


Article

Two-Threshold-Variable Integer-Valued Autoregressive Model

Jiayue Zhang ¹, Fukang Zhu ^{1,*}  and Huaping Chen ²¹ School of Mathematics, Jilin University, Changchun 130012, China² School of Mathematics and Statistics, Henan University, Kaifeng 475004, China; chenhp0107@henu.edu.cn

* Correspondence: fzhu@jlu.edu.cn

Abstract: In the past, most threshold models considered a single threshold variable. However, for some practical applications, models with two threshold variables may be needed. In this paper, we propose a two-threshold-variable integer-valued autoregressive model based on the binomial thinning operator and discuss some of its basic properties, including the mean, variance, strict stationarity, and ergodicity. We consider the conditional least squares (CLS) estimation and discuss the asymptotic normality of the CLS estimator under the known and unknown threshold values. The performances of the CLS estimator are compared via simulation studies. In addition, two real data sets are considered to underline the superior performance of the proposed model.

Keywords: two-threshold-variable; time series of counts; integer-valued autoregression; conditional least squares estimate

MSC: 62M10; 62M20



Citation: Zhang, J.; Zhu, F.; Chen, H. Two-Threshold-Variable Integer-Valued Autoregressive Model. *Mathematics* **2023**, *11*, 3586. <https://doi.org/10.3390/math11163586>

Academic Editor: Hsien-Chung Wu

Received: 21 July 2023

Revised: 7 August 2023

Accepted: 17 August 2023

Published: 18 August 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Integer-valued time series data often occur in real applications, such as in the number of births at a hospital for several consecutive months; in the number of workers that are fired from a factory each month; in the number of claims and in information transmission times of insurance companies every month; and, particularly, in health studies as the daily number of infected patients or deaths due to a virus. The binomial thinning operator proposed by Steutel and van Harn [1] has been widely used to construct the autoregressive model, i.e., the integer-valued autoregressive (INAR) model (Al-Osh and Alzaid [2], McKenzie [3]), which is a popular method to analyze the integer-valued time series data and is defined as follows:

$$X_t = \alpha \circ X_{t-1} + \epsilon_t, \quad t = 1, 2, \dots,$$

where $\{\epsilon_t\}$ is a sequence of independently and identified distributed (i.i.d.) random variables and is independent of $X_s, \forall s > t$, “ $\alpha \circ$ ” is the binomial thinning operator with

$$\alpha \circ X := \sum_{i=1}^X B_i, \quad \forall X \in \mathbb{N},$$

and B_i is independent of X and is a sequence of i.i.d. Bernoulli random variables with $P(B_i = 1) = \alpha \in (0, 1) = 1 - P(B_i = 0)$. See Du and Li [4], Silva and Oliveira [5], Silva and Silva [6], and Zhang et al. [7] for more extensions of the INAR model, among others.

However, the above INAR model and its extensions aim to analyze the integer-valued time series with a linear structure and are unavailable to analyze the integer-valued time series with a nonlinear structure. Scotto et al. [8] proposed a discrete counterpart of the conventional max-autoregressive process of order one. It is based on the binomial thinning operator and driven by a sequence of i.i.d. non-negative integer-valued random variables with either a regularly varying right tail or an exponential-type right tail. Aleksić and Ristić [9] introduced a new minification INAR model of the first-order to solve the problem, which can arise when the binomial thinning operator or the negative binomial thinning

operator is used. Namely, if one of these thinning operators is used in the construction of the minification model then it is possible that the model becomes zero constantly over time.

The threshold autoregressive (TAR) model proposed by Tong [10] provides an efficient method with which to handle continuous-valued time series data with a nonlinear structure. Boero and Marrocu [11] and Potter [12] applied threshold models to economy and finance. Dueker et al. [13] proposed a contemporaneous TAR model for the bond market. See Tong [14] for more discussion on continuous-valued threshold models. Li and Tong [15] proposed a faster approach (called the nested sub-sample search algorithm) to fit a threshold model using the least squares method. In an analogy to the TAR model, Monteiro et al. [16] introduced an integer-valued self-exciting threshold autoregressive process (SETINAR(2,1)), which is driven by an independent Poisson-distributed random variable. Wang et al. [17] proposed a self-excited threshold Poisson autoregressive model, which assumes a two-regime structure of the conditional mean process according to the magnitude of the lagged observations. Yang et al. [18] proposed an integer-valued threshold autoregressive process driven by an independent negative-binomial distributed random variable and the negative binomial thinning operator. To explore the relationship between stock return autocorrelation and trading volume, Zhang et al. [19] proposed a multiple-threshold-variable autoregressive model and applied it to analyze quarterly U.S. real GNP data. But the multiple-threshold-variable autoregressive model is restricted to a continuous-valued time series, and few studies have discussed a similar model for the nonnegative integer-valued time series. To fill this gap, we propose a new two-threshold-variable INAR (2-TINAR) model, which provides an alternative way to analyze nonnegative integer-valued time series with a nonlinear structure.

The paper is organized as follows. Section 2 defines the 2-TINAR model and establishes its stability properties. Section 3 considers conditional least squares (CLS) estimation. Section 4 gives a simulation study. Section 5 considers two real data applications to illustrate the effectiveness of the proposed model. Section 6 concludes.

2. Two-Threshold-Variable Integer-Valued Autoregressive Model

In this paper, we first give the definition of the 2-TINAR model and then discuss some properties of the model.

Definition 1. The 2-TINAR process $\{X_t\}$ is defined as

$$X_t = \sum_{j=1}^4 (\alpha_{j1} \circ X_{t-1} + \alpha_{j2} \circ X_{t-2} + \epsilon_{jt}) I_{jt}(r, s), \quad t = 1, 2, \dots, \quad (1)$$

where

(1) (r, s) are the threshold parameters and

$$\begin{aligned} I_{1t}(r, s) &= I(X_{t-1} > r, X_{t-2} > s) = I(X_{t-1} > r)I(X_{t-2} > s), \\ I_{2t}(r, s) &= I(X_{t-1} \leq r, X_{t-2} > s) = I(X_{t-1} \leq r)I(X_{t-2} > s), \\ I_{3t}(r, s) &= I(X_{t-1} \leq r, X_{t-2} \leq s) = I(X_{t-1} \leq r)I(X_{t-2} \leq s), \\ I_{4t}(r, s) &= I(X_{t-1} > r, X_{t-2} \leq s) = I(X_{t-1} > r)I(X_{t-2} \leq s); \end{aligned}$$

- (2) $\alpha_{ji} \in (0, 1), 0 < \sum_{i=1}^2 \alpha_{ji} < 1, i = 1, 2, j = 1, 2, 3, 4$, “ \circ ” is the binomial thinning operator, and the operators in $\alpha_{j1} \circ X_{t-1}$ and $\alpha_{j2} \circ X_{t-2}$ operate independently;
- (3) $\forall t, \epsilon_{jt} \sim \text{Pois}(\lambda_j)$ and for fixed j , ϵ_{jt} is i.i.d. and independent of $\alpha_{ji} \circ X_{t-i}$ and X_{t-i} .

In the following, we consider the properties of the 2-TINAR model, including stationarity, ergodicity, mean and variance, which will be given in the next three propositions, whose proofs are delegated to Appendix A.

Proposition 1. Let $Y_t = (X_t, X_{t-1})^\top$; then,

- (1) $\{Y_t\}$ is an irreducible, aperiodic, and positive recurrent Markov chain;
- (2) $\{Y_t\}$ is an ergodic sequence, and a strictly stationary process satisfying (1) exists.

Proposition 2. Assume that $\{X_t\}$ is generated from (1); then, $E(X_t^k) < \infty, k = 1, 2, 3$.

Proposition 3. Assume that $\{X_t\}$ is generated from (1) and $\mathcal{F} = \sigma\{X_{t-i}, i \geq 1\}$. Then, the mean and variance are

- (1) $E(X_t | \mathcal{F}_{t-1}) = \sum_{j=1}^4 (\alpha_{j1} X_{t-1} + \alpha_{j2} X_{t-2} + \lambda_j) I_{jt}(r, s);$
- (2) $E(X_t) = \sum_{j=1}^4 p_j (\alpha_{j1} u_j + \alpha_{j2} u_j^* + \lambda_j);$
- (3) $\text{Var}(X_t | \mathcal{F}_{t-1}) = \sum_{j=1}^4 (\alpha_{j1} (1 - \alpha_{j1}) X_{t-1} + \alpha_{j2} (1 - \alpha_{j2}) X_{t-2} + \lambda_j) I_{jt}(r, s);$
- (4) $\text{Var}(X_t) = \sum_{j=1}^4 (\alpha_{j1}^2 (p_j (v_j + u_j^2) - p_j^2 u_j^2) + \alpha_{j2}^2 (p_j (v_j^* + (u_j^*)^2) - p_j^2 (u_j^*)^2) + 2(\alpha_{j1} \alpha_{j2} w_j p_j - \alpha_{j1} \alpha_{j2} p_j^2 u_j u_j^*) + 2(\alpha_{j1} \lambda_j p_j u_j - \alpha_{j1} \lambda_j p_j^2 u_j) + 2(\alpha_{j2} \lambda_j p_j u_j^* - \alpha_{j2} \lambda_j p_j^2 u_j^*) + \alpha_{j1} (1 - \alpha_{j1}) p_j u_j + \alpha_{j2} (1 - \alpha_{j2}) p_j u_j^* + \lambda_j p_j) + \sum_{m=1}^6 C_m, m = 1, 2, \dots, 6,$
where $p_j, u_j, v_j, u_j^*, v_j^*, w_j$ and C_m are given in Appendix A.

3. Conditional Least Squares Estimation

In this section, we use the CLS method to estimate the parameters involved in the 2-TINAR model. Here, we consider the following two cases: the first one is that the threshold values r and s are known and the second one is that the threshold values r and s are unknown.

3.1. Known Case of (r, s)

In this part, we assume that $\{X_t, t = 1, \dots, n\}$ comprises observations generated by (1), $\psi_j = (\alpha_{j1}, \alpha_{j2}, \lambda_j)^\top$ is the vector of regression parameters, $j = 1, 2, 3, 4$, $\phi = (\psi_1^\top, \psi_2^\top, \psi_3^\top, \psi_4^\top)^\top = (\phi_1, \dots, \phi_{12})^\top$, and $g(\phi, X_{t-1}, X_{t-2}) = E_\phi(X_t | \mathcal{F}_{t-1}) = \sum_{j=1}^4 (\alpha_{j1} X_{t-1} + \alpha_{j2} X_{t-2} + \lambda_j) I_{jt}(r, s)$. Then, the CLS estimate $\hat{\psi}_{j,CLS} = (\hat{\alpha}_{j1,CLS}, \hat{\alpha}_{j2,CLS}, \hat{\lambda}_{j,CLS})^\top$ is obtained by minimizing the function

$$Q(\phi) = \sum_{t=3}^n (X_t - g(\phi, X_{t-1}, X_{t-2}))^2 = \sum_{t=3}^n q_t^2(\phi), \quad (2)$$

where $q_t(\phi) = X_t - \sum_{j=1}^4 (\alpha_{j1} X_{t-1} + \alpha_{j2} X_{t-2} + \lambda_j) I_{jt}(r, s)$. Then, the closed form of $\psi_{j,CLS}$ is obtained by the following equations:

$$\begin{cases} \frac{\partial Q(\phi)}{\partial \alpha_{j1}} = -2 \sum_{t=3}^n \left(X_t - \sum_{j=1}^4 (\alpha_{j1} X_{t-1} + \alpha_{j2} X_{t-2} + \lambda_j) I_{jt}(r, s) \right) X_{t-1} = 0, \\ \frac{\partial Q(\phi)}{\partial \alpha_{j2}} = -2 \sum_{t=3}^n \left(X_t - \sum_{j=1}^4 (\alpha_{j1} X_{t-1} + \alpha_{j2} X_{t-2} + \lambda_j) I_{jt}(r, s) \right) X_{t-2} = 0, \\ \frac{\partial Q(\phi)}{\partial \lambda_j} = -2 \sum_{t=3}^n \left(X_t - \sum_{j=1}^4 (\alpha_{j1} X_{t-1} + \alpha_{j2} X_{t-2} + \lambda_j) I_{jt}(r, s) \right) I_{jt}(r, s) = 0. \end{cases}$$

Denote

$$A_j = \begin{pmatrix} \sum_{t=3}^n I_{jt}(r, s) X_{t-1}^2 & \sum_{t=3}^n I_{jt}(r, s) X_{t-1} X_{t-2} & \sum_{t=3}^n I_{jt}(r, s) X_{t-1} \\ \sum_{t=3}^n I_{jt}(r, s) X_{t-1} X_{t-2} & \sum_{t=3}^n I_{jt}(r, s) X_{t-2}^2 & \sum_{t=3}^n I_{jt}(r, s) X_{t-2} \\ \sum_{t=3}^n I_{jt}(r, s) X_{t-1} & \sum_{t=3}^n I_{jt}(r, s) X_{t-2} & \sum_{t=3}^n I_{jt}(r, s) \end{pmatrix},$$

$$D_j = \left(\sum_{t=3}^n I_{jt}(r, s) X_t X_{t-1}, \sum_{t=3}^n I_{jt}(r, s) X_t X_{t-2}, \sum_{t=3}^n I_{jt}(r, s) X_t \right)^\top.$$

Then, $\psi_{j,CLS} = A_j^{-1} D_j$.

To study the asymptotic behaviour of the estimators, we make the following assumptions about the underlying process and the parameter space.

Assumption 1. If $\{X_t\}$ is generated from (1), then the parameter space ϕ is a compact subset of $\mathcal{D} \times \mathbb{R}_+^4$, and $\mathcal{D} = (0, 1) \times (0, 1) \times (0, 1) \times (0, 1) \times (0, 1) \times (0, 1) \times (0, 1) \times (0, 1)$ is a compact subset of \mathbb{R}_+^8 .

Assumption 2. The model (1) is identifiable, i.e., $p_\phi \neq p_{\phi_0}$, if $\phi \neq \phi_0$, where p_ϕ denotes the marginal distribution of $\{X_t\}$ with parameter ϕ .

For Assumption 1, assume the parameter space is compact so that the asymptotic properties of the CLS estimator can be guaranteed, which is common in INAR models. Parameter identifiability in Assumption 2 is a property that concerns whether the model parameters can be uniquely determined, which is the foundation for parameter estimation.

The following theorem establishes the asymptotic properties of the CLS estimator, whose proof will be given in Appendix A.

Theorem 1. Under the Assumptions 1 and 2, $\hat{\phi}_{CLS}$ is strongly consistent and asymptotically normally distributed with

$$\sqrt{T}(\hat{\phi}_{CLS} - \phi_0) \xrightarrow{d} N(0, V^{-1}WV^{-1}),$$

where $V = E\left(\frac{\partial E(X_t|\mathcal{F}_{t-1})}{\partial \phi} \frac{\partial E(X_t|\mathcal{F}_{t-1})}{\partial \phi^\top}\right)$ and $W = E\left((q_t(\phi))^2 \frac{\partial E(X_t|\mathcal{F}_{t-1})}{\partial \phi} \frac{\partial E(X_t|\mathcal{F}_{t-1})}{\partial \phi^\top}\right)$.

3.2. Unknown Case of (r, s)

Under the unknown case of (r, s) , we first estimate (r, s) , which is obtained by minimizing (2) by the following steps:

- (1) For each candidate for (r, s) in CR, we estimate $\hat{\phi}$ by minimizing $Q(\phi, r, s)$, i.e.,

$$\hat{\phi} = \arg \min_{\phi} Q(\phi, r, s).$$

- (2) The estimator for thresholds (r, s) is obtained by searching over all of the candidates for (r, s) in CR, i.e.,

$$(\hat{r}, \hat{s}) = \arg \min_{(r,s) \in CR} Q(\hat{\phi}, r, s),$$

where CR is the set of candidates for estimators for (r, s) with $CR = \{X_{\{i\}}, X_{\{i+1\}}\}, i = (0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8)$. To select a proper CR, we propose a validated method, which can put less pressure on the capacity of computing and guarantee the accuracy of estimates; $\{X_{\{i\}}, X_{\{i+1\}}\}$ are thirteen candidates for (\hat{r}, \hat{s}) . Actually, the estimators for (r, s) are searched for from $X_{\{0.20\}}$ to $X_{\{0.85\}}$, and this is a sufficient search range for the threshold. Hence, this method guarantees reasonable and sufficient search ranges without too much pressure on computing.

Based on the initial setting of the parameter space, both of the thresholds r and s are integers. Therefore, the consistency of \hat{r} means that $\hat{r} = r$ utterly, and so does \hat{s} . It is feasible that we estimate the other parameters by assuming that thresholds (r, s) are known, which is similar to the discussion in Wang et al. [17] because the validity of the estimates with unknown thresholds, as for the other parameters, is asymptotically identical to that obtained with known (r, s) . Hence, in this subsection, we treat the thresholds r and s as known parameters and consider the consistency for $\phi = (\alpha_{j1}, \alpha_{j2}, \lambda_j)^\top, j = 1, 2, 3, 4$.

4. Simulation Study

In this section, we illustrate the finite sample property of the CLS estimates under the known and unknown cases of (r, s) . In the simulation, we use $m = 10,000$ replications and set the sample size is $T = 500, 1000, 2000$ and $10,000$.

4.1. Known Case of (r, s)

In the simulation study, we consider the following parameter combinations:

(C1) $(0.3, 0.2, 7, 0.2, 0.25, 6, 0.2, 0.3, 8, 0.3, 0.2, 6)$ with $(r, s) = (13, 11)$,

(C2) $(0.3, 0.35, 15, 0.3, 0.35, 20, 0.3, 0.4, 25, 0.35, 0.25, 15)$ with $(r, s) = (55, 53)$.

As discussed in Li and Tong [15], if the proportion of observations in one regime to the whole is less than 5%, the estimation results may not be reliable. To illustrate the reasonableness of (r, s) given in (C1) and (C2), we give two sample paths of the sample generated by the 2-TINAR model with (C1) and (C2) in Figure 1, which shows that the proportion of the observations in each range is no less than 20%. In Figure 1, circle means sample point, red dot line means the value of r , and blue dot line means the value of s .

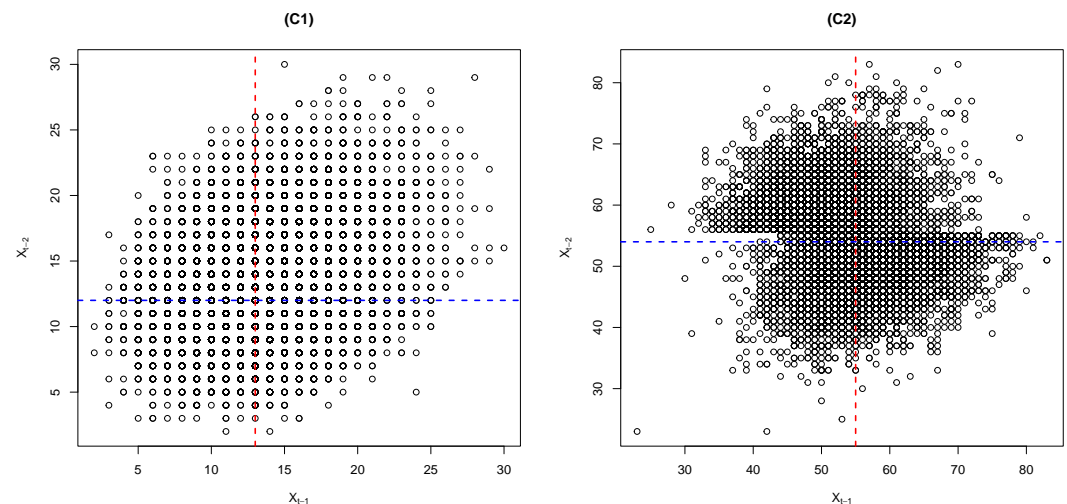


Figure 1. Sample paths for combinations (C1) and (C2).

The mean and standard deviation (SD) of the estimates are summarized in Table 1, from which we obtain that the CLS method performs reasonably well when (r, s) is known because the mean gradually approaches the true value of the parameter and SD decreases gradually, when the sample size is increasing.

To further account for the reasonableness of the CLS estimates, we present the boxplots of the parameter combinations (C1) in Figure 2 (the boxplots of (C2) are similar, and we omit them), and the QQ-plots of the parameter combinations (C1) and (C2) indicate the asymptotic normality of the CLS estimator. For saving space, we omit the QQ-plots, which are available upon request. All of them highlight that the good performances of the CLS estimate under the case of (r, s) are known.

Table 1. The mean and standard deviation (in brackets) of the estimates of 2-TINAR model with known (r, s) .

T	α_{11}	α_{12}	λ_1	α_{21}	α_{22}	λ_2	α_{31}	α_{32}	λ_3	α_{41}	α_{42}	λ_4
(C1) = (0.3, 0.2, 7, 0.2, 0.25, 6, 0.2, 0.3, 8, 0.3, 0.2, 6)												
500	0.2913 (0.1106)	0.1884 (0.0908)	7.3424 (2.1188)	0.2087 (0.1225)	0.2495 (0.1116)	6.0019 (2.0209)	0.2090 (0.1285)	0.3034 (0.1692)	8.1320 (2.1968)	0.3301 (0.2155)	0.2921 (0.2147)	6.5232 (4.1374)
1000	0.2951 (0.0781)	0.1936 (0.0658)	7.1725 (1.4782)	0.2022 (0.0918)	0.2493 (0.0802)	5.9952 (1.4289)	0.2000 (0.0990)	0.2974 (0.1280)	8.0702 (1.5499)	0.3060 (0.1634)	0.2373 (0.1635)	6.1198 (3.1374)
2000	0.2978 (0.0548)	0.1965 (0.0467)	7.0862 (1.0395)	0.2004 (0.0657)	0.2497 (0.0565)	5.9980 (0.9988)	0.1991 (0.0728)	0.2981 (0.0921)	8.0362 (1.0918)	0.3014 (0.1208)	0.2110 (0.1281)	6.0076 (2.3074)
10,000	0.2995 (0.0244)	0.1994 (0.0209)	7.0170 (0.4650)	0.2001 (0.0294)	0.2499 (0.0252)	5.9998 (0.4464)	0.1999 (0.0328)	0.2997 (0.0413)	8.0048 (0.4897)	0.3001 (0.0537)	0.1998 (0.0645)	6.0013 (1.0270)
(C2) = (0.3, 0.35, 15, 0.3, 0.35, 20, 0.3, 0.4, 25, 0.35, 0.25, 15)												
500	0.2955 (0.1487)	0.3395 (0.1278)	17.3932 (11.5915)	0.3014 (0.1017)	0.3460 (0.1130)	20.2152 (8.1557)	0.2978 (0.1371)	0.3985 (0.1181)	25.3153 (6.2198)	0.3506 (0.1245)	0.2493 (0.1334)	15.7572 (8.8644)
1000	0.2948 (0.1088)	0.3448 (0.0898)	15.8756 (8.7210)	0.3007 (0.0713)	0.3482 (0.0793)	20.0721 (5.7577)	0.2974 (0.1053)	0.3996 (0.0891)	25.1709 (5.2321)	0.3503 (0.0870)	0.2479 (0.0994)	15.1737 (6.5817)
2000	0.2969 (0.0775)	0.3475 (0.0629)	15.3461 (6.4147)	0.3001 (0.0502)	0.3489 (0.0558)	20.0642 (4.0647)	0.2986 (0.0774)	0.4002 (0.0661)	25.0691 (4.1621)	0.3502 (0.0611)	0.2483 (0.0708)	15.0749 (4.7424)
10,000	0.2995 (0.0343)	0.3495 (0.0279)	15.0594 (2.8665)	0.3001 (0.0223)	0.3499 (0.0250)	20.0017 (1.8115)	0.2999 (0.0363)	0.4001 (0.0311)	25.0010 (2.1124)	0.3500 (0.0272)	0.2497 (0.0317)	15.0124 (2.1158)

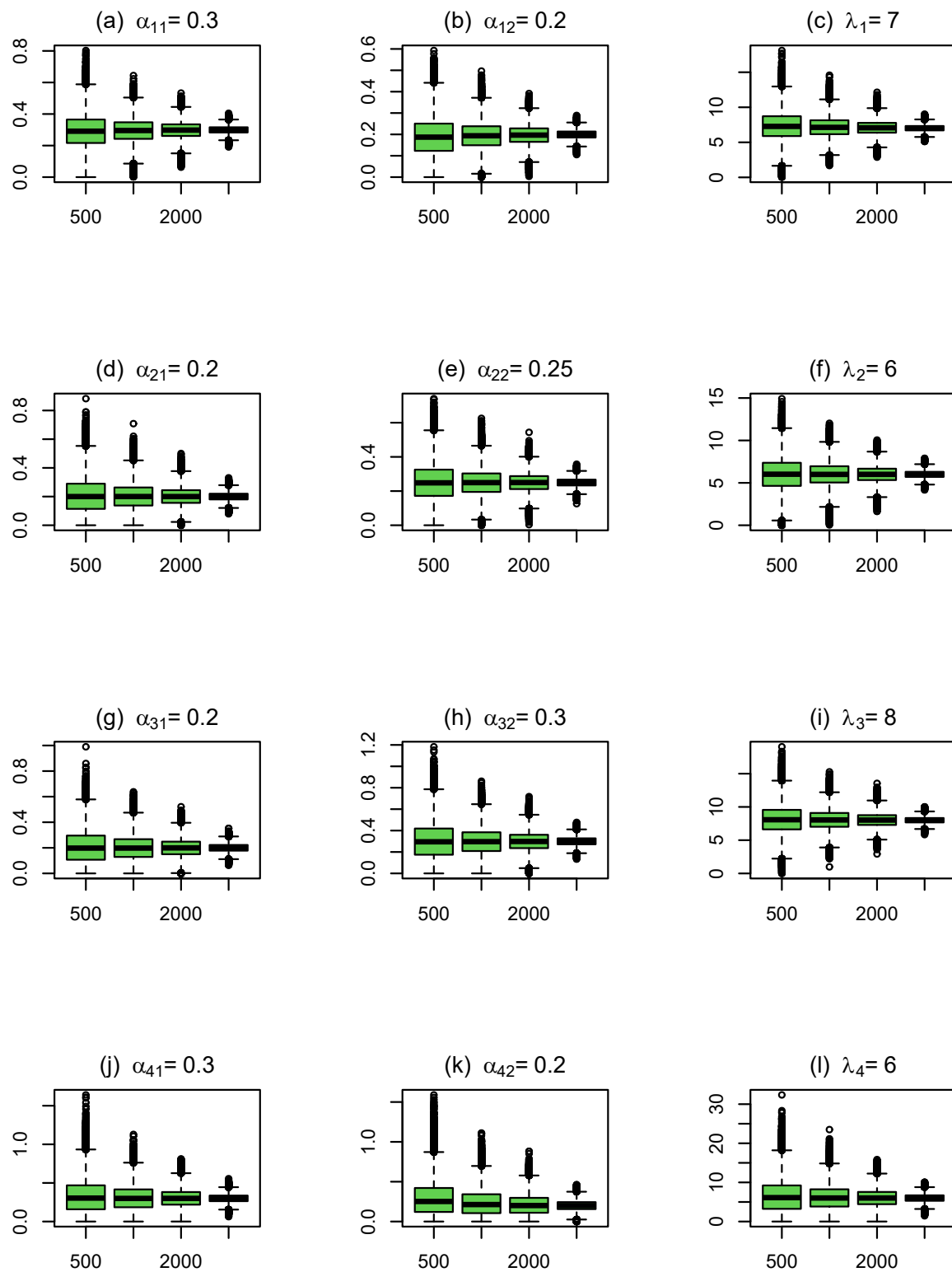


Figure 2. Boxplots of (C1).

4.2. Unknown Case of (r, s)

In this part, we first let (C3) $(0.2, 0.25, 3, 0.25, 0.35, 5, 0.3, 0.3, 4, 0.3, 0.25, 6, 13, 14)$ and (C4) $(0.30, 0.25, 6, 0.25, 0.35, 6, 0.4, 0.3, 9, 0.3, 0.35, 8, 30, 31)$. Then, we use the approach given in Section 3.2 to obtain (\hat{r}, \hat{s}) .

To illustrate the reasonableness of (\hat{r}, \hat{s}) given in (C3) and (C4), we give two sample paths of the sample generated by 2-TINAR model with (C3) and (C4) in Figure 3, which shows that the proportion of the observations in each range is no less than 20%. In Figure 3, circle, red and blue dot line in figure are the same as Figure 1.

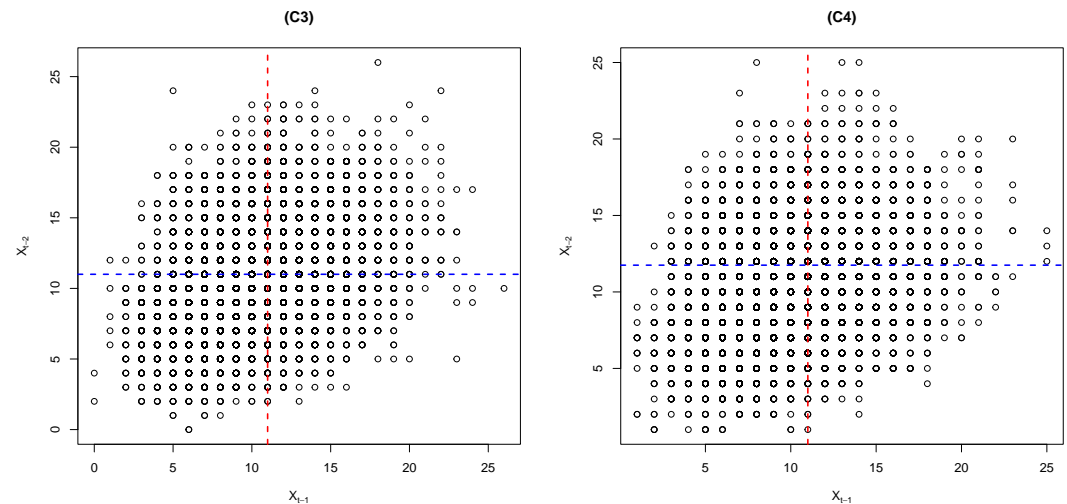


Figure 3. Sample paths of (C3) and (C4).

The mean and SD of the estimates are summarized in Table 2, from which we can see that with the increase in sample size, the mean gradually approaches the true value of the parameter and SD decreases gradually. The boxplots (C3) and (C4) are similar to (C1) and (C2). We present the boxplots of the parameter combinations (C3) in Figures 4 (the boxplots of (C4) are similar and we omit them).

From Figures 4, the median of the estimator is closer to the true value and the quartile range and overall range of the estimated values become narrower, both of which indicate the consistency of the estimators. The QQ-plots of the parameter combinations (C3) and (C4) indicate the asymptotic normality of the CLS estimator. For the same reason, we omit the QQ-plots. All of them highlight that the good performances of the CLS estimate under the case of (r, s) are unknown.

Table 2. The mean and standard deviation (in brackets) of the estimates of 2-TINAR model with unknown (r, s) .

T	α_{11}	α_{12}	λ_1	α_{21}	α_{22}	λ_2	α_{31}	α_{32}	λ_3	α_{41}	α_{42}	λ_4	r	s
(C3) = (0.2, 0.25, 3, 0.25, 0.35, 5, 0.3, 0.3, 4, 0.3, 0.25, 6, 13, 14)														
500	0.3165 (0.2458)	0.3882 (0.3007)	9.6383 (7.8787)	0.2897 (0.2093)	0.4308 (0.2878)	6.7583 (5.0400)	0.3092 (0.0655)	0.3000 (0.0590)	3.9558 (0.6470)	0.3908 (0.2612)	0.2547 (0.1715)	6.6819 (4.5692)	9.9851 (2.0175)	10.5800 (2.1507)
1000	0.2482 (0.1764)	0.2966 (0.2082)	6.6049 (5.6639)	0.2544 (0.1584)	0.3860 (0.2204)	5.6138 (3.8261)	0.3041 (0.0471)	0.3012 (0.0419)	3.9727 (0.4652)	0.3370 (0.1945)	0.2404 (0.1337)	6.2223 (3.4623)	11.3103 (1.9055)	12.0423 (2.0830)
2000	0.2156 (0.1327)	0.2634 (0.1540)	4.5700 (3.9383)	0.2486 (0.1235)	0.3584 (0.1626)	5.1460 (2.8763)	0.3014 (0.0325)	0.3007 (0.0291)	3.9891 (0.3225)	0.3127 (0.1352)	0.2427 (0.0971)	6.0176 (2.3564)	12.3400 (1.3604)	13.2323 (1.5242)
10,000	0.2003 (0.0627)	0.2486 (0.0725)	3.0470 (1.4592)	0.2503 (0.0569)	0.3491 (0.0717)	5.0088 (1.3053)	0.2996 (0.0133)	0.2998 (0.0127)	4.0057 (0.1368)	0.3004 (0.0537)	0.2501 (0.0411)	5.9923 (0.8949)	12.9962 (0.1104)	13.9960 (0.1166)
(C4) = (0.30, 0.25, 6, 0.25, 0.35, 6, 0.4, 0.3, 9, 0.3, 0.35, 8, 30, 31)														
500	0.5152 (0.4299)	0.4893 (0.4219)	22.8606 (19.7688)	0.2621 (0.1613)	0.3919 (0.2612)	12.3190 (10.0337)	0.3923 (0.0666)	0.2849 (0.0647)	9.5150 (1.8098)	0.3138 (0.2115)	0.3337 (0.1714)	11.0724 (7.6731)	26.5934 (3.9233)	27.4885 (4.0751)
1000	0.3794 (0.2716)	0.3437 (0.2547)	14.3485 (11.4391)	0.2502 (0.1216)	0.3564 (0.1978)	8.8832 (7.1729)	0.3966 (0.0465)	0.2933 (0.0458)	9.2336 (1.3281)	0.3006 (0.1619)	0.3395 (0.1253)	9.1092 (5.7439)	28.5735 (3.0484)	29.5242 (3.1956)
2000	0.3216 (0.2052)	0.2811 (0.1911)	10.0719 (7.6392)	0.2491 (0.0876)	0.3514 (0.1446)	6.8234 (4.7774)	0.3983 (0.0321)	0.2983 (0.0307)	9.0834 (0.9129)	0.3009 (0.1163)	0.3487 (0.0873)	8.1454 (4.1778)	29.6788 (1.5502)	30.6630 (1.6476)
10,000	0.3007 (0.1036)	0.2502 (0.0985)	6.4417 (4.0326)	0.2496 (0.0390)	0.3498 (0.0638)	6.0227 (2.2480)	0.3996 (0.0143)	0.2999 (0.0135)	9.0131 (0.4048)	0.3004 (0.0513)	0.3492 (0.0382)	8.0065 (1.8861)	30.0000 (0.0000)	31.0000 (0.0000)

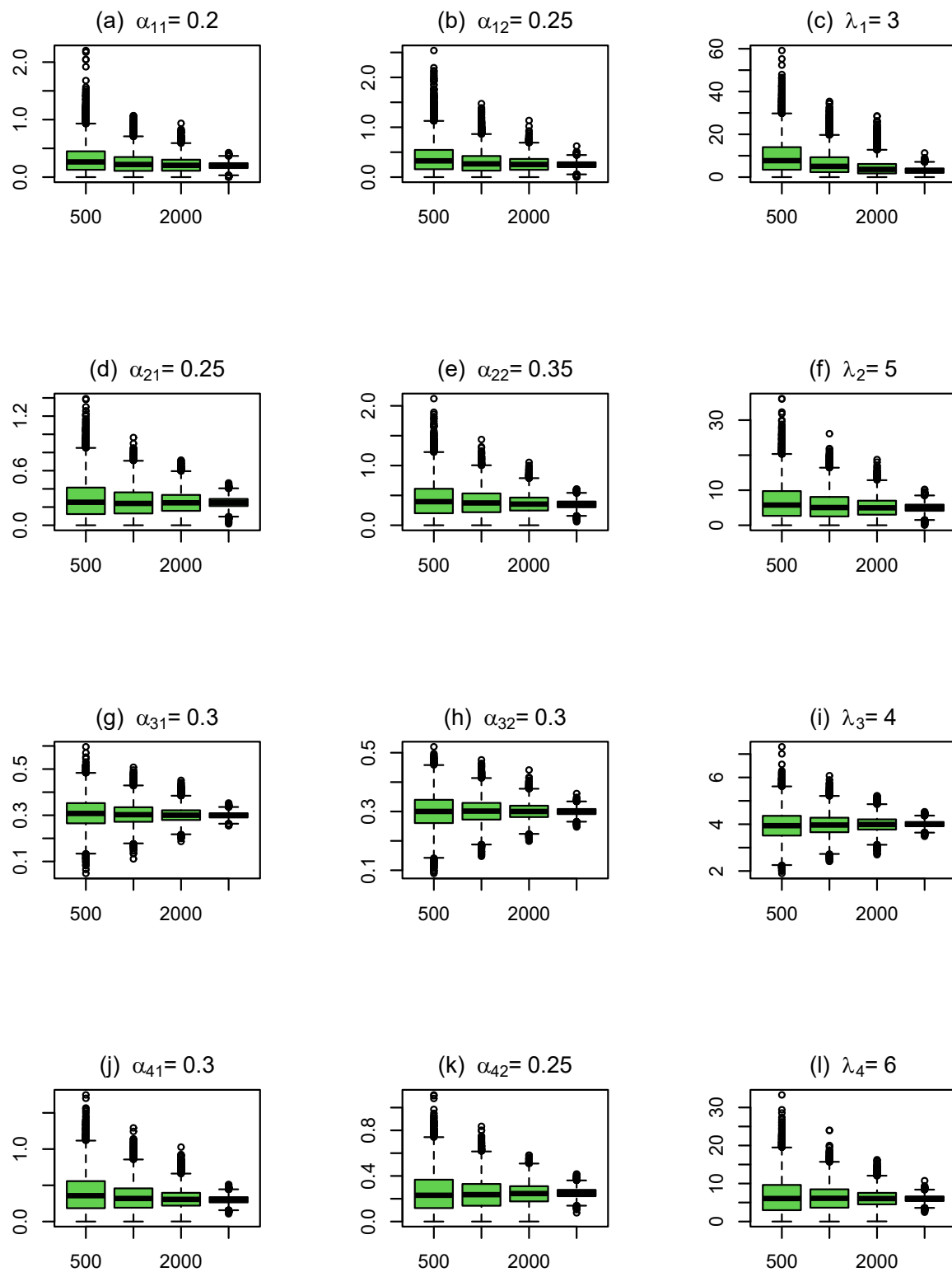


Figure 4. Boxplots of (C3).

5. Two Real Examples

In this section, we use 2-TINAR(2) models to study two stock datasets listed in the New York Stock Exchange (NYSE).

5.1. Siparex Croissance Stock

In this subsection, we consider the daily number of trades of a stock listed in the NYSE (Siparex Croissance). By computation, the mean is 10.0190 and the variance is 129.7295, which shows that this dataset is over-dispersed and implies that it may be better suited to the piecewise structure. Figure 5 shows the path of the data, whose autocorrelation (ACF) and partial autocorrelation functions (PACF) are presented in Figure 6.

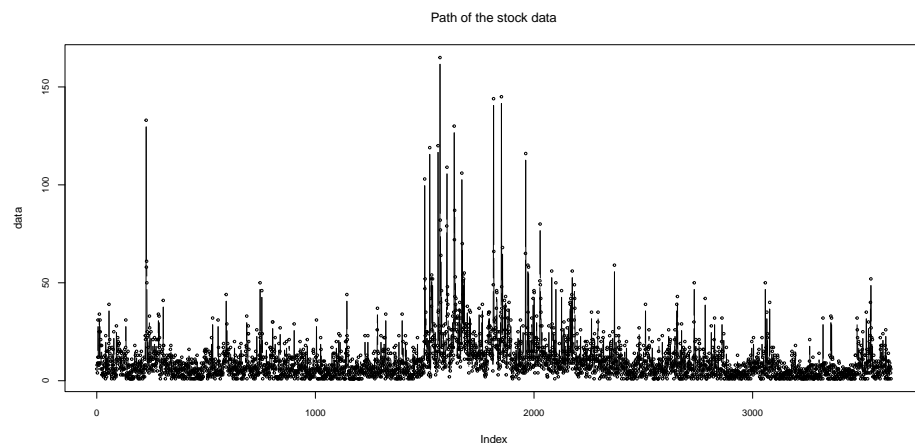


Figure 5. Path of the stock data.

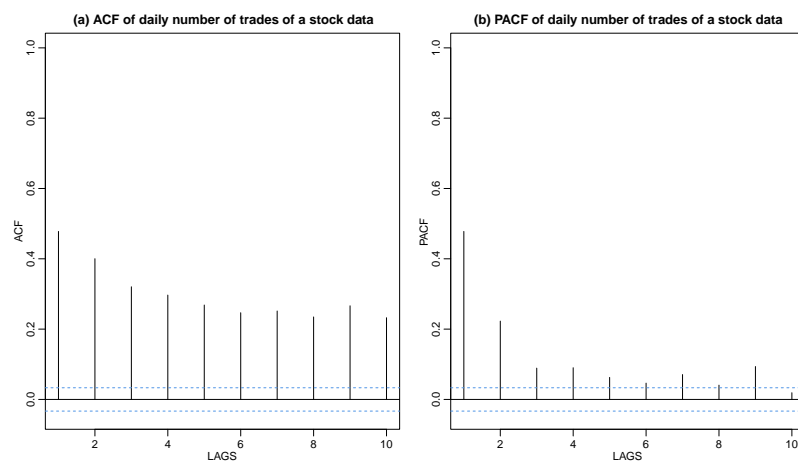


Figure 6. Daily number of trades of a stock data: (a). ACF (b). PACF.

We compare the proposed model with the max-INAR(1) model with geometric innovations (Scotto et al. [8]), the min-INAR(1) model (Aleksić and Ristić [9]), the Poisson INAR(2) (P-INAR) model (Du and Li [4]), and the SETINAR(2,1) model (Monteiro et al. [16]) with $Z_t \in \text{Pois}(\lambda_j)$ to fit the data set by the CLS method and compare their mean squared error (MSE) and mean absolute deviation error (MADE), where

$$MSE = \frac{1}{T-3} \sum_{t=3}^T (x_t - \hat{x}_t)^2, \quad MADE = \frac{1}{T-3} \sum_{t=3}^T |x_t - \hat{x}_t|.$$

For each model, we obtain the values of the CLS estimates of the parameters, include the SD of the estimates, and include the in-sample and out-of-sample MSE and MADE values.

For the in-sample values, all of the observations are used to estimate the parameters, while for the out-of-sample values, the first 3618 observations are used to estimate parameters; we predict that the last $m = 15$ observations, the r for SETINAR(2,1) model, and (r, s) for the 2-TINAR(2) model are obtained by the method described in Section 3.2.

The results of the CLS estimates, MSE and MADE, are summarized in Table 3, from which we can see that the max-INAR(1) model and the min-INAR(1) model are not well fitted, so these two models are not suitable for the kind of datum applied in this paper. The 2-TINAR(2) model takes the smallest MSE and MADE values; hence, 2-TINAR(2) is more appropriate for this data set.

5.2. Westar Energy Stock

In this subsection, we consider the number of trades in 5-min intervals between 9:45 a.m. and 4:00 p.m. of a stock listed in the NYSE (Westar Energy, Inc. (WR)), which belongs to the industry subsector conventional electricity. The time period covered is the first quarter of 2005 (3 January 2005–31 March 2005) with 61 trading days; the sample size is $T = 4575$. Data are taken from the Trades and Quotes (TAQ) dataset. By computation, the mean is 9.6070 and the variance is 34.8908, which shows it is overdispersed and implies that the piecewise structure may be more suitable for this set of data. Figure 7 shows the path of the data, whose autocorrelation (ACF) and partial autocorrelation functions (PACF) are presented in Figure 8.

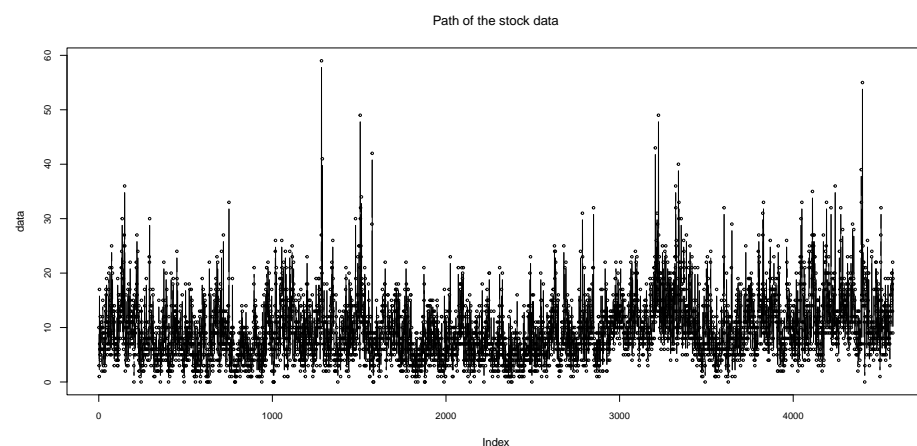


Figure 7. Path of the stock data.

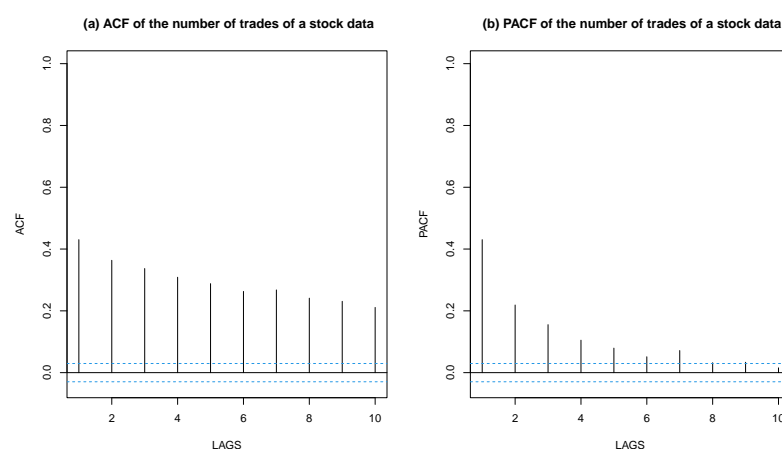


Figure 8. The number of trades of a stock data: (a). ACF (b). PACF.

Like in Section 5.1, we compare the proposed model with the max-INAR(1) model with geometric innovations (Scotto et al. [8]), the min-INAR(1) model (Aleksić and

Ristić [9]), the Poisson INAR(2) (P-INAR) model (Du and Li [4]), and the SETINAR(2,1) model (Monteiro et al. [16]) with $Z_t \in \text{Pois}(\lambda_j)$ to fit the data set by the CLS method and compare their MSE and MADE. For each model, we obtain the values of the CLS estimates of the parameters, include the SD of the estimates, and include the in-sample and out-of-sample MSE and MADE values.

For the in-sample values, all observations are used to estimate parameters, while for the out-of-sample values, the first 4560 observations are used to estimate the parameters; we predict that the last $m = 15$ observations, r for the SETINAR(2,1) model and (r, s) for the 2-TINAR(2) model, are obtained by the method described in Section 3.2. The results of the CLS estimates, MSE and MADE, are summarized in Table 4, from which we can see that the max-INAR(1) model and the min-INAR(1) model are not well fitted, so these two models are not suitable for this kind of data again. The 2-TINAR(2) model takes the smallest MSE and MADE values; hence, 2-TINAR(2) is more appropriate for this data set.

Obviously, we can see that as one of the main novelties of the proposed model, it means that the regime of r and s is determined by considering lots of past information, which makes it more practical. Compared with other models, the 2-TINAR(2) model distinguishes innovations in different regions, which makes our model more flexible, but at the same time it also increases the parameters of the model and makes the model more complex. Therefore, we can find that the 2-TINAR(2) model is more suitable for the analysis of the scattered dataset, such as two real examples mentioned in this paper, but it is not suitable for data with small variances and variations.

Table 3. Fitting results of the Siparex Croissance data.

Model in Sample	Estimate												MSE	MADE
max-INAR(1)	$\hat{\alpha}_1$ 0.4967 (1.8621)	\hat{q} 0.1208 (0.0627)											21.1591	3.9308
min-INAR(1)	$\hat{\alpha}_1$ 1.5074 (50.3267)	$\hat{\mu}$ 8.7789 (414.5962)											37.8780	5.2950
P-INAR(2)	$\hat{\alpha}_1$ 0.3714 (0.0316)	$\hat{\alpha}_2$ 0.2228 (0.0279)	$\hat{\lambda}$ 4.0663 (0.2790)										21.1386	3.8917
SETINAR(2) $r = 9$	$\hat{\alpha}_{11}$ 0.4015 (0.0373)	$\hat{\alpha}_{12}$ 0.2295 (0.0312)	λ_1 3.7634 (0.2895)	$\hat{\alpha}_{21}$ 0.3497 (0.0485)	$\hat{\alpha}_{22}$ 0.2155 (0.0392)	λ_2 5.0422 (1.1459)							21.3504	3.8975
2-TINAR(2) $r = 14, s = 17$	$\hat{\alpha}_{11}$ 0.3544 (0.0739)	$\hat{\alpha}_{12}$ 0.1413 (0.0583)	λ_1 9.0932 (2.4291)	$\hat{\alpha}_{21}$ 0.7353 (0.1802)	$\hat{\alpha}_{22}$ 0.1464 (0.0796)	λ_2 2.8976 (2.5865)	$\hat{\alpha}_{31}$ 0.3666 (0.0379)	$\hat{\alpha}_{32}$ 0.2699 (0.0314)	λ_3 3.7258 (0.2639)	$\hat{\alpha}_{41}$ 0.3490 (0.0630)	$\hat{\alpha}_{42}$ 0.2022 (0.1057)	λ_4 4.2198 (1.5101)	20.5577	3.7951
out of sample														
max-INAR(1)	$\hat{\alpha}_1$ 0.4967 (1.8660)	\hat{q} 0.1208 (0.0629)											25.3317	4.0658
min-INAR(1)	$\hat{\alpha}_1$ 1.5074 (49.6399)	$\hat{\mu}$ 8.7789 (408.9382)											40.6800	4.4990
P-INAR(2)	$\hat{\alpha}_1$ 0.3717 (0.0316)	$\hat{\alpha}_2$ 0.2224 (0.0279)	$\hat{\lambda}$ 4.0767 (0.2796)										22.3961	3.1670
SETINAR(2) $r = 9$	$\hat{\alpha}_{11}$ 0.4058 (0.0374)	$\hat{\alpha}_{12}$ 0.2282 (0.0313)	λ_1 3.7572 (0.2898)	$\hat{\alpha}_{21}$ 0.3497 (0.0485)	$\hat{\alpha}_{22}$ 0.2155 (0.0392)	λ_2 5.0422 (1.1459)							22.9152	3.2409
2-TINAR(2) $r = 14, s = 17$	$\hat{\alpha}_{11}$ 0.3544 (0.0737)	$\hat{\alpha}_{12}$ 0.1413 (0.0582)	λ_1 9.0932 (2.4241)	$\hat{\alpha}_{21}$ 0.7353 (0.1798)	$\hat{\alpha}_{22}$ 0.1464 (0.0795)	λ_2 2.8976 (2.5812)	$\hat{\alpha}_{31}$ 0.3710 (0.0380)	$\hat{\alpha}_{32}$ 0.2681 (0.0315)	λ_3 3.7213 (0.2636)	$\hat{\alpha}_{41}$ 0.3490 (0.0629)	$\hat{\alpha}_{42}$ 0.2022 (0.1055)	λ_4 4.2198 (1.5070)	17.3452	2.9123

Table 4. Fitting results of the WR data.

Model in Sample		Estimate											MSE	MADE		
max-INAR(1)	$\hat{\alpha}_1$	\hat{q}												150.0196	11.5280	
	0.2461 (1.2089)	0.3982 (2.4770)														
min-INAR(1)	$\hat{\alpha}_1$	$\hat{\mu}$												175.1617	12.6286	
	1.7625 (50.1349)	1.6307 (27.6055)														
P-INAR(2)	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\lambda}$										25.1577	4.4766		
	0.3362 (0.0192)	0.2187 (0.0172)	4.2795 (0.1804)													
SETINAR(2) $r = 14$	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	λ_1	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	λ_2								23.9746	4.4016	
	0.3325 (0.0227)	0.2424 (0.0186)	4.0595 (0.1909)	0.2586 (0.0634)	0.1499 (0.0396)	6.9125 (1.1693)										
2-TINAR(2) $r = 11, s = 12$	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	λ_1	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	λ_2	$\hat{\alpha}_{31}$	$\hat{\alpha}_{32}$	λ_3	$\hat{\alpha}_{41}$	$\hat{\alpha}_{42}$	λ_4	23.5336	4.4005		
	0.1855 (0.0848)	0.0458 (0.0919)	10.8460 (2.2813)	0.4619 (0.0962)	0.0985 (0.0585)	5.3952 (1.4106)	0.3116 (0.0233)	0.3009 (0.0225)	3.8049 (0.1971)	0.3660 (0.0884)	0.1952 (0.0942)	4.3789 (1.6380)				
out of sample																
max-INAR(1)	$\hat{\alpha}_1$	\hat{q}												179.5401	12.6882	
	0.2461 (1.2109)	0.3982 (2.4811)														
min-INAR(1)	$\hat{\alpha}_1$	$\hat{\mu}$												178.7166	12.7448	
	1.7625 (50.2126)	1.6307 (27.6483)														
P-INAR(2)	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\lambda}$										73.5537	7.7076		
	0.3359 (0.0192)	0.2172 (0.0172)	4.2877 (0.1808)													
SETINAR(2) $r = 14$	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	λ_1	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	λ_2									73.8454	7.7045
	0.3310 (0.0227)	0.2412 (0.0186)	4.0759 (0.1909)	0.2630 (0.0637)	0.1479 (0.0397)	6.8355 (1.1741)										
2-TINAR(2) $r = 11, s = 12$	$\hat{\alpha}_{11}$	$\hat{\alpha}_{12}$	λ_1	$\hat{\alpha}_{21}$	$\hat{\alpha}_{22}$	λ_2	$\hat{\alpha}_{31}$	$\hat{\alpha}_{32}$	λ_3	$\hat{\alpha}_{41}$	$\hat{\alpha}_{42}$	λ_4	71.6180	7.5889		
	0.1885 (0.0849)	0.0458 (0.0922)	10.7683 (2.3067)	0.4550 (0.0961)	0.0984 (0.0583)	5.4278 (1.4070)	0.3105 (0.0233)	0.3004 (0.0224)	3.8135 (0.1969)	0.3706 (0.0887)	0.1892 (0.0948)	4.3368 (1.6435)				

6. Conclusions

In this paper, we propose a new two-threshold-variable INAR(2) model, which is a generalization of existing INAR models. We consider the CLS estimate with known (r, s) and unknown (r, s) , respectively. To verify the asymptotic behaviour of the estimators, we give the results of the simulation in each case. A superior performance of the proposed model in real example is demonstrated.

In model (1), we use X_{t-1} and X_{t-2} as threshold variables, while other variables can also be used as threshold variables, i.e., the method considered here can be easily extended to other INAR models, such as INAR models with explanatory variable or covariate defined in Enciso-Mora et al. [20],

$$X_t = \alpha \circ X_{t-1} + Z_t, \quad Z_t \sim \text{Pois}(\exp(w_t \gamma)),$$

or

$$X_t = \alpha_t \circ X_{t-1} + Z_t, \quad \alpha_t = [1 + \exp(w_t \delta)]^{-1}, \quad Z_t \sim \text{Pois}(\exp(w_t \gamma)),$$

where w_t is explanatory variable or covariate; then, we can use X_{t-1} and w_t as threshold variables.

Furthermore, there should be more efficient methods of determining the search range for the threshold estimates, and the possibility of extending this model to the high-dimensional situation is worthy of attention. Moreover, we can extend the results to the three-threshold-variable case. These remains topics of future study. Extensions to the models in Chen et al. [21], Qian and Zhu [22], Su and Zhu [23] and Zhang et al. [7] are similar. Details will be discussed in a future project.

Author Contributions: Conceptualization, F.Z.; methodology, F.Z. and J.Z.; software, J.Z. and H.C.; validation, J.Z. and H.C.; formal analysis, J.Z., F.Z. and H.C.; investigation, J.Z. and H.C.; resources, F.Z.; data curation, J.Z. and H.C.; writing—original draft preparation, J.Z.; writing—review and editing, F.Z. and H.C.; visualization, J.Z. and H.C.; supervision, F.Z. All authors have read and agreed to the published version of the manuscript.

Funding: Zhu's work is supported by National Natural Science Foundation of China (No. 12271206), the Natural Science Foundation of Jilin Province (No. 20210101143JC), and the Science and Technology Research Planning Project of Jilin Provincial Department of Education (No. JJKH20231122KJ). Chen's work is supported by Natural Science Foundation of Henan Province (No. 222300420127) and Postdoctoral Research in Henan Province (No. 202103051).

Data Availability Statement: No new data were created in this review.

Acknowledgments: We thank three reviewers for their insightful and constructive comments, which greatly improved the overall presentation.

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

Appendix A

Proof of Proposition 1. (1) It is easy to see that $\{Y_t\}$ is a Markov chain with state space \mathbb{N}_0 , and the transition probability of the 2-TINAR(2) process is

$$P(Y_t = (j, l)^\top | Y_{t-1} = (l, m)^\top) = \sum_{k=0}^{\min(m, j)} \binom{m}{k} \alpha_{j2}^k (1 - \alpha_{j2})^{m-k} P(R = j - k),$$

$$\text{note that } P(R = j - k) = \sum_{b=0}^{\min(l, j-k)} \binom{j-k}{b} \alpha_{j1}^b (1 - \alpha_{j1})^{l-b} P(\epsilon_{jt} = j - k - b),$$

$P(Y_t = (j, l)^\top | Y_{t-1} = (l, m)^\top) > 0$; thus, we can see that $\{Y_t\}$ is an irreducible, aperiodic Markov chain. To prove that Y_t is positive recurrent, let $Y_t = A * Y_{t-1} + (\epsilon_{jt}, 0)^\top$, $\rho(A_j) < 1$,

where $A_j = \begin{pmatrix} \alpha_{j1} & \alpha_{j2} \\ 1 & 0 \end{pmatrix}$. The definition of $*$ can be expressed as $\begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix} * \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} := \begin{pmatrix} \alpha_1 \circ X_1 + \alpha_2 \circ X_2 \\ \alpha_3 \circ X_3 + \alpha_4 \circ X_4 \end{pmatrix}$; then, based on Proposition 2.1 in Yang et al. [18], we can know that $\{Y_t\}$ is a positive recurrent Markov chain.

(2) It is the direct conclusion of (1); thus, we can see the existence of a strictly stationary distribution of (1). \square

Proof of Proposition 2. Note that three of the four values $I_{1t}(r, s)$, $I_{2t}(r, s)$, $I_{3t}(r, s)$, $I_{4t}(r, s)$ must be equal to 0; thus, from Silva and Oliveira [5], let $\max \rho(A_j) < 1$, $\max \rho(A_j \otimes A_l) < 1$, and $\max \rho(A_j \otimes A_l \otimes A_k) < 1$. We can know that $\{X_t\}$ is a 3-order stationarity process. The 3-order joint moment of $X_t, X_{t+s_1}, X_{t+s_2}$, for $s_1, s_2 \in \mathbb{R}$ is a function of 2 variables defined by $\mu_X(s_1, s_2) = E(X_t, X_{t+s_1}, X_{t+s_2})$ with $\mu_X = E(X_t)$. Similar to Silva and Silva [6], we let $\mu_{ji} = E(B_i) = \alpha_{ji}, \sigma_{ji}^2 = \text{Var}(B_i) = \alpha_{ji}(1 - \alpha_{ji}), \gamma_{ji} = E(B_i^3) = \alpha_{ji}, \gamma_{\epsilon_{jt}} = E(\epsilon_{jt}^3), C_{ji} = \sum_{i=1}^2 (\gamma_{ji} - 3\alpha_{ji}\sigma_{ji}^2 - \alpha_{ji}^3)$. Then, for $k > 0$,

$$\begin{aligned} \mu_X(0, 0) &\leq \sum_{i=1}^2 \sum_{l=1}^2 \sum_{k=1}^2 \max(\alpha_{ji}) \max(\alpha_{jl}) \max(\alpha_{jk}) \mu_X(j-l, j-k) \\ &\quad + 3 \sum_{i=1}^2 \sum_{l=1}^2 \max(\alpha_{jl}) \max(\sigma_{ji}^2) \mu_X(i-l) + 3\lambda_j \sum_{i=1}^2 \sum_{l=1}^2 \max(\alpha_{ji}) \max(\alpha_{jl}) \mu_X(i-l) \\ &\quad + 3\lambda_j \sum_{i=1}^2 \max(\sigma_{ji}^2) \mu_X + \sum_{i=1}^2 \max(C_{ji}) \mu_X + \gamma_{\epsilon_{jt}}, \\ \mu_X(0, k) &\leq \sum_{i=1}^2 \max(\alpha_{ji}) \mu_X(0, k-i) + \lambda_j \mu_X(0), \\ \mu_X(k, k) &\leq \sum_{i=1}^2 \sum_{l=1}^2 \max(\alpha_{ji}) \max(\alpha_{jl}) \mu_X(k-i, k-l) + \sum_{i=1}^2 \max(\sigma_{ji}^2) \mu_X(k-i) + 2\lambda_j \mu_X(k), \\ \mu_X(k, m) &\leq \sum_{i=1}^2 \max(\alpha_{ji}) \mu_X(k, m-i) + \lambda_j \mu_X(k), m > k, \end{aligned}$$

the second-order moment of $\{X_t\}$ is $\mu_X(0) \leq \sum_{i=1}^2 \max(\alpha_{ji}) \mu_X(i) + \lambda_j \mu_X + \max(V_{jp})$, where $V_{jp} = \lambda_j + \mu_X \sum_{i=1}^2 \sigma_{ji}^2$. \square

Proof of Proposition 3. (1) and (3) are obvious, so we just present the proofs of (2) and (4), which are obtained by similar arguments after some tedious calculations.

(2). Let

$$\begin{aligned} p_1 &= P(X_{t-1} > r, X_{t-2} > s), p_2 = P(X_{t-1} \leq r, X_{t-2} > s), \\ p_3 &= P(X_{t-1} \leq r, X_{t-2} \leq s), p_4 = P(X_{t-1} > r, X_{t-2} \leq s), \\ u_1 &= E(X_{t-1} | X_{t-1} > r, X_{t-2} > s), u_2 = E(X_{t-1} | X_{t-1} \leq r, X_{t-2} > s), \\ u_3 &= E(X_{t-1} | X_{t-1} \leq r, X_{t-2} \leq s), u_4 = E(X_{t-1} | X_{t-1} > r, X_{t-2} \leq s), \\ u_1^* &= E(X_{t-2} | X_{t-1} > r, X_{t-2} > s), u_2^* = E(X_{t-2} | X_{t-1} \leq r, X_{t-2} > s), \\ u_3^* &= E(X_{t-2} | X_{t-1} \leq r, X_{t-2} \leq s), u_4^* = E(X_{t-2} | X_{t-1} > r, X_{t-2} \leq s), \\ v_1 &= \text{Var}(X_{t-1} | X_{t-1} > r, X_{t-2} > s), v_2 = \text{Var}(X_{t-1} | X_{t-1} \leq r, X_{t-2} > s), \\ v_3 &= \text{Var}(X_{t-1} | X_{t-1} \leq r, X_{t-2} \leq s), v_4 = \text{Var}(X_{t-1} | X_{t-1} > r, X_{t-2} \leq s), \\ v_1^* &= \text{Var}(X_{t-2} | X_{t-1} > r, X_{t-2} > s), v_2^* = \text{Var}(X_{t-2} | X_{t-1} \leq r, X_{t-2} > s), \\ v_3^* &= \text{Var}(X_{t-2} | X_{t-1} \leq r, X_{t-2} \leq s), v_4^* = \text{Var}(X_{t-2} | X_{t-1} > r, X_{t-2} \leq s), \\ w_1 &= E(X_{t-1} X_{t-2} | X_{t-1} > r, X_{t-2} > s), w_2 = E(X_{t-1} X_{t-2} | X_{t-1} \leq r, X_{t-2} > s), \\ w_3 &= E(X_{t-1} X_{t-2} | X_{t-1} \leq r, X_{t-2} \leq s), w_4 = E(X_{t-1} X_{t-2} | X_{t-1} > r, X_{t-2} \leq s). \end{aligned}$$

Hence,

$$\begin{aligned} E(X_t) &= E[E(X_t|\mathcal{F}_{t-1})] = E\left(\sum_{j=1}^4 (\alpha_{j1}X_{t-1} + \alpha_{j2}X_{t-2} + \lambda_j)I_{jt}(r, s)\right) \\ &= \sum_{j=1}^4 I_{jt}(r, s)(\alpha_{j1}EX_{t-1} + \alpha_{j2}EX_{t-2} + \lambda_j) = \sum_{j=1}^4 p_j(\alpha_{j1}u_j + \alpha_{j2}u_j^* + \lambda_j). \end{aligned}$$

(4). According to the variance formula, the variance of X_t is

$$\begin{aligned} \text{Var}(X_t) &= \text{Var}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s) + (\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s) \\ &\quad + (\alpha_{31} \circ X_{t-1} + \alpha_{32} \circ X_{t-2} + \lambda_3)I_{3t}(r, s) + (\alpha_{41} \circ X_{t-1} + \alpha_{42} \circ X_{t-2} + \lambda_4)I_{4t}(r, s)) \\ &= \text{Var}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s)) + \text{Var}((\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s)) \\ &\quad + \text{Var}((\alpha_{31} \circ X_{t-1} + \alpha_{32} \circ X_{t-2} + \lambda_3)I_{3t}(r, s)) + \text{Var}((\alpha_{41} \circ X_{t-1} + \alpha_{42} \circ X_{t-2} + \lambda_4)I_{4t}(r, s)) \\ &\quad + 2\text{Cov}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s), (\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s)) \\ &\quad + 2\text{Cov}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s), (\alpha_{31} \circ X_{t-1} + \alpha_{32} \circ X_{t-2} + \lambda_3)I_{3t}(r, s)) \\ &\quad + 2\text{Cov}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s), (\alpha_{41} \circ X_{t-1} + \alpha_{42} \circ X_{t-2} + \lambda_4)I_{4t}(r, s)) \\ &\quad + 2\text{Cov}((\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s), (\alpha_{31} \circ X_{t-1} + \alpha_{32} \circ X_{t-2} + \lambda_3)I_{3t}(r, s)) \\ &\quad + 2\text{Cov}((\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s), (\alpha_{41} \circ X_{t-1} + \alpha_{42} \circ X_{t-2} + \lambda_4)I_{4t}(r, s)) \\ &\quad + 2\text{Cov}((\alpha_{31} \circ X_{t-1} + \alpha_{32} \circ X_{t-2} + \lambda_3)I_{3t}(r, s), (\alpha_{41} \circ X_{t-1} + \alpha_{42} \circ X_{t-2} + \lambda_4)I_{4t}(r, s)) \\ &:= V_1 + V_2 + V_3 + V_4 + C_1 + C_2 + C_3 + C_4 + C_5 + C_6. \end{aligned} \quad (\text{A1})$$

In the following, we compute these quantities in (A1). First, we have

$$\begin{aligned} V_1 &= \text{Var}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s)) \\ &= \text{Var}(E((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s)|\mathcal{F})) \\ &\quad + E(\text{Var}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s)|\mathcal{F})) \\ &= \text{Var}(\alpha_{11}X_{t-1}I_{1t} + \alpha_{12}X_{t-2}I_{1t} + \lambda_1I_{1t}) + E(\alpha_{11}(1 - \alpha_{11})X_{t-1}I_{1t} + \alpha_{12}(1 - \alpha_{12})X_{t-2}I_{1t} + \lambda_1I_{1t}) \\ &= \alpha_{11}^2 \text{Var}(I_{1t}X_{t-1}) + \alpha_{12}^2 \text{Var}(I_{1t}X_{t-2}) + 2\text{Cov}(\alpha_{11}X_{t-1}I_{1t}, \alpha_{12}X_{t-2}I_{1t}) + 2\text{Cov}(\alpha_{11}X_{t-1}I_{1t}, \lambda_1I_{1t}) \\ &\quad + 2\text{Cov}(\alpha_{12}X_{t-2}I_{1t}, \lambda_1I_{1t}) + \alpha_{11}(1 - \alpha_{11})p_1u_1 + \alpha_{12}(1 - \alpha_{12})p_1u_1^* + \lambda_1p_1 \\ &= \alpha_{11}^2(p_1(v_1 + u_1^2) - p_1^2u_1^2) + \alpha_{12}^2(p_1(v_1^* + (u_1^*)^2) - p_1^2(u_1^*)^2) \\ &\quad + 2(\alpha_{11}\alpha_{12}w_1p_1 - \alpha_{11}\alpha_{12}p_1^2u_1u_1^*) + 2(\alpha_{11}\lambda_1p_1u_1 - \alpha_{11}\lambda_1p_1^2u_1) + 2(\alpha_{12}\lambda_1p_1u_1^* - \alpha_{12}\lambda_1p_1^2u_1^*) \\ &\quad + \alpha_{11}(1 - \alpha_{11})p_1u_1 + \alpha_{12}(1 - \alpha_{12})p_1u_1^* + \lambda_1p_1. \end{aligned} \quad (\text{A2})$$

By the same arguments as above, it follows that

$$\begin{aligned} V_j &= \text{Var}((\alpha_{j1} \circ X_{t-1} + \alpha_{j2} \circ X_{t-2} + \lambda_j)I_{jt}(r, s)) \\ &= \alpha_{j1}^2(p_j(v_j + u_j^2) - p_j^2u_j^2) + \alpha_{j2}^2(p_j(v_j^* + (u_j^*)^2) - p_j^2(u_j^*)^2) \\ &\quad + 2(\alpha_{j1}\alpha_{j2}w_jp_j - \alpha_{j1}\alpha_{j2}p_j^2u_ju_j^*) + 2(\alpha_{j1}\lambda_jp_ju_j - \alpha_{j1}\lambda_jp_j^2u_j) + 2(\alpha_{j2}\lambda_jp_ju_j^* - \alpha_{j2}\lambda_jp_j^2u_j^*) \\ &\quad + \alpha_{j1}(1 - \alpha_{j1})p_ju_j + \alpha_{j2}(1 - \alpha_{j2})p_ju_j^* + \lambda_jp_j, \quad j = 2, 3, 4. \end{aligned} \quad (\text{A3})$$

We can see that C_j takes the form

$$\begin{aligned} C_1 &= 2\text{Cov}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s), (\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s)) \\ &= -2E((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s))E((\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s)) \\ &= -2(\alpha_{11}u_1p_1 + \alpha_{12}u_1^*p_1 + \lambda_1p_1)(\alpha_{21}u_2p_2 + \alpha_{22}u_2^*p_2 + \lambda_2p_2), \end{aligned} \quad (\text{A4})$$

$$\begin{aligned} C_2 &= 2\text{Cov}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s), (\alpha_{31} \circ X_{t-1} + \alpha_{32} \circ X_{t-2} + \lambda_3)I_{3t}(r, s)) \\ &= -2(\alpha_{11}u_1p_1 + \alpha_{12}u_1^*p_1 + \lambda_1p_1)(\alpha_{31}u_3p_3 + \alpha_{32}u_3^*p_3 + \lambda_3p_3), \end{aligned} \quad (\text{A5})$$

$$\begin{aligned} C_3 &= 2\text{Cov}((\alpha_{11} \circ X_{t-1} + \alpha_{12} \circ X_{t-2} + \lambda_1)I_{1t}(r, s), (\alpha_{41} \circ X_{t-1} + \alpha_{42} \circ X_{t-2} + \lambda_4)I_{4t}(r, s)) \\ &= -2(\alpha_{11}u_1p_1 + \alpha_{12}u_1^*p_1 + \lambda_1p_1)(\alpha_{41}u_4p_4 + \alpha_{42}u_4^*p_4 + \lambda_4p_4), \end{aligned} \quad (\text{A6})$$

$$C_4 = 2\text{Cov}((\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s), (\alpha_{31} \circ X_{t-1} + \alpha_{32} \circ X_{t-2} + \lambda_3)I_{3t}(r, s)) \\ = -2(\alpha_{21}u_2p_2 + \alpha_{22}u_2^*p_2 + \lambda_2p_2)(\alpha_{31}u_3p_3 + \alpha_{32}u_3^*p_3 + \lambda_3p_3), \quad (\text{A7})$$

$$C_5 = 2\text{Cov}((\alpha_{21} \circ X_{t-1} + \alpha_{22} \circ X_{t-2} + \lambda_2)I_{2t}(r, s), (\alpha_{41} \circ X_{t-1} + \alpha_{42} \circ X_{t-2} + \lambda_4)I_{4t}(r, s)) \\ = -2(\alpha_{21}u_2p_2 + \alpha_{22}u_2^*p_2 + \lambda_2p_2)(\alpha_{41}u_4p_4 + \alpha_{42}u_4^*p_4 + \lambda_4p_4), \quad (\text{A8})$$

$$C_6 = 2\text{Cov}((\alpha_{31} \circ X_{t-1} + \alpha_{32} \circ X_{t-2} + \lambda_3)I_{3t}(r, s), (\alpha_{41} \circ X_{t-1} + \alpha_{42} \circ X_{t-2} + \lambda_4)I_{4t}(r, s)) \\ = -2(\alpha_{31}u_3p_3 + \alpha_{32}u_3^*p_3 + \lambda_3p_3)(\alpha_{41}u_4p_4 + \alpha_{42}u_4^*p_4 + \lambda_4p_4), \quad (\text{A9})$$

then, (4) follows by replacing (A2)–(A9) in (A1). \square

Proof of Theorem 2. It can easily be seen that the conditions in Klimko and Nelson [24] are verified; we can see that

$$g(\phi, X_{t-1}, X_{t-2}), \frac{\partial g(\phi, X_{t-1}, X_{t-2})}{\partial \phi_i}, \frac{\partial g^2(\phi, X_{t-1}, X_{t-2})}{\partial \phi_i \partial \phi_j}, \frac{\partial g^3(\phi, X_{t-1}, X_{t-2})}{\partial \phi_i \partial \phi_j \partial \phi_k}$$

satisfy all the regularity conditions for $i, j, k = 1, \dots, 12$. Thus, the CLS estimator is strongly consistent. Moreover, when proving asymptotic normality we first have to check the following conditions:

$$(1) \quad E(X_t | X_{t-1}, \dots, X_0) = E(X_t | X_{t-1}, X_{t-2}), t \geq 3, a.s.;$$

$$(2) \quad E(q_i^2(\phi) | \frac{\partial g(\phi, X_{t-1}, X_{t-2})}{\partial \phi_i} \frac{\partial g(\phi, X_{t-1}, X_{t-2})}{\partial \phi_j}) < \infty, i, j = 1, \dots, 12;$$

$$(3) \quad V \text{ is non singular,}$$

then, we know from Klimko and Nelson [24] that the CLS estimation is asymptotically normal with the asymptotic variance $V^{-1}WV^{-1}$. \square

References

1. Steutel, F.W.; van Harn, K. Discrete analogues of self-decomposability and stability. *Ann. Probab.* **1979**, *7*, 893–899. [\[CrossRef\]](#)
2. Al-Osh, M.A.; Alzaid, A.A. First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.* **1987**, *8*, 261–275. [\[CrossRef\]](#)
3. McKenzie, E. Some simple models for discrete variate time series. *Water Resour. Bull.* **1985**, *21*, 645–650. [\[CrossRef\]](#)
4. Du, J.; Li, Y. The integer-valued autoregressive (INAR(p)) model. *J. Time Ser. Anal.* **1991**, *12*, 129–142.
5. Silva, I.; Oliveira, V.L. Difference equations for the higher-order moments and cumulants of the INAR(p) model. *J. Time Ser. Anal.* **2005**, *26*, 17–36. [\[CrossRef\]](#)
6. Silva, I.; Silva, M.E. Parameter estimation for INAR processes based on high-order statistics. *REVSTAT Stat. J.* **2009**, *7*, 105–117.
7. Zhang, J.; Zhu, F.; Mamode Khan, N. A new INAR model based on Poisson-BE2 innovations. *Commun. Stat.-Theory Methods* **2023**, *52*, 6063–6067. [\[CrossRef\]](#)
8. Scotto, M.G.; Weiß, C.H.; Möller, T.A.; Gouveia, S. The max-INAR(1) model for count processes. *TEST* **2018**, *27*, 850–870. [\[CrossRef\]](#)
9. Aleksić, M.S.; Ristić, R.M. A geometric minification integer-valued autoregressive model. *Appl. Math. Model.* **2021**, *90*, 265–280. [\[CrossRef\]](#)
10. Tong, H. On a threshold model. In *Pattern Recognition and Signal Processing*; Sijthoff and Noordhoff: Amsterdam, The Netherlands, 1978.
11. Boero, G.; Marrocu, E. The performance of SETAR models: a regime conditional evaluation of point, interval and density forecasts. *Int. J. Forecast.* **2004**, *20*, 305–320. [\[CrossRef\]](#)
12. Potter, S.M. A nonlinear approach to U.S. GNP. *J. Appl. Econom.* **1995**, *10*, 109–125. [\[CrossRef\]](#)
13. Dueker, M.; Martin, S.; Spagnolo, F. Contemporaneous threshold autoregressive models: estimation, testing and forecasting. *J. Econom.* **2007**, *141*, 517–547. [\[CrossRef\]](#)
14. Tong, H. Threshold models in time series analysis 30 years on. *Stat. Its Interface* **2011**, *4*, 107–118. [\[CrossRef\]](#)
15. Li, D.; Tong, H. Nested sub-sample search algorithm for estimation of threshold models. *Stat. Sin.* **2016**, *26*, 1543–1554. [\[CrossRef\]](#)
16. Monteiro, M.; Scotto, M.G.; Pereira, I. Integer-valued self-exciting threshold autoregressive processes. *Commun. Stat.-Theory Methods* **2012**, *41*, 2717–2737. [\[CrossRef\]](#)
17. Wang, C.; Liu, H.; Yao, J.; Davis, R.A.; Li, W.K. Self-excited threshold Poisson autoregression. *J. Am. Stat. Assoc.* **2014**, *109*, 776–787. [\[CrossRef\]](#)
18. Yang, K.; Wang, D.; Jia, B.; Li, H. An integer-valued threshold autoregressive process based on negative binomial thinning. *Stat. Pap.* **2018**, *59*, 1131–1160. [\[CrossRef\]](#)

19. Zhang, X.; Li, D.; Tong, H. On the least squares estimation of multiple-threshold-variable autoregressive models. *J. Bus. Econ. Stat.* **2023**, 1–14. . [[CrossRef](#)]
20. Enciso-Mora, V.; Neal, P.; Subba Rao, T. Integer valued AR processes with explanatory variables. *Sankhyā* **2009**, *71*, 248–263.
21. Chen, H.; Li, Q.; Zhu, F. Two classes of dynamic binomial integer-valued ARCH models. *Braz. J. Probab. Stat.* **2020**, *34*, 685–711. [[CrossRef](#)]
22. 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](#)]
23. Su, B.; Zhu, F. Comparison of BINAR(1) models with bivariate negative binomial innovations and explanatory variables. *J. Stat. Comput. Simul.* **2021**, *91*, 1616–1634. [[CrossRef](#)]
24. Klimko, L.A.; Nelson, P.I. On conditional least squares estimation for stochastic processes. *Ann. Stat.* **1978**, *6*, 629–642. [[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.