepsilon(c) = \theta^c$, and combining this with Equation (46) yields:*

$$\theta^c = \left(1 - p_{00}^{(X)}\right) W(\alpha, \theta),$$

where $W(\alpha, \theta) = (\alpha\theta)^{-1}(1 + \alpha - \theta)$. From here, taking logarithms leads to the continuous relaxation:

$$c \approx \frac{\log\left(\left(1 - p_{00}^{(X)}\right) W(\alpha, \theta)\right)}{\log(\theta)}.$$

Hence, although c is integer-valued, this expression provides a convenient heuristic starting value around which the grid search can be centered, thereby improving numerical efficiency.

5. Monte Carlo Study

The theoretical properties of the SMM estimators established above provide the asymptotic foundation of the zero-based SMM estimator, while the explicit zero-run structure guarantees identification of the structural parameters. To assess the finite-sample behavior of the estimator and to examine the speed of convergence toward the asymptotic regime, we now conduct Monte Carlo (MC) simulations of the MIN–SB process under different innovation specifications and parameter configurations. The simulation study is designed to complement the theoretical results established in Section 4, where consistency and AN property of the SMM estimators were derived under geometric ergodicity of the underlying Markov structure. The main goals of the MC analysis are threefold:

1. to evaluate finite-sample bias and dispersion of the parameter estimates;
2. to examine how quickly the AN property appears;
3. to assess the practical identifiability of the threshold parameter.

The simulation design, parameter estimation procedure, and the results thus obtained are described in more detail below.

5.1. Simulation Design

The data generating procedure for the observation $Y_1, Y_2, \ldots, Y_T$ follows the stationary MIN–SB model defined by Equations (9)–(11). To this end, two innovation structures are considered, corresponding to the explicit formulas derived in Section 4.2:

- geometric innovations $\varepsilon_t \sim \text{Geom}(\theta)$;
- Poisson innovations, $\varepsilon_t \sim \text{Pois}(\theta)$.

For geometric innovations, the true parameter vector is fixed at $(\alpha, \theta, c) = (0.6, 0.6, 1)$, while for Poisson innovations, the true parameter vector is $(\alpha, \theta, c) = (0.6, 2, 1)$. These values correspond to a moderately persistent latent process with non-negligible zero inflation and clearly active threshold dynamics, thus representing a challenging but realistic estimation scenario. To study convergence towards the asymptotic regime predicted by

Theorem 4, four sample sizes are considered, with the lengths $T \in \{50, 150, 500, 1500\}$. For each configuration, $M = 300$ independent MC replications are generated using standard routines in the 'R' statistical software environment (version 4.5.2).

The estimation is performed using a moments-based approach that relies on the zero-related statistics of the observed process (Y_t). Specifically, the estimation procedure follows the zero-based simulated moments framework presented in Section 4.1, i.e., the empirical moment vector $\widehat{\mathbf{m}}_T$ consists of:

1. estimator of marginal probability of zero $\widehat{p}_Y(0)$;
2. estimator of zero-to-zero transition probability $\widehat{p}_{00}^{(Y)}$;
3. estimator of the expected length of zero runs $\widehat{L}_0^{(Y)}$.

These statistics directly reflect the structural decomposition of the MIN–SB model discussed in Section 3 (cf. Theorem 3), where zero inflation and temporal dependence arise from distinct components of the latent dynamics. Hence, the chosen moments provide an empirically stable alternative to higher-order moments, whose estimation is typically unreliable in zero-inflated count processes.

In addition, due to the discrete and non-smooth role of the threshold parameter c, a special profile SMM strategy is adopted for it. More precisely, a finite grid $\mathcal{C} = \{0, 1, 2, 3\}$ is considered, and for each fixed $c \in \mathcal{C}$ the remaining parameters (α, θ) are estimated by minimizing the SMM criterion function:

$$Q_T(\alpha, \theta \,|\, c) = \big(\widehat{\mathbf{m}}_T - \mathbf{m}(\alpha, \theta, c)\big)^\top W \big(\widehat{\mathbf{m}}_T - \mathbf{m}(\alpha, \theta, c)\big),$$

where $\mathbf{m}(\alpha, \theta, c)$ is its theoretical counterpart. The threshold estimate $\hat{c}$ is then selected as the grid value minimizing the profiled criterion function, and the final estimates $(\hat{\alpha}, \hat{\theta})$ are obtained conditionally on $\hat{c}$. In doing so, note that the estimation grid includes the boundary value $c = 0$, allowing for the possibility that the threshold mechanism is inactive. Further, the weighting matrix W is chosen as the inverse of the estimated long-run covariance matrix of the moment process, yielding the efficient simulated moments estimator discussed in Remark 5.

For each fixed value of the threshold parameter c, the continuous parameters were estimated using the `optim()` routine in the R environment with the Nelder–Mead simplex algorithm [41]. Hence, this procedure separates the discrete optimization over c from the smooth optimization over (α, θ), leading to improved numerical stability and clearer finite-sample behavior. The final threshold estimate is then obtained as the grid value minimizing the profiled criterion, while the continuous parameter estimates are obtained conditionally on this choice. For each MC simulation, descriptive statistics of obtained SMM estimates (Min, Mean, Max) are calculated, along with the corresponding estimation errors, i.e., bias, standard deviation (StDev), and root mean square error (RMSE). To further examine the asymptotic properties of the SMM estimates, normality was assessed using the Jarque–Bera (JB), Shapiro–Wilk (SW), and Anderson–Darling (AD) tests. While the JB test is reported for completeness, the primary conclusions are based on the SW and AD tests, which exhibit superior performance in finite samples. Finally, test statistics and the corresponding p-values are calculated using software procedures from the R package "nortest" [42].

5.2. Monte Carlo Results

Tables 2 and 3 report MC results for smaller sample sizes ($T = 50$ and $T = 150$), providing insight into the finite-sample behavior of the estimator. As expected, the estimators exhibit increased variability in shorter samples, with wider dispersion and slightly higher bias, particularly for the innovation parameter (θ) and the threshold parameter

(c). This effect is more pronounced for $T = 50$, reflecting the limited information content of short series. Nevertheless, the estimators of the continuous parameters (α, θ) remain reasonably accurate under both innovation specifications, with biases that are relatively small compared to their standard deviations. RMSE values are close to the corresponding SDs, indicating that estimation errors are primarily driven by sampling variability rather than systematic bias. Normality diagnostics suggest that, even in small samples, the sampling distributions of the continuous estimators approaching Gaussian behavior, although some deviations remain visible. In contrast, the threshold estimator $\hat{c}$ exhibits the expected discrete behavior induced by grid-based selection, and its performance is therefore better interpreted through selection frequencies (for that purpose, see Table 6).

Table 2. MC results for parameter estimates of the MIN–SB model ($T = 50$).

Statistic	Geometric Innovations			Poisson Innovations		
	$\hat{\alpha}$	$\hat{\theta}$	$\hat{c}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{c}$
Min.	0.3755	0.4321	0.0000	0.3079	1.3626	0.0000
Mean	0.5760	0.6279	1.3267	0.5851	1.9805	1.5033
Max.	1.0612	0.8570	3.0000	0.9175	2.3546	3.0000
Bias	−0.0240	0.0279	0.3267	−0.0149	−0.0195	0.5033
SD	0.1472	0.1103	0.8628	0.1376	0.2047	0.9126
RMSE	0.1489	0.1136	0.8907	0.1382	0.2053	1.0480
JB stat.	8.4902 *	7.7736 *	–	4.2482	3.1075	–
(p-values)	(0.0143)	(0.0205)	–	(0.1195)	(0.2115)	–
SW stat.	0.9873	0.9810 *	–	0.9860 **	0.9890 *	–
(p-values)	(0.1073)	(0.0203)	–	(0.0051)	(0.0232)	–
AD stat.	1.0428 **	0.6534	–	0.7723 *	0.8895 *	–
(p-values)	(0.0095)	(0.0873)	–	(0.0444)	(0.0228)	–

* $0.01 < p < 0.05$; ** $0.00 < p < 0.01$.

Table 3. MC results for parameter estimates of the MIN–SB model ($T = 150$).

Statistic	Geometric Innovations			Poisson Innovations		
	$\hat{\alpha}$	$\hat{\theta}$	$\hat{c}$	$\hat{\alpha}$	$\hat{\theta}$	$\hat{c}$
Min.	0.3882	0.4454	0.0000	0.3163	1.4700	0.0000
Mean	0.5941	0.6041	1.2833	0.5926	1.9851	1.4833
Max.	0.9498	0.8210	3.0000	0.8007	2.3440	3.0000
Bias	−0.0059	0.0041	0.2833	−0.0074	−0.0149	0.4833
SD	0.0931	0.0673	0.8143	0.0848	0.1709	0.8359
RMSE	0.0931	0.0687	0.8145	0.0860	0.1713	0.9644
JB stat.	7.2348 *	4.1846	–	5.5062	4.3984	–
(p-values)	(0.0269)	(0.1234)	–	(0.0637)	(0.1109)	–
SW stat.	0.9883	0.9925 *	–	0.9896 *	0.9883 *	–
(p-values)	(0.2043)	(0.0107)	–	(0.0312)	(0.0158)	–
AD stat.	0.6935	0.5476	–	0.8112 *	0.8345 *	–
(p-values)	(0.0695)	(0.1577)	–	(0.0355)	(0.0311)	–

* $0.01 < p < 0.05$.

In addition, Tables 4 and 5 present results for larger sample sizes, where the transition toward the asymptotic regime becomes evident. For both innovation specifications, systematic deviations of the continuous parameter estimators become negligible, while sampling variability decreases substantially, in line with the expected $T^{-1/2}$ convergence rate. The empirical distributions approach Gaussian shapes, supporting the AN established in Theorem 4. Geometric innovations exhibit slightly faster stabilization, whereas

Poisson innovations display somewhat higher variability, reflecting their inherently larger dispersion. As expected, the threshold estimator $\hat{c}$ remains discrete and does not satisfy the AN condition. However, its distribution becomes increasingly concentrated around the true value as the sample size grows, confirming the stability of the grid-based selection procedure. Taken together, the MC results reveal a clear convergence pattern. As the sample size increases, biases rapidly vanish and variability decreases, while the empirical distributions of the continuous estimators evolve toward symmetric Gaussian forms, in agreement with the asymptotic theory. Overall, the results confirm that the proposed SMM estimator achieves stable and accurate inference across a wide range of sample sizes.

Table 4. MC results for parameter estimates of the MIN–SB model ($T = 500$).

Statistic	Geometric Innovations			Poisson Innovations		
	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\theta}}$	$\hat{c}$	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\theta}}$	$\hat{c}$
Min.	0.4740	0.4919	0.0000	0.4377	1.6031	0.0000
Mean	0.5987	0.6036	0.9333	0.5983	1.9924	1.4033
Max.	0.7562	0.7140	3.0000	0.7548	2.3336	3.0000
Bias	−0.0013	0.0036	−0.0667	−0.0017	−0.0076	0.4033
SD	0.0504	0.0409	0.7466	0.0450	0.1358	0.7976
RMSE	0.0503	0.0410	0.7483	0.0450	0.1358	0.8926
JB stat.	2.4406	3.8278	–	3.6872	5.9004	–
(p-values)	(0.2951)	(0.1475)	–	(0.1582)	(0.0523)	–
SW stat.	0.9894	0.9922 *	–	0.9913	0.9922 *	–
(p-values)	(0.0977)	(0.0396)	–	(0.0743)	(0.0396)	–
AD stat.	0.5266	0.6458	–	0.6203	0.7781 *	–
(p-values)	(0.1780)	(0.0912)	–	(0.1055)	(0.0429)	–

* $0.01 < p < 0.05$.

Table 5. MC results for parameter estimates of the MIN–SB model ($T = 1500$).

Statistic	Geometric Innovations			Poisson Innovations		
	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\theta}}$	$\hat{c}$	$\hat{\boldsymbol{\alpha}}$	$\hat{\boldsymbol{\theta}}$	$\hat{c}$
Min.	0.5182	0.5048	0.0000	0.5306	1.6767	0.0000
Mean	0.5998	0.6003	0.9400	0.5996	2.0013	1.1900
Max.	0.6871	0.6816	3.0000	0.6665	2.3368	3.0000
Bias	−0.0002	0.0003	−0.0600	−0.0004	0.0013	0.1900
SD	0.0300	0.0321	0.6667	0.0247	0.1134	0.7786
RMSE	0.0299	0.0321	0.6683	0.0246	0.1132	0.8406
JB stat.	2.9855	2.5066	–	0.4059	1.7559	–
(p-values)	(0.2248)	(0.2856)	–	(0.8163)	(0.4156)	–
SW stat.	0.9948	0.9938	–	0.9966	0.9944	–
(p-values)	(0.6601)	(0.2314)	–	(0.7675)	(0.3463)	–
AD stat.	0.4696	0.5819	–	0.1608	0.2896	–
(p-values)	(0.2460)	(0.1290)	–	(0.9474)	(0.6115)	–

In contrast, the threshold estimator $\hat{c}$ reflects its discrete, grid-based construction and is therefore more appropriately interpreted through selection frequencies rather than normality-based arguments (see Table 6). Nevertheless, the MC results indicate that the threshold estimator $\hat{c}$ may exhibit somewhat greater finite-sample variability than the continuous parameter estimators. Therefore, neighboring threshold values should be interpreted with some caution in empirical applications, particularly for moderate sample sizes, even when they induce qualitatively similar zero-regime dynamics. This behavior is also reflected in the stability of the fitted zero-based statistics across nearby values of the

threshold parameter, highlighting the role of c as a regime indicator governing the activation of gated innovations rather than as a continuously varying parameter. The insensitivity to small deviations in $\hat{c}$ further supports the robustness of its structural interpretation in finite samples. From a practical perspective, inference on c is more robustly based on its empirical selection profile across the grid. This feature is characteristic of discrete structural parameters estimated via grid-search procedures in moment-based frameworks, and is consistent with analogous threshold parameters in nonlinear autoregressive models (see, e.g., [43,44]). As the sample size increases, the selection frequencies become increasingly concentrated around the true value, providing empirical evidence of the consistency of the proposed estimation procedure.

To further investigate the distributional properties of the estimators, Figure 3 displays Q–Q plots of the estimates of α and θ against the corresponding Gaussian quantiles. The plots show an approximately linear pattern in the central region, suggesting that the AN approximation provides a reasonable description of the finite-sample behavior of the continuous parameters. In doing so, minor deviations in the tails reflect finite-sample effects induced by the threshold mechanism and the zero-inflated structure, rather than systematic departures from AN. In contrast, the Q–Q plots corresponding to the threshold estimator $\hat{c}$ show a characteristic step-like pattern, which is a direct consequence of the grid-based estimation procedure and the discrete nature of the parameter space. Since c is integer and is chosen by minimizing a profile over a finite grid, its sampling distribution is inherently discrete. Therefore, AN does not apply to $\hat{c}$, and its behavior on the finite sample is more conveniently interpreted in terms of selection frequencies than through a normal approximation.

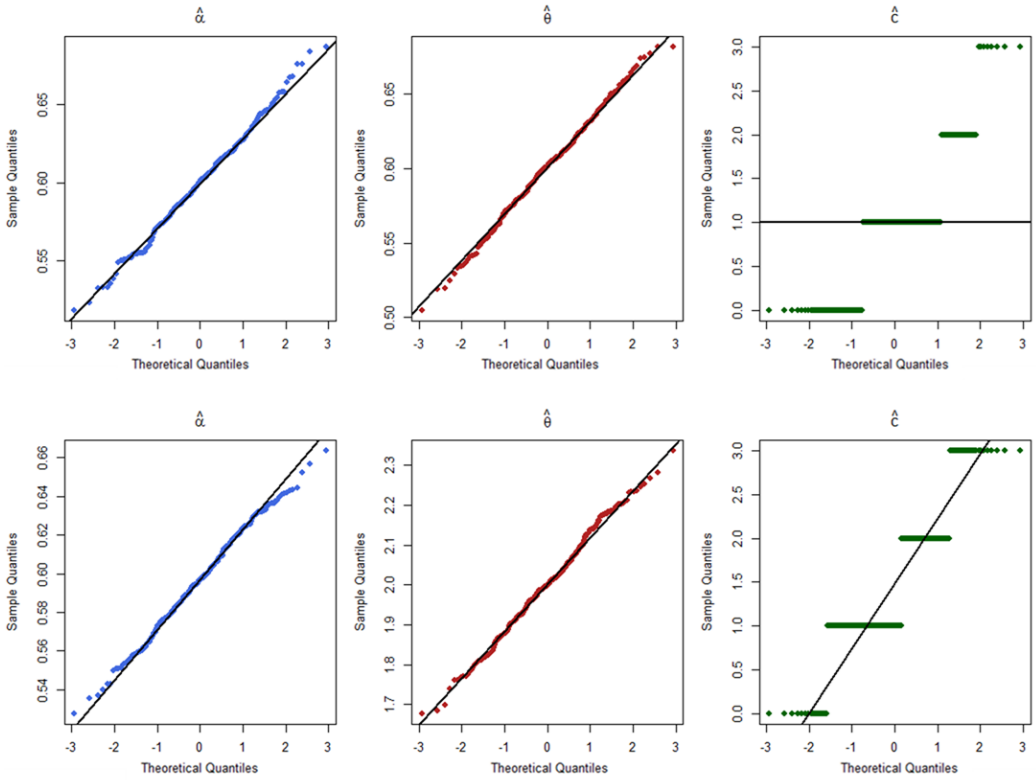


Figure 3. Q–Q plots of parameter estimates of the MIN–SB process with the sample size $T = 1500$ and two types of innovations. Panels above: Geometric innovations with the true parameters $\alpha = \theta = 0.6, c = 1$. Panels below: Poisson innovations with the true parameters $\alpha = 0.6, \theta = 2, c = 1$.

Table 6. Selection frequencies of the threshold estimator $\hat{c}$ for different sample sizes.

$\hat{c}$	Geometric Innovations				Poisson Innovations			
	$T = 50$	$T = 150$	$T = 500$	$T = 1500$	$T = 50$	$T = 150$	$T = 500$	$T = 1500$
0	63	73	85	68	35	34	36	44
1	128	158	159	190	131	120	130	141
2	87	50	47	34	82	113	111	88
3	22	19	9	8	52	33	23	27

6. Application in the Fire Safety Dynamics

This section discusses the practical application of the proposed MIN–SB framework, with a special emphasis on fire safety. To this end, the dataset used in the study is first described, as well as its basic statistical and dynamical properties. Then, using the theoretical results obtained above, the MIN–SB fitting of observed data is performed. Finally, in order to verify the effectiveness of the proposed model, its comparative analysis is performed with the standard INAR benchmark.

6.1. Dataset and Estimation Results

As already mentioned, the dataset used in this study refers to the number of people injured and killed in fires in the territory of the Republic of Serbia in the period from 2009 to 2023. The data were obtained from the Ministry of Internal Affairs of the Republic of Serbia, and based on them it was determined that a total of 6013 people were injured in fires during the observed time period, and 1370 of them died in $T = 5769$ independent events. Dynamics of these series, termed as Series A (daily number of injured) and Series B (daily number of fatalities) is shown in Figure 4, while their descriptive statistical analysis is presented in following Table 7. According to this, several characteristics of both observed series can be observed. First, the distributions are highly asymmetric and heavy–tailed, as indicated by large skewness and kurtosis values. In particular, the extremely high kurtosis of Series A suggests the presence of occasional extreme events, which is a common feature of accident-related count processes. Second, both series A and B show pronounced zero inflation, as their number of zero occurrences is 2445 and 4376, respectively. Thus, the empirical probability of zero is approximately 0.45 for Series A and nearly 0.80 for Series B, indicating that a large proportion of days involve no injuries or fatalities.

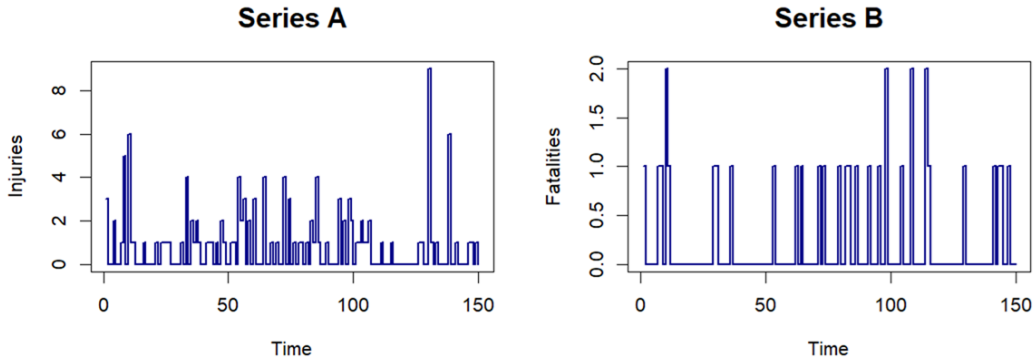


Figure 4. Daily dynamics of the number of injured (Series A) and fatalities (Series B) in fire accidents in the Republic of Serbia.

Moreover, the estimated zero–transition probabilities $\widehat{p}_{00}$ are slightly larger than the marginal zero probabilities $\widehat{p}_0$, which is consistent with the theoretical results in Remark 4 and indicates a pronounced clustering tendency of zero occurrences. Finally, the empirical relationship between the mean and variance indicates excessive over-dispersion in both series, particularly pronounced in the case of Series B. These empirical features (heavy tails,

over-dispersion, and strong zero inflation) motivate the use of flexible count time series models such as the proposed MIN–SB framework. In addition, according to the results of augmented Dickey–Fuller (ADF) test, also shown in Table 7, the presence of a unit root in both series is strongly rejected ($p < 0.01$), which confirms that the observed processes are stationary and suitable for modeling with count time series.

Table 7. Summary and zero statistics, along with testing the stationarity of observed real-world data.

Statistics	**Series A**	**Series B**
Minimum	0	0
Maximum	36	7
Mode	0	0
Median	1	0
Mean	1.098	0.250
Standard deviation	1.564	0.565
Variance	2.447	0.320
Skewness	4.406	3.085
Kurtosis	57.37	15.65
Zero count	2445	4376
Zero proportion $(\widehat{p}_0)$	0.4463	0.7988
Zero-to-zero transition $(\widehat{p}_{00})$	0.4744	0.8101
Expected zero length $(\widehat{L}_0)$	1.9027	5.2596
ADF stat.	-14.15	-13.45
(p-value)	(<0.01)	(<0.01)

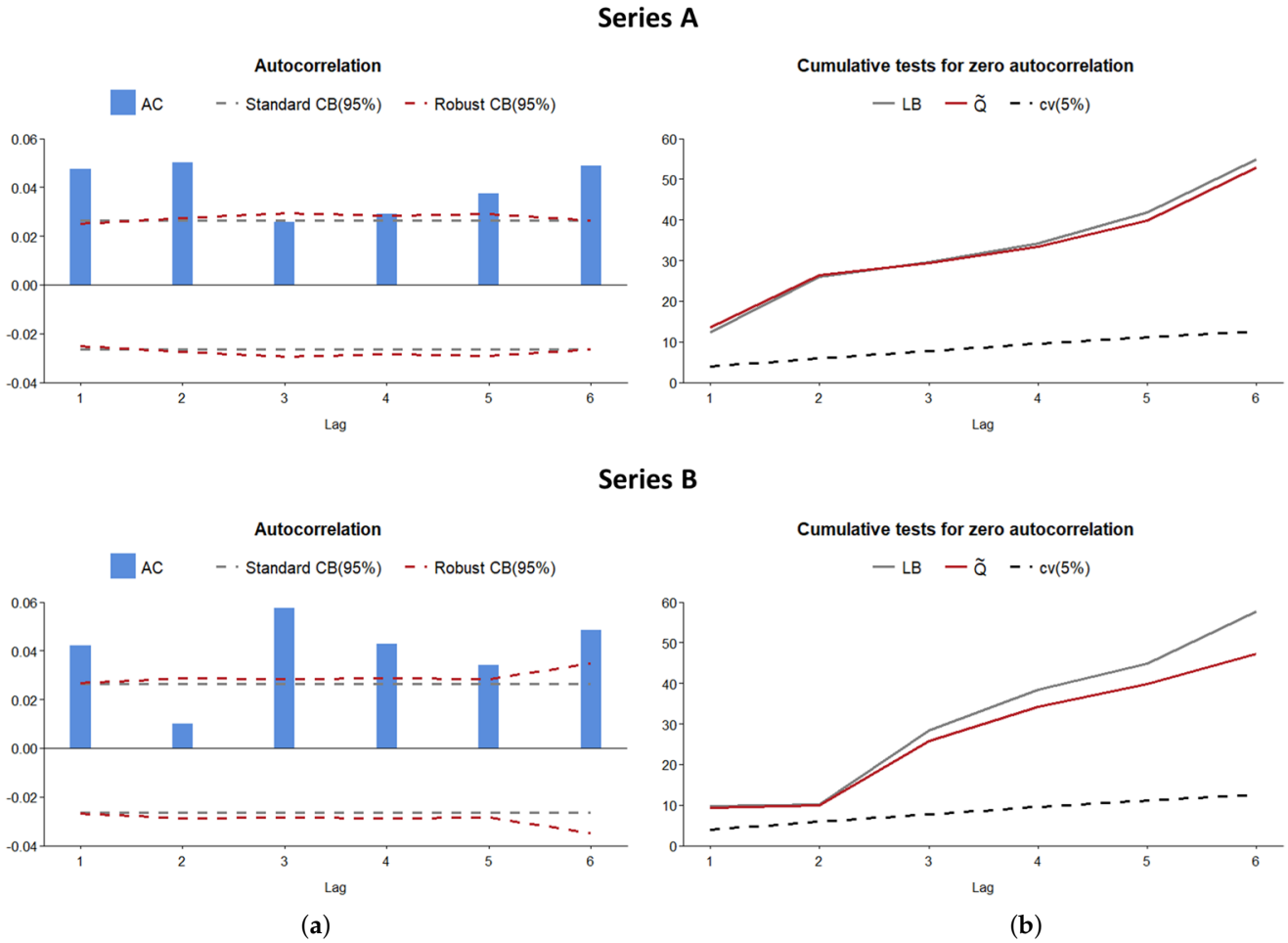
In addition, note that the presence of weak autocorrelation, shown in Figure 5a, suggests that both series can be modeled by an MIN–SB process, but also the possibility that they are uncorrelated. Thus, the following null hypotheses are tested here:

$$H_0 : \rho_x(k) = 0, \text{ at individual lag } \; k = 1, \ldots, m,$$
$$H_0 : \rho_x(k) = 0, \text{ for cumulative } \; k = 1, \ldots, m,$$

where $\rho_x(k)$ is the kth-order autocorrelations of the series (x_t). Using some recent results of the zero-autocorrelation testing, proposed by Dalla et al. [45] and Giraitis et al. [46], the first hypothesis is tested using the standard t and robust $\hat{t}$ statistics. On the other hand, to test cumulative absence of autocorrelation, the well-known Ljung-Box (LB) cumulative statistic is used:

$$LB_m(k) = n(n+2) \sum_{k=1}^{m} \frac{\widehat{\rho}_k^2}{n-k},$$

where $\widehat{\rho}_k$ is estimated k-th correlations of the series (x_t). Furthermore, the Q_m statistic given by Equation (17) in Dalla et al. [45] is also used, and whole testing procedure is carried-out using the function "ac.test()" within the R-package "testcorr" [47], as is presented in Figure 5. In the case of individual testing, shown in Figure 5a, it is noticeable that the hypothesis of the absence of correlation is rejected in most cases. Namely, for series A, at the significance level of $\alpha = 5\%$ (shown by the dashed lines), the null hypothesis is not rejected only for $k = 3$. Here, the corresponding p-values of t and $\hat{t}$ statistics are just slightly greater than the significance level (0.058 and 0.088, respectively). Similarly, within Series B, the only case of zero correlation occurs only at $k = 2$. However, according to the results of the cumulative tests, shown in Figure 5b, it is clear that the null hypothesis of non-correlation of the observed series is not valid at any level. Therefore, the assumption of their possible modeling using the MIN–SB process can be justified.



(**a**) (**b**)

Figure 5. Tests of zero autocorrelation of the observed time series: (**a**) testing of individual ACF leg; (**b**) cumulative ACF testing.

Further, to compare the MIN–SB framework with the INAR(1) benchmark modeling, both of these stochastic models are formed based on the observed data. For the INAR(1) model, parameters are estimated using the conditional least squares (CLS) approach, which provides consistent estimates of the binomial thinning parameter α and the innovation mean μ_ε. Since innovations are assumed to follow a member of the power-series (PS) family, the structural parameter θ can be obtained through the relationship between θ and the mean of the innovation distribution. Specifically, for geometric innovations, the parametric transformation $\theta = \mu_\varepsilon/(1+\mu_\varepsilon)$ is applied, while for Poisson innovations, the well-known equality $\theta = \mu_\varepsilon$ is valid.

In addition to RMSE, we also report information criteria to facilitate model comparison. Since the likelihood function of the MIN–SB model is not available in closed form, standard likelihood-based information criteria cannot be directly computed. Instead, we employ a pseudo-likelihood approach based on frequency matching. Specifically, the empirical and fitted frequencies are compared via residual sums of squares, which can be interpreted as arising from a Gaussian quasi-likelihood approximation (see, e.g., [48]). Therefore, AIC and BIC values for all models, including INAR(1), are computed as follows:

$$\mathrm{AIC} = n\log(\mathrm{RSS}/n) + 2p, \qquad \mathrm{BIC} = n\log(\mathrm{RSS}/n) + p\log(n),$$

where RSS denotes the residual sum of squares and p is the number of estimated parameters. Note that this ensures that the reported values are directly comparable across both competing models.

The following Table 8 reports estimation results and model comparison for Series A. The results clearly indicate that the MIN–SB specification provides a substantially better fit than the classical INAR(1) benchmark. In particular, the Poisson MIN–SB model achieves the lowest RMSE as well as the smallest AIC and BIC values, indicating superior explanatory performance. The estimated thinning parameter is relatively small ($\hat{\alpha} \approx 0.09$), suggesting that the persistence of the process is weak and that most variation is driven by innovation shocks rather than autoregressive propagation. The estimated threshold parameter $\hat{c} = 0$ indicates that the threshold mechanism is effectively inactive for this series, suggesting that zero-inflation arises primarily from the innovation distribution rather than from regime switching. In contrast, the INAR(1) model exhibits considerably larger prediction errors and higher information criteria values, indicating that the standard thinning-based structure is less capable of capturing the observed distributional properties of the data.

Table 8. Model comparison for Series A (number of injured in fires).

Model	Innovations	$\hat{\alpha}$	$\hat{\theta}$	$\hat{c}$	RMSE	AIC	BIC
MIN–SB	Geometric	0.1991	0.8576	0	0.4041	257.5	262.3
	Poisson	0.0871	9.2946	0	0.2941	233.9	238.8
INAR(1)	Geometric	0.0473	0.5111	–	0.5638	280.1	283.3
	Poisson	0.0473	1.0454	–	1.8914	369.7	372.9

Results for Series B, presented in Table 9, lead to similar conclusions. The MIN–SB model again provides the best overall fit, particularly under geometric innovations, which yield the lowest RMSE and information criteria values. Compared with Series A, the estimated threshold parameter $\hat{c} = 1$ suggests a slightly stronger gating mechanism, indicating that only sufficiently large innovations contribute to the latent intensity of the process. This is consistent with the empirical structure of the fatality series, where long sequences of zeros are frequently observed and non-zero outcomes tend to occur sporadically. Although the INAR(1) model with geometric innovations performs reasonably well, its goodness-of-fit measures remain slightly inferior to those of the MIN–SB model, confirming the advantage of incorporating the minification mechanism in modeling rare-event count processes. In particular, the estimated value of the threshold parameter suggests that the activation of the innovation mechanism is primarily governed by low-count regimes, indicating that clustering of small values plays a dominant role in the observed dynamics.

Table 9. Model comparison for Series B (number of fatalities in fires).

Model	Innovations	$\hat{\alpha}$	$\hat{\theta}$	$\hat{c}$	RMSE	AIC	BIC
MIN–SB	Geometric	0.1848	0.7684	1	0.0969	37.52	37.76
	Poisson	0.0523	4.9814	3	0.4285	61.31	61.55
INAR(1)	Geometric	0.0420	0.5111	–	0.1240	39.47	39.63
	Poisson	0.0420	0.2394	–	0.9532	72.10	72.26

Finally, Figure 6 illustrates the results obtained above, showing the fit of the empirical distributions of the observed series to the "best-fit" INAR and MIN–SB distributions (also marked in bold in previous tables). Overall, the MIN–SB model reproduces the empirical distributions with high accuracy, indicating that it is indeed a flexible framework for modeling fire-related casualty data. Thus, by combining threshold-driven innovation activation with a minification mechanism, the resulting model successfully captures the

simultaneous presence of zero inflation, overdispersion, and weak temporal dependence observed in the data, as well as other key structural features of the data.

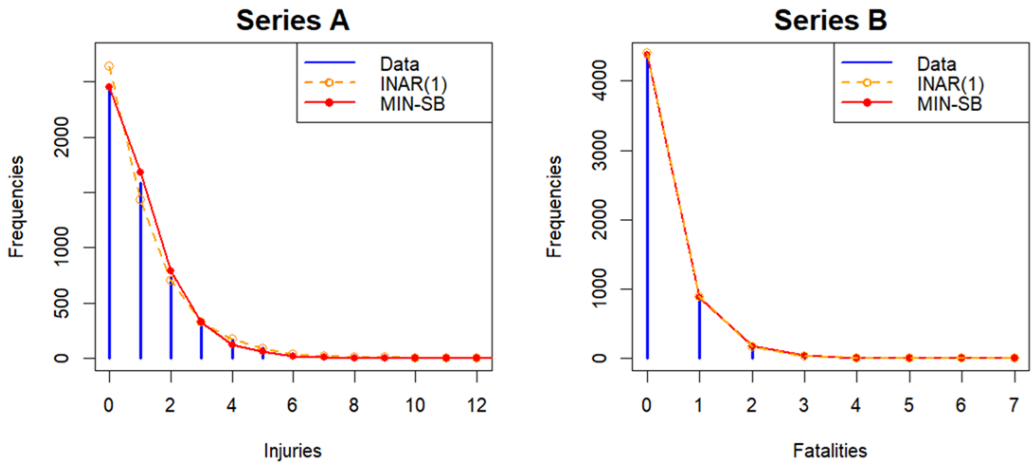


Figure 6. Frequency distributions of the observed data fitted by INAR(1) and MIN–SB process.

6.2. Discussion and Practical Implications

The empirical analysis provides several insights into the stochastic structure of the analyzed fire-related time series and highlights the practical relevance of the proposed MIN–SB framework. A key observation emerging from Table 7 is the pronounced role of zeros in both series. In particular, the fatality series exhibits extremely high zero inflation and long zero runs, while the injury series shows a moderate but still substantial proportion of zero observations. This empirical feature supports the modeling philosophy underlying the MIN–SB process, where the dynamics of the series are partly governed by a threshold-based mechanism controlling the occurrence and persistence of zeros. The estimation procedure used in this study directly reflects the theoretical structure of the model. Specifically, the simulated moment conditions are constructed from three quantities that characterize the zero dynamics of the process: the marginal zero probability p_0, the zero-to-zero transition probability p_{00}, and the expected length of zero runs $\mathbb{E}(L_0)$. These statistics admit a clear probabilistic interpretation within the MIN–SB framework and therefore provide a natural link between the theoretical results derived in Sections 2–4 and their empirical implementation.

The two analyzed series exhibit notably different stochastic regimes. For Series A (number of injured persons), the empirical zero statistics indicate moderate zero inflation and relatively short zero runs. The estimated MIN–SB parameters reveal a very small persistence parameter $\hat{\alpha}$, together with a threshold estimate $\hat{c} = 0$, suggesting that the dynamics are largely innovation-driven with weak latent dependence. In accordance with the definition of the MIN–SB model, stated in Section 3, the MIN–SB model is then close to the ordinary minification benchmark, which in this case shows better performance than the usual INAR-approach. Among the considered specifications, the Poisson innovation model provides the best goodness-of-fit according to RMSE and information criteria. From a practical perspective, this result indicates that daily injury counts are mainly influenced by external stochastic shocks, such as incident severity or environmental conditions, rather than by strong temporal propagation of previous events.

In contrast, the fatality series (Series B) exhibits extremely strong zero inflation and long sequences of zero observations. The estimated MIN–SB model yields a non-trivial threshold estimate $\hat{c} = 1$, indicating the presence of a gating mechanism that activates the latent process only when sufficiently large innovations occur. In this case, the geometric innovation specification provides the best empirical fit. This finding is intuitively plausible,

since the geometric distribution assigns relatively high probability to small counts and therefore naturally accommodates prolonged zero regimes. The results suggest that fatal events tend to appear in rare bursts separated by extended periods without observations.

A comparison with the classical INAR(1) benchmark further illustrates the advantage of the proposed approach. While INAR models rely solely on a thinning mechanism to generate temporal dependence, they do not explicitly account for the persistence of zero observations. Consequently, they must compromise between fitting the marginal distribution and reproducing the observed zero dynamics. In contrast, the MIN–SB framework incorporates a threshold-based mechanism that explicitly controls zero occurrences, allowing the model to capture both marginal and dynamic zero features simultaneously. This structural difference explains the consistently better performance of the MIN–SB models in terms of RMSE, AIC, and BIC, which indicate a clear preference for the MIN–SB specification over the competing model.

From an applied perspective, the results suggest that fire-related casualty data may exhibit heterogeneous stochastic mechanisms even within the same domain. Injury counts appear largely innovation-driven, whereas fatality counts are dominated by prolonged zero regimes and stronger internal persistence. Models capable of explicitly incorporating zero-transition dynamics therefore provide a more realistic statistical representation of such processes and may lead to improved predictive performance. These findings demonstrate that the MIN–SB framework offers a flexible and interpretable tool for modeling zero-inflated count time series. By combining threshold-driven innovation activation with the minification mechanism, the model successfully captures both innovation-driven variability and persistent zero clustering observed in real-world fire safety data.

7. Conclusions

In this paper, a new integer-valued time series model, termed the minification Split-BREAK (MIN–SB) process, has been introduced. The model combines a thinning-type latent dynamics with a threshold-based innovation activation mechanism and a minification operator, allowing it to capture zero inflation, zero clustering, and weak temporal dependence simultaneously. The main stochastic properties of the MIN–SB process were investigated, including stationarity conditions, moment structure, and the behavior of zero runs. In addition, the model was shown to admit a hidden-state interpretation, which provides a natural explanation for the clustering of zeros observed in many rare-event count processes. Parameter estimation was developed using a simulated method of moments framework based on zero-related statistics. MC experiments demonstrated favorable finite-sample performance of the proposed estimator and confirmed the asymptotic properties derived in the theoretical section.

The practical usefulness of the model was illustrated through an application to fire-related casualty data in Serbia. The empirical results showed that the MIN–SB specification provides a more accurate description of the data than the classical INAR(1) benchmark, particularly in the presence of strong zero inflation and clustering. Two important special cases of the model were examined in detail, namely the MIN–SB process with geometric and Poisson innovations. The empirical analysis suggests that the choice of innovation distribution reflects the dominant mechanism generating the observed counts, with Poisson innovations being more appropriate for innovation-driven variability and geometric innovations better capturing persistent zero regimes. These findings also suggest that incorporating threshold-based mechanisms can substantially improve both interpretability and predictive performance in applications involving highly zero-inflated data.

Several directions for further research naturally arise from the present study. Possible extensions include higher-order dynamics, multivariate settings, or alternative innovation

distributions, as well as on developing likelihood-based or Bayesian estimation procedures. Another promising direction is the application of the MIN–SB framework to other types of rare-event count processes arising in risk analysis, reliability, or safety studies. The results highlight the importance of modeling zero-inflation and threshold-driven dynamics in count time series. The proposed MIN–SB framework provides a flexible yet tractable approach that captures these features more effectively than standard models.

Author Contributions: Conceptualization, V.S.S., N.M. and H.S.B.; methodology, V.S.S., N.M. and K.T.; software, V.S.S., N.M. and K.T.; validation, N.M., K.T. and H.S.B.; formal analysis, V.S.S., H.S.B. and S.A.; data curation, V.S.S., N.M. and K.T.; writing—original draft preparation, V.S.S., N.M. and K.T.; writing—review and editing, H.S.B. and S.A.; visualization, V.S.S. and H.S.B.; supervision, V.S.S., N.M. and S.A.; project administration, K.T. and S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia (PNURSP2026R744).

Data Availability Statement: The data presented in this study are available on request from the corresponding author, subject to approval from the Ministry of Internal Affairs of the Republic of Serbia regarding their further distribution.

Acknowledgments: The authors The authors gratefully acknowledge Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R744), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia, for financial support of this project. We would also like to thank the Editor and anonymous reviewers for their careful reading and constructive and insightful comments, which greatly helped to improve the quality and clarity of the manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A.

Lemma A1 (Variance bound for Lipschitz functions)**.** *Let X be a square-integrable RV, and let $f : \mathbb{R} \to \mathbb{R}$ be an L-Lipschitz function, i.e.,*

$$|f(x) - f(y)| \leq L|x - y|, \qquad \forall x, y \in \mathbb{R}.$$

Then, the inequality $\mathrm{Var}(f(X)) \leq L^2 \,\mathrm{Var}(X)$ *holds.*

Proof. Let X' be an independent copy of the square-integrable RV X. Then, the following identity is valid:

$$\mathrm{Var}(X) = \frac{1}{2}\,\mathbb{E}\big[(X - X')^2\big].$$

From here, for any L-Lipschitz function $f(x)$, it follows:

$$\mathrm{Var}(f(X)) = \frac{1}{2}\,\mathbb{E}\big[(f(X) - f(X'))^2\big] \leq \frac{L^2}{2}\,\mathbb{E}\big[(X - X')^2\big].$$

According to $\mathbb{E}[(X - X')^2] = 2\,\mathrm{Var}(X)$, the stated inequality is immediately obtained. □

Lemma A2 (Conditional probability given a zero observation)**.** *Let (X_t, Y_t) denote the latent and observable components of the stationary MIN–SB process. Then:*

$$\mathbb{P}(X_t = 0 \mid Y_{t-1} = 0) = \frac{p_X(0)p_{00}^{(X)} + p_\varepsilon(0)\Big(\Psi_X\big((1+\alpha)^{-1}\big) - p_X(0)\Big)}{p_X(0) + (1 - p_X(0))p_\varepsilon(0)}, \tag{A1}$$

where $p_\varepsilon(k) = \mathbb{P}(\varepsilon_t = k)$, $p_X(0) = \mathbb{P}(X_t = 0)$, $p_{00}^{(X)} = \mathbb{P}(X_t = 0 \mid X_{t-1} = 0)$, *and* $\Psi_X(u)$ *is the PGF of RVs* (X_t).

Proof. By definition of the observed process, the equality $Y_{t-1} = \min(X_{t-1}, \varepsilon_{t-1})$ holds, and therefore:

$$\{Y_{t-1} = 0\} = \{X_{t-1} = 0\} \cup \{\varepsilon_{t-1} = 0,\ X_{t-1} > 0\},$$

where the events on the right-hand side are disjointed. Applying probability to the above equality yields:

$$\mathbb{P}(Y_{t-1} = 0) = \mathbb{P}(X_{t-1} = 0) + \mathbb{P}(X_{t-1} > 0, \varepsilon_{t-1} = 0),$$

and since ε_{t-1} is independent of X_{t-1}, one obtains:

$$\mathbb{P}(Y_{t-1} = 0) = p_X(0) + (1 - p_X(0))\, p_\varepsilon(0).$$

Furthermore, using the law of total probability with respect to X_{t-1}:

$$\mathbb{P}(X_t = 0, Y_{t-1} = 0) = \mathbb{P}(X_t = 0, X_{t-1} = 0) + \mathbb{P}(X_t = 0, X_{t-1} > 0, \varepsilon_{t-1} = 0).$$

Hence, for the first term we have:

$$\mathbb{P}(X_t = 0, X_{t-1} = 0) = \mathbb{P}(X_t = 0 \mid X_{t-1} = 0) P(X_{t-1} = 0) = p_X(0) p_{00}^{(X)},$$

and for the second term we condition on $X_{t-1} = k > 0$:

$$\mathbb{P}(X_t = 0, X_{t-1} = k, \varepsilon_{t-1} = 0) = p_X(k) p_\varepsilon(0) \mathbb{P}(X_t = 0 \mid X_{t-1} = k, \varepsilon_{t-1} = 0).$$

According to the definition of the NBT operator (cf. Equation (3)), we get:

$$\mathbb{P}(X_t = 0 \mid X_{t-1} = k, \varepsilon_{t-1} = 0) = \mathbb{P}(\alpha \circ X_{t-1} = 0 \mid X_{t-1} = k) = \left(\frac{1}{1+\alpha}\right)^k,$$

and summing over all $k \geq 1$ yields:

$$\begin{aligned}\mathbb{P}(X_t = 0, X_{t-1} > 0, \varepsilon_{t-1} = 0) &= p_\varepsilon(0) \sum_{k=1}^{\infty} p_X(k) \left(\frac{1}{1+\alpha}\right)^k \\ &= p_\varepsilon(0) \Big(\Psi_X\big((1+\alpha)^{-1}\big) - p_X(0) \Big).\end{aligned}$$

Finally, combining the above expressions in the conditional probability formula:

$$\mathbb{P}(X_t = 0 \mid Y_{t-1} = 0) = \frac{\mathbb{P}(X_t = 0,\ Y_{t-1} = 0)}{\mathbb{P}(Y_{t-1} = 0)}.$$

immediately gives Equation (A1) and completes the proof. □

Lemma A3 (PGF tail identity)**.** *Let* X *be a non-negative integer-valued random variable with PGF* $\Psi_X(u) = \mathbb{E}[u^X]$, $|u| \leq 1$. *Then, for any* $0 < \theta < 1$, *the following identity holds:*

$$\sum_{j=1}^{\infty} j\, \theta^j\, \mathbb{P}(X \geq j) = \frac{\theta}{(1-\theta)^2} \Big[1 - \Psi_X(\theta)\Big] - \frac{\theta^2}{1-\theta}\, \Psi'_X(\theta). \tag{A2}$$

Proof. Using the tail representation $\mathbb{P}(X \geq j) = \sum_{k=j}^{\infty} \mathbb{P}(X = k)$, one obtains:

$$\sum_{j=1}^{\infty} j\,\theta^j\,\mathbb{P}(X \geq j) = \sum_{k=0}^{\infty} \mathbb{P}(X = k) \sum_{j=1}^{k} j\,\theta^j. \tag{A3}$$

The inner sum is the standard finite geometric identity:

$$\sum_{j=1}^{k} j\,\theta^j = \frac{\theta\big(1 - (k+1)\theta^k + k\theta^{k+1}\big)}{(1-\theta)^2},$$

so substituting this expression into Equation (A3) yields:

$$\sum_{j=1}^{\infty} j\,\theta^j\,\mathbb{P}(X \geq j) = \frac{\theta}{(1-\theta)^2}\Big[1 - \mathbb{E}[\theta^X] - (1-\theta)\mathbb{E}[X\theta^X]\Big].$$

Finally, using the identity $\mathbb{E}[X\theta^X] = \theta\,\Psi'_X(\theta)$, Equation (A2) is easily obtained. □

References

1. Li, Q.; Chen, H.; Liu, X. A New Bivariate Random Coefficient INAR(1) Model with Applications. *Symmetry* **2022**, *14*, 39. [CrossRef]
2. Maya, R.; Chesneau, C.; Krishna, A.; Irshad, M.R. Poisson Extended Exponential Distribution with Associated INAR(1) Process and Applications. *Stats* **2022**, *5*, 755–772. [CrossRef]
3. Yu, K.; Tao, T. An Observation-Driven Random Parameter INAR(1) Model Based on the Poisson Thinning Operator. *Entropy* **2023**, *25*, 859. [CrossRef] [PubMed]
4. Huang, J.; Zhu, F.; Deng, D. A Mixed Generalized Poisson INAR Model with Applications. *J. Stat. Comput. Sim.* **2023**, **93**, 1851–1878. [CrossRef]
5. Almuhayfith, F.E.; Krishna, A.; Maya, R.; Irshad, M.R.; Bakouch, H.S.; Almulhim, M. Higher-Order INAR Model Based on a Flexible Innovation and Application to COVID-19 and Gold Particles Data. *Axioms* **2024**, *13*, 32. [CrossRef]
6. Pang, Y.; Wang, D.; Goh, M. A Two-Step Estimation Method for a Time-Varying INAR Model. *Axioms* **2024**, *13*, 19. [CrossRef]
7. Ghodake, A.; Awale, M.; Bakouch, H.S.; Alomair, G.; Daghestani, A.F. A Seasonal Transmuted Geometric INAR Process: Modeling and Applications in Count Time Series. *Mathematics* **2025**, *13*, 2334. [CrossRef]
8. Tavares, L.V. An Exponential Markovian Stationary Process. *J. Appl. Probab.* **1980**, *17*, 1117–1120. [CrossRef]
9. Tavares, L.V. A Non-Gaussian Markovian Model to Simulate Hydrologic Processes. *J. Hydrol.* **1980**, *46*, 281–287. [CrossRef]
10. Sim, C.H. Simulation of Weibull and Gamma Autoregressive Stationary Process. *Commun. Stat.–Simul. Comput.* **1986**, *15*, 1141–1146. [CrossRef]
11. Lewis, P.A.; McKenzie, E.D. Minification Processes and Their Transformations. *J. Appl. Probab.* **1991**, *28*, 45–57. [CrossRef]
12. Kalamkar, V.A. Minification Processes with Discrete Marginals. *J. Appl. Probab.* **1995**, *32*, 692–706. [CrossRef]
13. Aleksić, M.; Ristić, M. A Geometric Minification Integer-Valued Autoregressive Model. *Appl. Math. Model.* **2021**, *90*, 265–280. [CrossRef]
14. Stojanović, M. An EM Algorithm for Estimation of the Parameters of the Geometric Minification INAR Model. *J. Stat. Comput. Simul.* **2022**, *92*, 2941–2955. [CrossRef]
15. Qian, L.; Zhu, F. A New Minification Integer-Valued Autoregressive Process Driven by Explanatory Variables. *Aust. N. Z. J. Stat.* **2022**, *64*, 478–494. [CrossRef]
16. Stojanović, V.S.; Bakouch, H.S.; Bojičić, R.; Alomair, G.; Alghamdi, S.A. Poisson-Lindley Minification INAR Process with Application to Financial Data. *AIMS Math.* **2024**, *9*, 22627–22654. [CrossRef]
17. Jovanović, M.; Stojanović, V.; Kuk, K.; Popović, B.; Čisar, P. Asymptotic Properties and Application of GSB Process: A Case Study of the COVID-19 Dynamics in Serbia. *Mathematics* **2022**, *10*, 3849. [CrossRef]
18. Stojanović, V.S.; Bakouch, H.S.; Ljajko, E.; Božović, I. Laplacian Split-BREAK Process with Application in Dynamic Analysis of the World Oil and Gas Market. *Axioms* **2023**, *12*, 622. [CrossRef]
19. Ljajko E.; Stojanović, V.S.; Tošić, M.; Božović, I. Cauchy Split-Break Process: Asymptotic Properties and Application in Securities Market Analysis. *UPB Sci. Bull. Ser. A—Appl. Math. Phys.* **2023**, *85*, 139–154.
20. Stojanović, V.S.; Bakouch, H.S.; Gajtanović, Z.; Almuhayfith, F.E.; Kuk, K. Integer-Valued Split-BREAK Process with a General Family of Innovations and Application to Accident Count Data Modeling. *Axioms* **2024**, *13*, 40. [CrossRef]
21. Bourguignon, M.; Vasconcellos, K.L.P.; First Order Non-Negative Integer Valued Autoregressive Processes with Power Series Innovations. *Braz. J. Probab. Stat.* **2015**, 29, 71–93. [CrossRef]

22. Stojanović, V.; Ljajko, E.; Tošić, M. Parameters Estimation in Non-Negative Integer-Valued Time Series: Approach Based on Probability Generating Functions. *Axioms* **2023**, *12*, 112. [CrossRef]
23. Stojanović, V.S.; Bakouch, H.S.; Ljajko, E.; Qarmalah, N. Zero-and-One Integer-Valued AR(1) Time Series with Power Series Innovations and Probability Generating Function Estimation Approach. *Mathematics* **2023**, *11*, 1772. [CrossRef]
24. Ristić, M.S.; Bakouch, H.S.; Nastić, A.S. A New Geometric First-Order Integer-Valued Autoregressive (NGINAR(1)) Process. *J. Stat. Plan. Inference* **2009**, *139*, 2218–2226. [CrossRef]
25. Zhang, Q.; Wang, D.; Fan, X. A Negative Binomial Thinning-Based Bivariate INAR(1) Process. *Stat. Neerl.* **2020**, *74*, 517–537. [CrossRef]
26. Fan, Y.; Cheng, J.; Wang, D. A New Integer-Valued Threshold Autoregressive Process Based on Modified Negative Binomial Operator Driven by Explanatory Variables. *Stat. Pap.* **2024**, *65*, 5873–5901. [CrossRef]
27. Kang, Y.; Zhu, F.; Wang, D.; Wang, S. A Zero-Modified Geometric INAR(1) Model for Analyzing Count Time Series with Multiple Features. *Can. J. Stat.* **2024**, *52*, 873–899. [CrossRef]
28. Karlis, D.; Mamode Khan, N.; Sunecher, Y. The Negative Binomial INAR(1) Process under Different Thinning Processes: Can We Separate between the Different Models? *Stats* **2024**, *7*, 793–807. [CrossRef]
29. Pitman, J. *Probability*; Springer: New York, NY, USA, 1993; p. 372. [CrossRef]
30. Xu, L.; Zhang, S.; Zhang, R. A Note on Foster–Lyapunov Drift Condition for Recurrence of Markov Chains on General State Spaces. *J. Theor. Prob.* **2018**, *31*, 1923–1928. [CrossRef]
31. Meyn, S.; Tweedie, R.L. *Markov Chains and Stochastic Stability*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2012. [CrossRef]
32. Gallegos-Herrada, M.A.; Ledvinka, D.; Rosenthal, J.S. Equivalences of Geometric Ergodicity of Markov Chains. *J. Theor. Probab.* **2024**, *37*, 1230–1256. [CrossRef]
33. Douc, R.; Moulines, E.; Priouret, P.; Soulier, P. Ergodic Theory for Markov Chains. In *Markov Chains*; Springer Series in Operations Research and Financial Engineering; Springer: Cham, Switzerland, 2018. [CrossRef]
34. Maiti, R.; Biswas, A.; Das, S. Time Series of Zero-Inflated Counts and their Coherent Forecasting. *J. Forecast.* **2015**, *34*, 694–707. [CrossRef]
35. Li, C.; Cui, S.; Wang, D. Monitoring the Zero-Inflated Time Series Model of Counts with Random Coefficient. *Entropy* **2021**, *23*, 372. [CrossRef]
36. Duffie, D.; Singleton, K.J. Simulated Moments Estimation of Markov Models of Asset Prices. *Econometrica* **1993**, *61*, 929–952. [CrossRef]
37. Kristensen, D.; Shin, Y. Estimation of Dynamic Models with Nonparametric Simulated Moments. *J. Econom.* **2012**, *167*, 76–94. [CrossRef]
38. Jones, G.L. On the Markov Chain Central Limit Theorem. *Prob. Surv.* **2004**, *1*, 299–320. [CrossRef]
39. Newey, W.K.; McFadden, D. Large Sample Estimation and Hypothesis Testing. In *Handbook of Econometrics*; Elsevier: Amsterdam, The Netherlands, 1994; Volume IV.
40. Ferguson, T.S. *A Course in Large Sample Theory*, 1st ed.; Chapman & Hall: London, UK, 1996. Available online: https://stars.library.ucf.edu/etextbooks/390 (accessed on 11 March 2026).
41. Nelder, J.A.; Mead, R. A simplex algorithm for function minimization. *Comput. J.* **1965**, *7*, 308–313. [CrossRef]
42. Gross, L. Tests for Normality. R Package Version 1.0-2. 2013. Available online: http://CRAN.R-project.org/package=nortest (accessed on 23 February 2026).
43. Chan, K.S. Consistency and Limiting distribution of the Least Squares Estimator of a Threshold Autoregressive Model. *Ann. Stat.* **1993**, *21*, 520–533. [CrossRef]
44. Hansen, B.E. Threshold Autoregression in Economics. *Stat. Interface* **2011**, *4*, 123–127. [CrossRef]
45. Dalla, V.; Giraitis, L.; Phillips, P.C.B. Robust Tests for White Noise and Cross-Correlation. *Econom. Theory* **2022**, *38*, 913–941. [CrossRef]
46. Giraitis, L.; Li, Y.; Phillips, P.C.B. Robust Inference on Correlation under General Heterogeneity. *J. Econom.* **2024**, *244*, 105691. [CrossRef]
47. Dalla, V.; Giraitis, L.; Phillips, P.C.B. Testing Zero Correlation. R Package Version 0.3.0. Available online: https://cran.r-project.org/web/packages/testcorr/index.html (accessed on 1 March 2026).
48. White, H. Maximum Likelihood Estimation of Misspecified Models. *Econometrica* **1982**, 50, 1–25. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.