An INAR(1) Time Series Model via a Modified Discrete Burr–Hatke with Medical Applications

Masoumeh Shirozhan[1] · Naushad Ali Mamode Khan[2] · Hassan S. Bakouch[3,4]

Abstract

This paper introduces a flexible discrete transmuted record type discrete Burr–Hatke (TRT-DBH) model that seems suitable for handling over-dispersion and equi-dispersion in count data analysis. Further to the elegant properties of the TRT-DBH, we propose, in the time series context, a first-order integer-valued autoregressive process with TRT-DBH distributed innovations [TRBH-INAR(1)]. The moment properties and inferential procedures of this new INAR(1) process are studied. Some Monte Carlo simulation experiments are executed to assess the consistency of the parameters of the TRBH-INAR(1) model. To further motivate its purpose, the TRBH-INAR(1) is applied to analyze the series of the COVID-19 deaths in Netherlands and the series of infected cases due to the Tularaemia disease in Bavaria. The proposed TRBH-INAR(1) model yields superior fitting criteria than other established competitive INAR(1) models in the literature. Further diagnostics related to the residual analysis and forecasting based on the TRBH-INAR(1) model are also discussed. Based on modified Sieve bootstrap predictors, we provide integer forecasts of future death of COVID-19 and infected of Tularemia.

Keywords Transmuted record type · INAR(1) model · Modified empirical likelihood · COVID-19 disease · Tularemia disease

1 Introduction

Count data are commonly encountered in everyday life phenomena, including insurance, economics, social sciences, medicines, transport and among an unlimited number of areas. In these applications, the count observations are normally expressed as positive integers that are collected on a daily, weekly, or monthly sequential basis. Hence, such repeated observations are more likely serially correlated. On the other side, these series of counts are commonly over-dispersed due to some outliers or presence of physical or latent effects and, in some cases, can be equi-dispersed or under-dispersed as well. Thus, there is an important need to appropriately model the series of counts via the suitable distribution that can accommodate as fully as possible different statistical features.

The prevalent probability models in the literature include the geometric, Poisson, Poisson mixtures (Karlis and Xekalaki 2005), Conway–Maxwell Poisson (Shmueli et al. 2005; Sellers and Shmueli 2010; Sellers et al. 2012) distributions. However, owing to the complex nature and some unique properties of the natural phenomena, such as skewness, dispersion, monotone or unimodal failure rate, inflation or deflation, these conventional density functions may not be fully relevant, as similarly argued in El-Morshedy et al. (2020), Eliwa and El-Morshedy (2021) and the

✉ Masoumeh Shirozhan
ms.shirozhan@gmail.com

Naushad Ali Mamode Khan
n.mamodekhan@uom.ac.mu

Hassan S. Bakouch
hassan.bakouch@science.tanta.edu.eg;
h.bakouch@qu.edu.sa

1 Water and Wastewater Company, Ardabil, Ardabil Province, Iran

2 Department of Economics and Statistics, University of Mauritius, Moka, Mauritius

3 Department of Mathematics, College of Science, Qassim University, Buraydah, Saudi Arabia

4 Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt

references therein. This leads to introducing some other more flexible distributions for positive counts that emerge from discretizing some continuous functions. Examples of such distributions have been comprehensively studied by Gómez-Déniz and Calderín-Ojeda (2011), Chakraborty and Chakravarty (2012), Nekoukhou et al. (2013), Bakouch et al. (2014), Hussain et al. (2016), Bahti and Bakouch (2019), Altun (2020) and references therein.

In this same sense, this paper introduces a modified discrete Burr–Hatke (BH) model, based on the transmuted record type (TRT) constructor introduced by Shakil and Ahsanullah (2011). The discrete version of BH model has been recently proposed by El-Morshedy et al. (2020) to model count events exhibiting huge over-dispersion with various skewness features. The discrete version of BH stands out as a main competitor to the traditional count models as it yields far more superior fitting criteria. This paper focuses on the TRT construction strategy because it leads to skewed distributions and is compatible with one side long-tailed data. The transmuted distributions are special cases of extremal distributions (Kozubowski and Podgórski 2016).

Further to the choice of the probability model, there is a need to investigate the relevant time series structures related to the counts of correlated nature. For repeated count observations, McKenzie (1986), McKenzie (1988) and Al-Osh and Alzaid (1987) introduced the thinning-based INAR(1) processes. The classical INAR(1) model consists of two important components: a survival part that relates the current observation with its previous lagged via the thinning, in particular, the binomial thinning operation (Steutel and van Harn 1979) and a random innovation or error component. In the original INAR(1) model, the innovation was allowed to follow the benchmark Poisson, while the binomial thinning was defined with the fixed or random coefficient.

Over the years, in view of obtaining better fitting criteria or information criteria in severely over-dispersed or zero inflated data series, several authors have proposed a vast number of alterations to either the innovation terms. The INAR(1) model based on the geometric innovations was introduced by Jazi et al. (2012), which can handle the over-dispersed count data sets. The INAR(1) model with the binomial thinning operator and Poisson–Lindley innovation was established by Lívio et al. (2018), and several estimation methods, including the conditional least squares, Yule–Walker and conditional maximum likelihood, were used for estimating the parameters.

Recently, to cover some unique properties of real data sets, other distributions were designated for innovation of the IANR(1) models, such as power series (Bourguignon and Vasconcellos 2015), Poisson-transmuted exponential

(Altun and Mamode Khan 2021) and Bell (Huang and Zhu 2021).

Borges et al. (2017) introduced a new operator called ρ-negative binomial thinning operator and provided a new INAR(1) process with geometric marginals that can be applied for phenomena with excess zeros. Liu and Zhu (2021) introduced a new flexible thinning operator named extended binomial that has two parameters. Considering the extended binomial operator, they defined a new INAR(1) model and estimated the unknown parameters through the two-step conditional least squares and conditional maximum likelihood methods.

Ristić et al. (2013) were the first to propose the INAR(1) process having a dependent count series. Shirozhan et al. (2019) combined the Pegram operator with the dependent thinning operator. A new dependent negative binomial thinning operator based on the inflated geometric counting series is introduced by Shamma et al. (2020).

In classical models, the counting series are expected to be independent, which is not often the case in real-world situations like contagious diseases. Also, the binomial thinning operator is not suitable for zero inflated demands. After an outbreak has gone, we occasionally come across data sets with too many zeros. As a result, we need a count model that can deal with zero inflation. All earlier drawbacks stimulate us to introduce a new INAR(1) model based on the generalized negative binomial (GNB) thinning operator with flexible discrete innovations. The most notable feature of the utilized thinning operator is that it may be used for data sets with additional observations. The principal aim of this paper is devoted to introducing an INAR(1) model with dependent counting series with flexible innovations, where the dependency of the count series makes the thinning operator more suitable for modeling practical count data sets. Some clinical data sets demonstrate the applicability of the suggested model.

The following is the outline of the paper. Section 2 introduces a distribution using the TRT approach and the discrete BH baseline distribution. Also, the survival, hazard rate, probability generating functions and non-central moments of the proposed distribution are provided. The GNB thinning operator is reviewed in Sect. 3. With the proposed discrete innovations, the INAR(1) process is developed, which is based on the GNB thinning operator, and some properties of the process are investigated, including the conditional mean and variance. Several parametric and nonparametric estimation methods for the proposed INAR(1) process are reported in Sect. 4. Finally, in Sect. 5, two real-life count data are utilized to analyze the application of the introduced INAR(1) model, demonstrating our model's suitability in contrast to several relevant INAR(1) models.

2 A Modified Version of Discrete Burr–Hatke Distribution

In this section, we provide a modified discrete distribution based on transmuted record type method and discrete Burr–Hatke baseline distribution. The survival and hazard rate function, along with some statistical properties of the distribution, are also given.

First, we consider the DBH distribution, which is introduced by El-Morshedy et al. (2020). The cumulative distribution function (CDF) and probability mass function (PMF) of DBH distribution are represented, respectively, as

$$H_Y(y, \lambda) = 1 - \frac{\lambda^{y+1}}{y+2}, \quad 0 < \lambda < 1, \ y = 0, 1, 2, \ldots,$$

$$h_Y(y, \lambda) = \left(\frac{1}{y+1} - \frac{\lambda}{y+2} \right) \lambda^y.$$

Now, we review the TRT method, which is defined as

$$Z \overset{d}{=} \begin{cases} Y_{U_{(1)}} & \text{w.p.} \quad 1 - \gamma \\ Y_{U_{(2)}} & \text{w.p.} \quad \gamma \end{cases}, \qquad 0 < \gamma < 1,$$

where $Y_{U_{(1)}}$ and $Y_{U_{(2)}}$ are, respectively, the first and second upper records [for more details, see Shakil and Ahsanullah (2011)].

Hence, the CDF of TRT distribution is shown as

$$F_Z(z, \lambda, \gamma) = H_Y(z) + \gamma(1 - H_Y(z)) \ln(1 - H_Y(z)),$$

where $H_Y(z)$ is an arbitrary CDF baseline distribution. Consider the DBH baseline distribution, the CDF of the proposed distribution is represented as

$$F_z(z, \lambda, \gamma) = 1 - \frac{\lambda^{z+1}}{z+2} + \gamma \left(\frac{\lambda^{z+1}}{z+2} \right) \ln \left(\frac{\lambda^{z+1}}{z+2} \right)$$

$$= 1 - \frac{\lambda^{z+1}}{z+2} \left[1 - \gamma \ln \left(\frac{\lambda^{z+1}}{z+2} \right) \right],$$

and the PMF is

$$f_z(z, \lambda, \gamma) = F_z(z, \lambda, \gamma) - F_z(z - 1, \lambda, \gamma)$$

$$= \lambda^z \left[\frac{1}{z+1} - \frac{\lambda}{z+2} - \gamma \left(\frac{1}{z+1} \ln \left(\frac{\lambda^z}{z+1} \right) \right. \right.$$

$$\left. \left. - \frac{\lambda}{z+2} \ln \left(\frac{\lambda^{z+1}}{z+2} \right) \right) \right].$$

We call this distribution as the transmuted record type-discrete Burr–Hatke.

The survival and hazard rate functions (HRF) of the TRT-DBH are demonstrated as below

$$S(z, \lambda, \gamma) = \frac{\lambda^{z+1}}{z+2} \left[1 - \gamma \ln \left(\frac{\lambda^{z+1}}{z+2} \right) \right],$$

$$\text{HRF } (z, \lambda, \gamma) = \frac{f_z(z, \lambda, \gamma)}{S(z-1, \lambda, \gamma)}$$

$$= 1 - \frac{\lambda(z+1) \left(1 - \gamma \ln \left(\frac{\lambda^{z+1}}{z+2} \right) \right)}{(z+2) \left(1 - \gamma \ln \left(\frac{\lambda^z}{z+1} \right) \right)}.$$

The PMF and HRF plots of TRT-DBH distribution are depicted in Figs. 1 and 2, for different combinations of the parameters.

It is clear that the HRF of TRT-DBH distribution has different shapes, including decreasing and unimodal, which dedicate the capability of TRT-DBH distribution to model different types of data sets.

Fig. 1 The PMF plots for TRT-DBH distribution with different combinations of parameters (λ, γ)
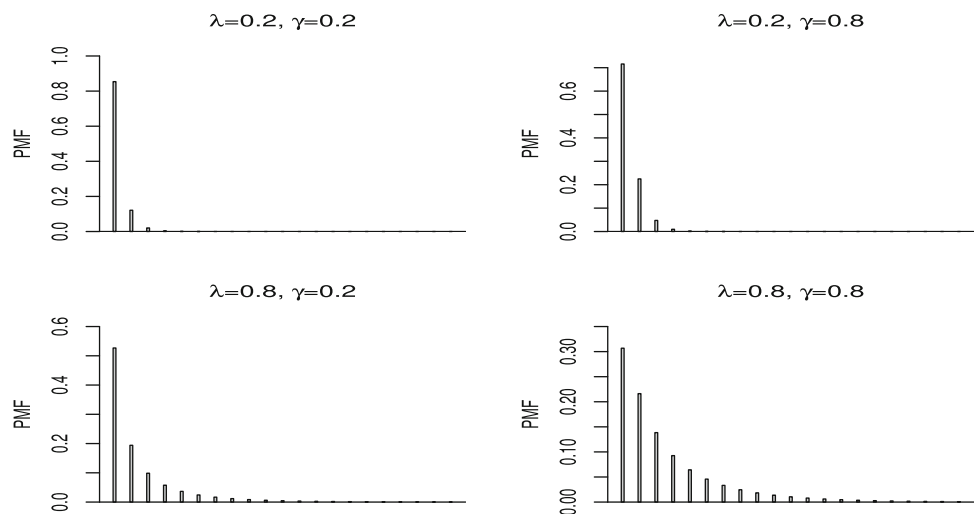
Fig. 2 The HRF plots for TRT-DBH distribution with different combinations of parameters (λ, γ)
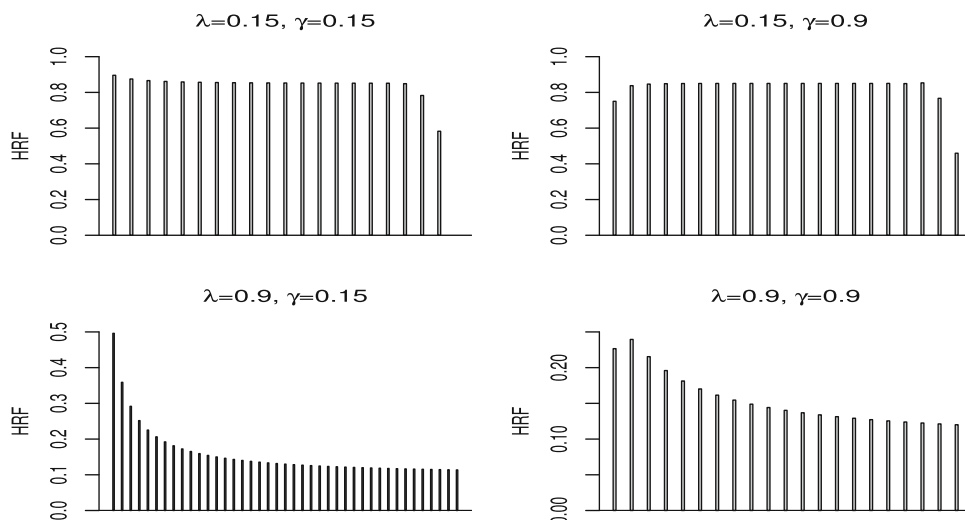
2.1 Some Statistical Properties of TRT-DBH Distribution

Now, some properties of the TRT-DBH distribution are investigated, such as the probability generating function, r-th non-central moments, and so on.

Let Z follow the TRT-DBH distribution with the parameters (λ, γ), then its probability generating function is obtained as follows

$$
\begin{aligned}
G_z(s) &= E(S^Z) \\
&= \sum_{z=0}^{\infty} s^z \left[\frac{\lambda^z}{z+1} - \frac{\lambda^{z+1}}{z+2} \right] - \gamma \sum_{z=0}^{\infty} s^z \\
&\quad \left[\frac{\lambda^z}{z+1} \ln\left(\frac{\lambda^z}{z+1} \right) - \frac{\lambda^{z+1}}{z+2} \ln\left(\frac{\lambda^{z+1}}{z+2} \right) \right] \\
&= \frac{1}{S\lambda} \left(1 - \frac{1}{s} \right) \sum_{z=0}^{\infty} \frac{(s\lambda)^{z+1}}{z+1} + \frac{1}{s} \\
&\quad - \gamma \ln(\lambda) \left(1 - \frac{1}{s} \right) \sum_{z=0}^{\infty} \left(1 - \frac{1}{z+1} \right) (S\lambda)^z \\
&\quad + \gamma \left(1 - \frac{1}{s} \right) \sum_{z=0}^{\infty} \frac{\ln(z+1)(s\lambda)^z}{z+1} \\
&= \left(1 - \frac{1}{s} \right) \left\{ \frac{-1}{s\lambda} \ln(1 - s\lambda) - \gamma \ln(\lambda) \right. \\
&\quad \left[\frac{1}{1 - s\lambda} - \frac{1}{s\lambda} \ln(1 - s\lambda) \right] \\
&\quad \left. - \gamma S\lambda \, \Phi'(s\lambda, 1, 2) \right\} + \frac{1}{s}, \qquad |s| < 1
\end{aligned}
$$

where $\Phi(a, b, c) = \sum_{n=0}^{\infty} \frac{a^n}{(n+c)^b}$, $|a| < 1$ is the LerchPhi function and we denote its derivative as

$$
\Phi'(a, b, c) = \frac{\partial}{\partial b} \Phi^{(0,1,0)}(a, b = b_0, c) \mid_{b=b_0} .
$$

Noted that both λ and s are restricted $(0 < \lambda < 1, |s| < 1)$; hence, the condition of the LerchPhi function is satisfied.

The r-th non-central moments of TRT-DBH distribution are represented as

$$
\begin{aligned}
E(Z^r) &= \sum_{z=0}^{\infty} z^r f_z(z, \lambda, \gamma) \\
&= \sum_{z=0}^{\infty} \left(z^r - (z-1)^r \right) \frac{\lambda^z}{z+1} \\
&\quad - \gamma \sum_{z=0}^{\infty} \left(z^r - (z-1)^r \right) \frac{\lambda^z}{z+1} \ln\left(\frac{\lambda^z}{z+1} \right).
\end{aligned}
$$

It is concluded that the first and second moments of the TRT-DBH distribution are as follows

$$
\begin{aligned}
\mu_z = E(Z) &= -1 - \frac{\ln(1-\lambda)}{\lambda} - \gamma \left[\ln(\lambda) \right. \\
&\quad \left(\frac{1}{1-\lambda} + \frac{\ln(1-\lambda)}{\lambda} \right) + \lambda \, \Phi'(\lambda, 1, 2) \right], \\
E(Z^2) &= \frac{3-\lambda}{1-\lambda} + \frac{3\ln(1-\lambda)}{\lambda} \\
&\quad - \gamma \ln(\lambda) \left(\frac{5\lambda - 3}{(1-\lambda)^2} - \frac{3\ln(1-\lambda)}{\lambda} \right) \\
&\quad - \gamma \left(2\,\Phi'(\lambda, 0, 1) - 3\lambda\,\Phi'(\lambda, 1, 2) \right).
\end{aligned} \qquad (1)
$$

Accordingly, based on the first and second moments, the variance of TRT-DBH distribution can be obtained in closed form. The Fisher dispersion index (FDI) is defined as the variance to mean ratio, which indicates whether a certain distribution is suitable for under or over-dispersed data sets. If FDI $< (>)1$, the distribution is under-dispersed (over-dispersed).

The numerical mean, variance, skewness, kurtosis and FDI of TRT-DBH distribution are provided in Table 1, for different combinations of the parameters. Based on

Table 1 Some statistical properties of the TRT-DBH distribution

Measures	(λ, γ)								
	(0.1, 0.1)	(0.1, 0.5)	(0.1, 0.9)	(0.5, 0.1)	(0.5, 0.5)	(0.5, 0.9)	(0.9, 0.1)	(0.9, 0.5)	(0.9, 0.9)
Mean	0.07072	0.13916	0.20761	0.46341	0.77189	1.08036	1.86500	3.09126	4.31753
Variance	0.07830	0.15151	0.21535	0.87540	1.49050	1.91528	14.5526	27.2993	37.0385
Skewness	4.34336	3.03405	2.34709	3.05341	2.39579	1.98291	4.55818	3.60665	3.07481
Kurtosis	26.2771	15.3736	11.2771	23.3821	20.0279	20.3337	47.7409	35.8777	32.8224
FDI	1.10723	1.08872	1.03728	1.88904	1.93099	1.77282	7.80303	8.83112	8.57865
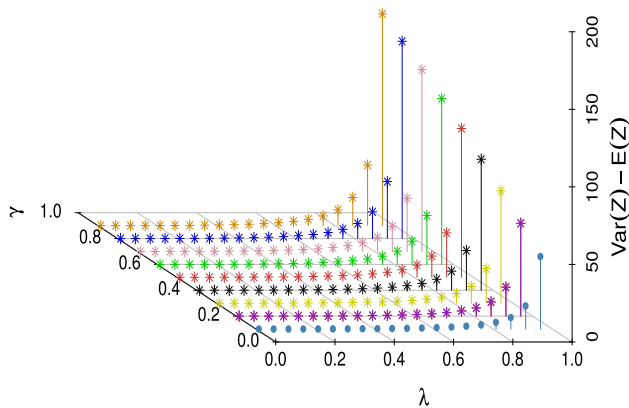
Fig. 3 The $Var(Z) - E(Z)$ plots for TRT-DBH distribution with different combinations of parameters (λ, γ)

Table 1, the mean and variance of TRT-DBH are increased by increasing values of parameters λ and γ. Also, for small values of λ and large values of γ, the FDI measure is near to one, which indicates the equi-dispersion of TRT-DBH distribution. For other combinations of (λ, γ), the FDI measure is more than one, so the TRT-DBH distribution is over-dispersion. The TRT-DBH distribution is severely skewed to right and leptokurtic. So, TRT-DBH distribution also has a perfect fit for right long-tailed data.

Based on Fig. 3, the values of $Var(Z) - E(Z)$ are always positive, which confirms the results of Table 1 and the over-dispersion nature of the TRT-DBH model.

3 Formulation of the INAR(1) Model with TRT-DBH Innovation

The purpose of this section is to introduce an INAR(1) time series model based on the TRT-DBH distribution. First, we review the definition of the GNB thinning operator defined by Shamma et al. (2020), and then, an INAR(1) model with TRT-DBH innovation is constructed.

Definition 1 (Shamma et al. 2020) Consider a sequence of independent identically distributed (iid) geometric random variables $\{V_i\}_{i\in\mathbb{N}}$ with parameter $\frac{\theta}{1+\theta}$ and Bernoulli random variable W with parameter $\frac{\alpha}{\theta}$, $0 \leq \alpha \leq \theta \leq 1$, where V_i and W are independent for all $i \in \mathbb{N}$. Define a sequence of dependent random variables $\{U_i\}_{i\in\mathbb{N}}$ as $U_i = V_i W$, $i \in \mathbb{N}$. It can be verified U_i has a mixture distribution as follows

$$P(U_i = u) = \begin{cases} 1 - \dfrac{\alpha}{1+\theta} & u = 0 \\ (\dfrac{\alpha}{\theta}) \dfrac{\theta^u}{(1+\theta)^{u+1}} & u = 1, 2, \ldots \end{cases}, \quad (2)$$

denoted as zero inflated geometric distribution $\left(ZIG(1 - \frac{\alpha}{\theta}, \frac{\theta}{1+\theta})\right)$.

Also,

$$E(U_i) = \alpha, Var(U_i) = \alpha(2\theta - \alpha + 1),$$
$$Cov(U_i, U_j) = \alpha(\theta - \alpha), \ i \neq j.$$

The random variable $\sum_{i=1}^{n} U_i$ is a mixture of zero and negative binomial $(n, \frac{\theta}{1+\theta})$ distributed random variables with proportions $1 - \frac{\alpha}{\theta}$ and $\frac{\alpha}{\theta}$, respectively, and the zero inflated negative binomial distribution is the name given to it.

Definition 2 (GNB thinning operator) Let X be a non-negative integer valued random variable and $\{U_i, i \in \mathbb{N}\}$ as noted above by (2). The operator "$\alpha *_\theta$", $0 \leq \alpha \leq \theta \leq 1$, defined as $\alpha *_\theta X_{t-1} = \sum_{i=1}^{X_{t-1}} U_{i,t}$ is called the GNB thinning operator.

Shamma et al. (2020) outlined the properties of the GNB thinning operator

3.1 The Proposed INAR(1) Model

The following recursive equation introduces the proposed stationary INAR(1) process $\{X_t\}$ as

$$X_t = \alpha *_\theta X_{t-1} + Z_t, \qquad \alpha < \theta < 1, \ t \geq 1, \quad (3)$$

where "$*_\theta$" is the GNB thinning operator, $\{Z_t\}$ be a sequence of TRT-DBH random variables with parameters (λ, γ) and given X_{t-1}, the random variables $\alpha *_\theta X_{t-1}$ and Z_t are independent of each other. We shall refer to this model as TRBH-INAR(1).

The one-step transition probabilities are

$$P_{0j} = P(X_t = j | X_{t-1} = 0) = P(Z_t = j),$$

and for $i \geq 1$, we get

$$P_{ij} = P(X_t = j | X_{t-1} = i) = (1 - \frac{\alpha}{\theta})P(Z_t = j) \\ + \frac{\alpha}{\theta} \sum_{k=0}^{j} \binom{i-1}{i+k-1} \frac{\theta^k}{(1+\theta)^{i+k}} P(Z_t = j - k), \quad (4)$$

where

$$P(Z_t = j) = \lambda^j \left[\frac{1}{j+1} - \frac{\lambda}{j+2} - \gamma \left(\frac{1}{j+1} \ln \left(\frac{\lambda^j}{j+1} \right) \right. \right. \\ \left. \left. - \frac{\lambda}{j+2} \ln \left(\frac{\lambda^{j+1}}{j+2} \right) \right) \right].$$

This model may be fitted to infectious illness data and can be used to describe the disease's transmission as follows: In the case of the INAR(1) model, if X_{t-1} represents the number of new patients throughout the time span $(t-2, t-1]$, $\alpha *_\theta X_{t-1}$ will be the number of surviving patients from the previous month, which may stimulate new patients or likely cure, and $\{Z_t\}$ will be the number of new patients infected in the current period.

Remark 1 Shamma et al. (2020) provided several properties of the GNB thinning operator as follows

  (i) $E(\alpha *_\theta X \mid X) = \alpha X$,
  (ii) $Var(\alpha *_\theta X \mid X) = \alpha(\theta - \alpha)X + \alpha(\theta + 1)X$,
  (iii) $E(\alpha *_\theta X) = \alpha E(X)$,
  (iv) $E(\alpha *_\theta X)$
            $= \alpha(\theta - \alpha)E^2(X) + \alpha(\theta + 1)E(X) + \alpha\theta Var(X)$.

The expectation and variance of the process $\{X_t\}$ is obtained as

$$E(X) = \frac{\mu_z}{1 - \alpha},$$
$$Var(X) = \frac{\alpha\mu_z(1 + \theta)}{(1-\alpha)(1-\alpha\theta)} + \frac{\alpha(\theta - \alpha)\mu_z^2}{(1-\alpha)^2(1-\alpha\theta)} \\ + \frac{\sigma_Z^2}{1 - \alpha\theta},$$

where μ_z and σ_Z^2 are the mean and variance of TRT-DBH distribution, respectively.

Proposition 1 *The Fisher dispersion index of $\{X_t\}$ is obtained as*

$$I_X = \frac{Var(X)}{E(X)} \\ = \frac{\alpha(\theta - \alpha)\mu_z^2 + \alpha(\theta + 1)(1 - \alpha)\mu_z + (1 - \alpha)^2 \sigma_Z^2}{(1 - \alpha\theta)(1 - \alpha)\mu_z}.$$

This readily demonstrates that I_X is more than one and obviously is over-dispersed.

Proof In order to confirm $I_Z X \geq 1$, it is required to show $Var(X) - E(X) \geq 0$. Hence, we show the following inequality always holds

$$\alpha(\theta - \alpha)\mu_z^2 + (1 - \alpha)(2\alpha\theta - 1 + \alpha)\mu_z + (1 - \alpha)^2 \sigma_Z^2 \geq 0,$$

which can be rewritten as

$$\alpha(\theta - \alpha)\mu_z^2 + (1 - \alpha)^2 \left(\sigma_Z^2 - \mu_z \right) + 2\alpha\theta(1 - \alpha)\mu_z \geq 0.$$

Since the TRT-DBH model is over-dispersed or equi-dispersed, so $I_Z > 1$ (i.e., $\sigma_z^2 - \mu_z \geq 0$), and the proof is concluded.   □

Proposition 2 *Suppose $\{X_t\}$ is a stationary process defined by (3), then for $\alpha < \theta < 1$ and $t \geq 1$,*

  (i)   The conditional expectation is

$$E(X_t \mid X_{t-k}) = \alpha^k X_{t-k} + \frac{1 - \alpha^k}{1 - \alpha} \mu_z. \quad (5)$$

When $k \to \infty$, then $\lim_{k \to \infty} E(X_t \mid X_{t-k}) = \frac{\mu_z}{1 - \alpha}$, which is the process's unconditional expectation.

  (ii)   The conditional variance is

$$Var(X_t \mid X_{t-1}) = \alpha(\theta - \alpha)X_{t-1}^2 \\ + \alpha(\theta + 1)X_{t-1} + \sigma_Z^2, \quad (6)$$

and

$$Var(X_t | X_{t-k}) = \alpha^k(\theta^k - \alpha^k)X_{t-k}^2 + \alpha^k \frac{(1+\theta)(1-\theta^k)}{1-\theta} X_{t-k} \\ + 2\mu_z \alpha^k \left(\frac{1-\theta^k}{1-\theta} - \frac{1-\alpha^k}{1-\alpha} \right) X_{t-k} \\ + \frac{\alpha\mu_z(1+\theta)}{1-\theta} \left(\frac{1-\alpha^{k-1}}{1-\alpha} - \frac{\theta(1-(\alpha\theta)^{k-1})}{1-\alpha\theta} \right) \\ + \mu_z^2 \left(\frac{1-(\alpha\theta)^{k-1}}{1-\alpha\theta} - \frac{1-\alpha^{2k-2}}{1-\alpha^2} \right) \\ + 2\alpha\mu_z^2 \left[\frac{\theta}{1-\theta} \left(\frac{\alpha(1-\alpha^{k-1})}{1-\alpha} - \frac{\alpha\theta(1-(\alpha\theta)^{k-1})}{1-\alpha\theta} \right) \right. \\ \left. - \frac{\alpha}{1-\alpha} \left(\frac{\alpha(1-\alpha^{k-1})}{1-\alpha} - \frac{\alpha\theta(1-\alpha^{2k-2})}{1-\alpha^2} \right) \right] \\ + \sigma_Z^2 \frac{1-(\alpha\theta)^k}{1-\alpha\theta}.$$

hence,

$$\lim_{k \to \infty} Var(X_t \mid X_{t-k}) = \frac{\alpha \mu_z (1 + \theta)}{(1 - \alpha)(1 - \alpha\theta)}$$
$$+ \frac{\alpha(\theta - \alpha)\mu_z^2}{(1 - \alpha)^2(1 - \alpha\theta)} + \frac{\sigma_Z^2}{1 - \alpha\theta},$$

which is the process's unconditional variance.

(iii) The autocorrelation function of the process $\{X_t\}$ is represented as

$$\rho(k) = Corr(X_t, X_{t-k}) = \alpha^k.$$

Proof See Appendix A. □

3.2 Different Estimation Method

The conditional maximum likelihood, modified conditional least square, modified maximum empirical likelihood, and Yule–Walker estimation procedures for the parameters of the TRBH-INAR(1) model are discussed in this section.

3.2.1 Conditional Maximum Likelihood Estimation

The log-likelihood function is maximized in terms of the model parameters $\delta = (\alpha, \theta, \lambda, \gamma)$ in order to produce conditional maximum likelihood (CML) estimators. The log-likelihood function for sample observation $X_1, \ldots, X_n$ from the TRBH-INAR(1) model can be written as

$$\ell(\delta) = \log L(\delta \mid X_2, \ldots, X_n)$$
$$= \sum_{t=2}^{n} \log P(X_t = j \mid X_{t-1} = i),$$

where $P(X_t = j \mid X_{t-1} = i)$ is transition probability given by (4). The CML estimator of the unknown parameters are numerically obtained by maximizing the log-likelihood function with commands "nlm" or "optim" from statistical package "R".

3.2.2 Modified Conditional Least Square Estimation

The modified conditional least squares (MCLS) estimators of the parameters α, μ_z are found by minimizing the expression below

$$Q(\alpha, \mu_z) = \sum_{t=2}^{n} (X_t - E(X_t \mid X_{t-1}))^2$$
$$= \sum_{t=2}^{n} (X_t - \alpha X_{t-1} - \mu_z)^2, \tag{7}$$

where μ_z is a function of the parameters (λ, γ). The estimators are given by

$$\hat{\alpha}_{MCLS} = \frac{(n-1)\sum_{t=2}^{n} X_t X_{t-1} - \sum_{t=2}^{n} X_t \sum_{t=2}^{n} X_{t-1}}{(n-1)\sum_{t=2}^{n} X_{t-1}^2 - \left(\sum_{t=2}^{n} X_{t-1}\right)^2},$$

and

Table 2 Results of simulations of the TRBH-INAR(1) model's parameter estimations

n	CML					MMEL				
	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$	R.Time	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$	R.Time
$(\alpha, \theta, \lambda, \gamma) = (0.4, 0.8, 0.7, 0.3)$										
100	0.43075	0.81661	0.76381	0.28297	134.69	0.41534	0.77461	0.69378	0.31683	97.01
MSE	(0.01358)	(0.03378)	(0.00858)	(0.00355)		(0.01301)	(0.03197)	(0.00928)	(0.00232)	
200	0.42209	0.82608	0.74652	0.31603	311.03	0.41129	0.78035	0.70572	0.31138	153.86
MSE	(0.00874)	(0.01564)	(0.00618)	(0.00309)		(0.00975)	(0.01652)	(0.00839)	(0.00187)	
500	0.42182	0.81377	0.72775	0.30616	857.82	0.41028	0.81293	0.70374	0.30929	493.41
MSE	(0.00312)	(0.00602)	(0.00509)	(0.00219)		(0.00428)	(0.00943)	(0.00625)	(0.00138)	
1000	0.41721	0.80927	0.70715	0.30123	2115.21	0.41112	0.79797	0.69651	0.30199	652.44
MSE	(0.00218)	(0.00259)	(0.00477)	(0.00166)		(0.00235)	(0.00479)	(0.00439)	(0.00104)	
$(\alpha, \theta, \lambda, \gamma) = (0.2, 0.4, 0.9, 0.6)$										
100	0.23824	0.45879	0.89641	0.61591	112.15	0.23941	0.36262	0.94042	0.63544	65.74
MSE	(0.00799)	(0.03091)	(0.00082)	(0.03401)		(0.00952)	(0.04456)	(0.00398)	(0.04425)	
200	0.22906	0.44981	0.89978	0.60062	266.47	0.25133	0.37088	0.93545	0.62737	124.34
MSE	(0.00527)	(0.01183)	(0.00036)	(0.02152)		(0.00553)	(0.03457)	(0.00259)	(0.01327)	
500	0.22899	0.43896	0.90223	0.59378	763.11	0.22391	0.42868	0.92714	0.62215	460.02
MSE	(0.00356)	(0.00559)	(0.00014)	(0.01337)		(0.00526)	(0.01953)	(0.00189)	(0.00761)	
1000	0.21661	0.41588	0.90277	0.60904	1932.19	0.21868	0.41706	0.91664	0.61247	557.86
MSE	(0.00275)	(0.00354)	(0.00071)	(0.01015)		(0.00331)	(0.01366)	(0.00092)	(0.00285)	

Table 3 Results of simulations of the TRBH-INAR(1) model's parameter estimations

n	MCLS					YW				
	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$	R.Time	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$	R.Time
$(\alpha, \theta, \lambda, \gamma) = (0.4, 0.8, 0.7, 0.3)$										
100	0.34643	0.76204	0.80737	0.30327	273.25	0.38488	0.76703	0.78856	0.28814	60.14
MSE	(0.02229)	(0.37431)	(0.01459)	(0.00854)		(0.01532)	(0.28766)	(0.01158)	(0.00813)	
200	0.36568	0.78433	0.75426	0.29853	420.77	0.39773	0.76522	0.77414	0.29876	102.7
MSE	(0.012535)	(0.10733)	(0.01279)	(0.00822)		(0.00978)	(0.01932)	(0.01058)	(0.00769)	
500	0.38148	0.78581	0.72288	0.29787	847.28	0.39588	0.79944	0.75931	0.29683	199.61
MSE	(0.00591)	(0.03619)	(0.01158)	(0.00726)		(0.00441)	(0.01024)	(0.01088)	(0.00621)	
1000	0.39085	0.79011	0.71332	0.295236	1324.41	0.39865	0.80364	0.75087	0.29504	258.23
MSE	(0.00299)	(0.01821)	(0.01136)	(0.00602)		(0.00379)	(0.00823)	(0.01009)	(0.00589)	
$(\alpha, \theta, \lambda, \gamma) = (0.2, 0.4, 0.9, 0.6)$										
100	0.26071	0.32083	0.91666	0.63862	218.46	0.23588	0.36791	0.91948	0.56604	56.23
MSE	(0.01159)	(0.05267)	(0.00919)	(0.06116)		(0.01272)	(0.06227)	(0.00721)	(0.07331)	
200	0.24713	0.34577	0.91059	0.62406	387.89	0.21343	0.44966	0.92816	0.55631	94.93
MSE	(0.00761)	(0.05147)	(0.00825)	(0.06111)		(0.00628)	(0.05396)	(0.00606)	(0.06303)	
500	0.22984	0.35921	0.91424	0.62141	756.62	0.20261	0.43879	0.91098	0.54532	128.81
MSE	(0.00685)	(0.05224)	(0.00717)	(0.05103)		(0.00587)	(0.03977)	(0.00402)	(0.06297)	
1000	0.20334	0.36305	0.92775	0.59396	1251.96	0.19814	0.43378	0.93276	0.53304	160.28
MSE	(0.00649)	(0.04271)	(0.00623)	(0.04092)		(0.00558)	(0.02375)	(0.00294)	(0.04302)	

$$\hat{\mu}_{Z,MCLS} = \frac{\sum_{t=2}^{n} X_t - \hat{\alpha}_{MCLS} \sum_{t=2}^{n} X_{t-1}}{n-1}.$$

It is worth mention that the MCLS estimation of parameters (λ, γ) is obtained by finding the root of the equation (1) equal to the estimation of the $\hat{\mu}_{Z,MCLS}$.

The one-step conditional expectation of the process depends only on the parameters α and μ_z and is not possible to use it for the estimation of the parameter θ. Hence, the parameter θ can be estimated under the modified method proposed by Karlsen and Tjøstheim (1988). The parameter θ can be estimated by minimizing the following expression

$$T(\theta) = \sum_{t=2}^{n} (V_t - Var(X_t|X_{t-1}))^2, \tag{8}$$

where

$$V_t = (X_t - E(X_t|X_{t-1}))^2$$
$$= \left(X_t - \hat{\alpha}_{MCLS} X_{t-1} - \hat{\mu}_{Z,MCLS} \right)^2,$$

and $Var(X_t|X_{t-1})$ is defined in (6) with estimated values of the parameters $(\alpha, \lambda, \gamma)$ as below

$$Var(X_t \mid X_{t-1}) = \hat{\alpha}_{MCLS}(\theta - \hat{\alpha}_{MCLS}) X_{t-1}^2$$
$$+ \hat{\alpha}_{MCLS}(\theta + 1) X_{t-1} + \hat{\sigma}_{Z,MCLS}^2,$$

where σ_Z^2 is a function of the parameters (λ, γ) and can be estimated easily by $(\hat{\lambda}_{MCLS}, \hat{\gamma}_{MCLS})$.

3.2.3 Modified Maximum Empirical Likelihood Estimation

The nonparametric modified empirical likelihood (MEL) technique for the TRBH-INAR(1) model is discussed in this section, which comprises two phases. In the first step, we obtain the maximum MEL estimators for the parameters α and μ_z as follows. By taking the derivative of $Q(\alpha, \mu_z)$ defined in (7) with respect to $\boldsymbol{\beta} = (\alpha, \mu_z)$, we have the estimating equation

$$\frac{-1}{2} \frac{\partial Q(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \sum_{t=2}^{n} m_t(\boldsymbol{\beta}) = 0,$$

where $m_t(\boldsymbol{\beta}) = (m_{1,t}(\boldsymbol{\beta}), m_{2,t}(\boldsymbol{\beta}))'$ with $m_{1,t}(\boldsymbol{\beta}) = X_{t-1}(X_t - \alpha X_{t-1} - \mu_z)$, $m_{2,t}(\boldsymbol{\beta}) = X_t - \alpha X_{t-1} - \mu_z$. Following Qin and Lawless (1994), we can define the log MEL function as

$$L_{ME}(\boldsymbol{\beta}) = \sum_{t=1}^{n} \log \left(1 + d'(\boldsymbol{\beta}) m_t(\boldsymbol{\beta}) \right),$$

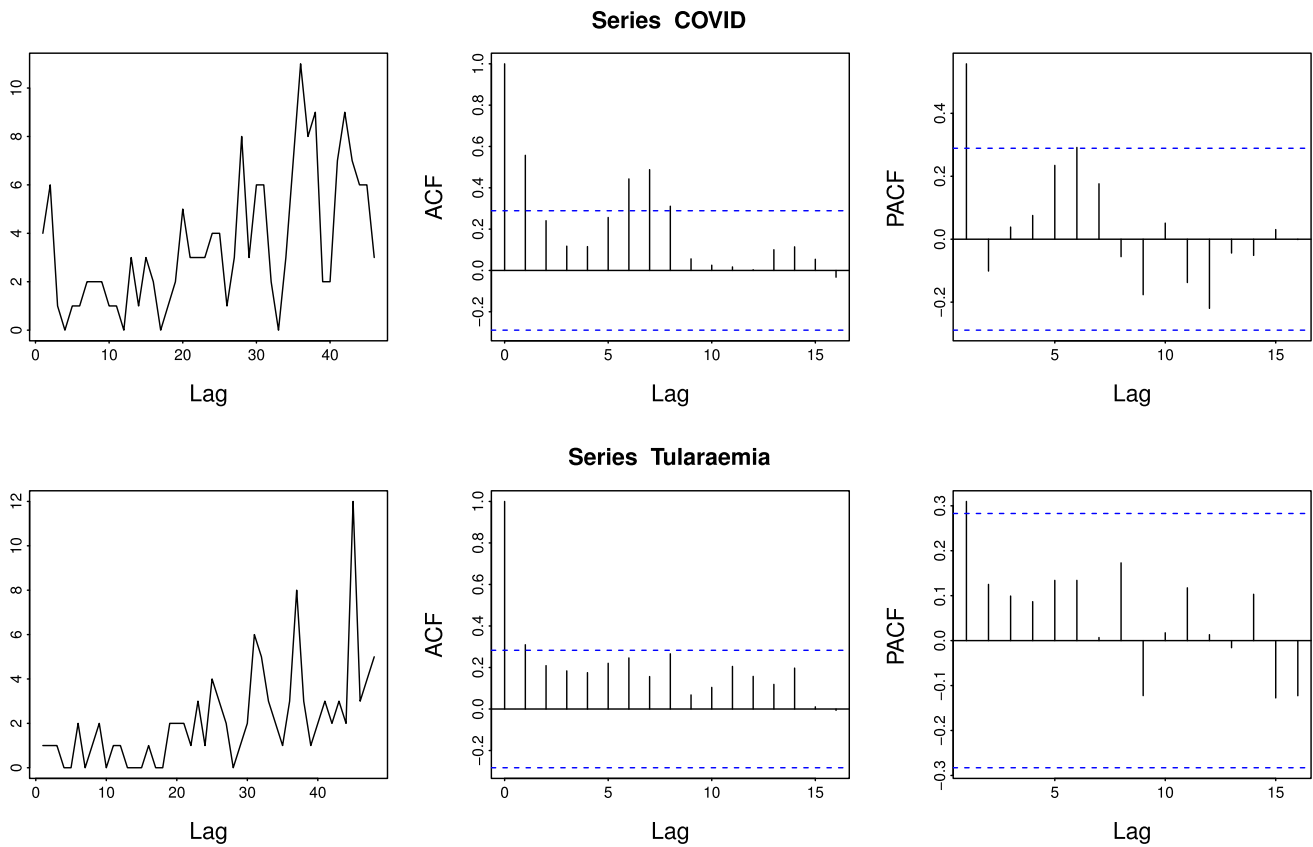where $d(\boldsymbol{\beta})$ satisfies

Fig. 4 The sample path, ACF and PACF of both data sets

Table 4 The CML estimates and some IC measures of COVID-19 data

Model	CML	AIC	BIC	HQIC	CAIC	RMS
PINAR(1)	$\hat{\lambda} = 2.13186$, $\hat{\alpha} = 0.40172$	211.27	214.92	212.64	213.84	2.34
GINAR(1)	$\hat{p} = 0.73919$, $\hat{\alpha} = 0.35818$	218.04	221.69	219.41	220.61	2.41
NGINAR(1)	$\hat{p} = 4.16264$, $\hat{\alpha} = 0.95784$	204.74	208.41	206.11	207.31	2.55
NBRCINAR(1)	$\hat{n} = 3.60698$, $\hat{p} = 0.50657$, $\hat{\rho} = 0.47929$	203.32	208.81	205.38	206.29	2.32
NBIINAR(1)	$\hat{n} = 3.07621$, $\hat{p} = 1.86301$, $\hat{\rho} = 0.53645$	201.83	207.32	203.89	204.81	2.31
GPQINAR(1)	$\hat{\lambda} = 1.43165$, $\hat{\theta} = 0.25134$, $\hat{\rho} = 0.4607$	204.38	209.86	206.43	207.35	2.48
NDCINAR(1)	$\hat{\alpha} = 0.34853$, $\hat{\theta} = 0.99957$, $\hat{\mu} = 3.70577$	213.90	219.38	215.95	214.47	2.38
ρ−NGINAR(1)	$\hat{\alpha} = 0.68822$, $\hat{\rho} = 3.73767$, $\hat{\mu} = 0.29627$	200.14	205.62	202.19	200.71	2.58
DCGINAR(1)	$\hat{\alpha} = 0.51582$, $\hat{\theta} = 0.66611$, $\hat{\mu} = 3.10469$	210.16	215.65	212.22	210.73	2.32
GADCINAR(1)	$\hat{\alpha} = 0.44552$, $\hat{\theta} = 0.80996$, $\hat{\mu} = 3.56932$	207.75	213.24	209.81	208.32	2.33
GNBINAR(1)	$\hat{\alpha} = 0.92765$, $\hat{\theta} = 0.96184$, $\hat{\mu} = 3.98312$	201.61	207.09	203.66	202.17	2.52
TRBH-INAR(1)	$\hat{\alpha} = 0.58847$, $\hat{\theta} = 0.80592$, $\hat{\lambda} = 0.66448$, $\hat{\gamma} = 0.53178$	196.94	204.26	199.68	200.44	2.27

Table 5 The CML estimates and some IC measures of Tularemia data

Model	CML	AIC	BIC	HQIC	CAIC	RMS
PINAR(1)	$\hat{\lambda} = 1.54083,\ \hat{\alpha} = 0.27962$	195.36	199.11	196.78	197.91	2.13
GINAR(1)	$\hat{p} = 0.64816,\ \hat{\alpha} = 0.34085$	184.05	187.80	185.47	186.60	2.14
NGINAR(1)	$\hat{p} = 2.15417,\ \hat{\alpha} = 0.54857$	180.33	184.07	181.75	182.88	2.18
NBRCINAR(1)	$\hat{n} = 2.42637,\ \hat{p} = 0.52029,\ \hat{\rho} = 0.38773$	183.13	188.74	185.25	186.06	2.13
NBIINAR(1)	$\hat{n} = 1.84523,\ \hat{p} = 1.60513,\ \hat{\rho} = 0.46674$	181.54	187.15	183.66	184.47	2.15
GPQINAR(1)	$\hat{\lambda} = 1.07131,\ \hat{\theta} = 0.26666,\ \hat{\rho} = 0.35176$	183.52	189.14	185.65	186.46	2.18
NDCINAR(1)	$\hat{\alpha} = 0.77226,\ \hat{\theta} = 0.99999,\ \hat{\mu} = 13.75551$	237.51	243.12	239.63	238.05	3.48
$\rho-$NGINAR(1)	$\hat{\alpha} = 1.19804,\ \hat{\rho} = 8659.928,\ \hat{\mu} = 0.31496$	177.04	182.65	179.16	177.58	4.41
DCGINAR(1)	$\hat{\alpha} = 0.62351,\ \hat{\theta} = 0.53917,\ \hat{\mu} = 2.10281$	178.49	184.11	180.61	179.04	2.23
GADCINAR(1)	$\hat{\alpha} = 0.37802,\ \hat{\theta} = 0.39667,\ \hat{\mu} = 1.93569$	185.42	191.03	187.54	185.96	2.14
GNBINAR(1)	$\hat{\alpha} = 0.34240,\ \hat{\theta} = 0.41512,\ \hat{\mu} = 2.04136$	182.49	188.11	184.61	183.04	2.13
TRBH-INAR(1)	$\hat{\alpha} = 0.33436,\ \hat{\theta} = 0.48157,\ \hat{\lambda} = 0.58878$ $\hat{\gamma} = 0.79411$	173.45	180.94	176.28	176.88	2.08
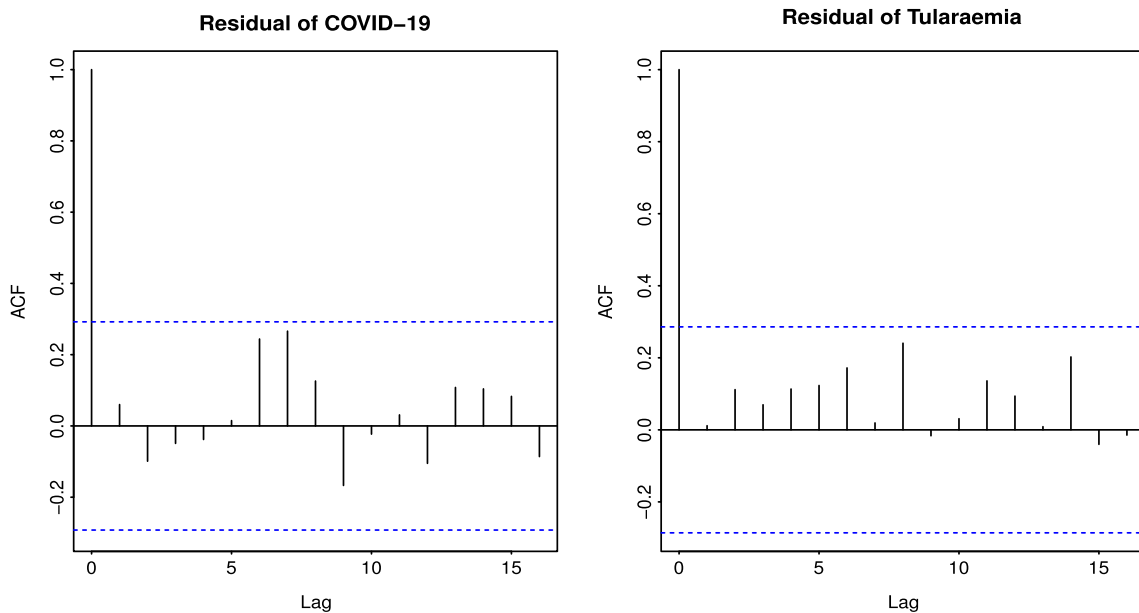
Fig. 5 The Pearson residuals ACF for the two data sets

$$\frac{1}{n}\sum_{t=1}^{n}\frac{m_t(\boldsymbol{\beta})}{1+d'(\boldsymbol{\beta})m_t(\boldsymbol{\beta})} = \boldsymbol{0}.$$

The maximum MEL estimator (MMELE) for the parameter $\boldsymbol{\beta}$ is defined by minimizing the above equation, i.e.,

$$\hat{\boldsymbol{\beta}}_{mmel} = \arg\min_{\boldsymbol{\beta}} L_{ME}(\boldsymbol{\beta}).$$

The MMEL estimation of the parameters (λ, γ) can be easily obtained based on the $\hat{\mu}_{Z,mmel}$ and finding the root of the Eq. (1).

The maximum MEL estimator for the parameter θ is obtained in the second phase. By considering the function $T(\theta)$ which is defined in (8), we have

$$\frac{-1}{2}\frac{\partial T(\theta)}{\partial \theta} = \sum_{t=2}^{n} m_t(\theta),$$
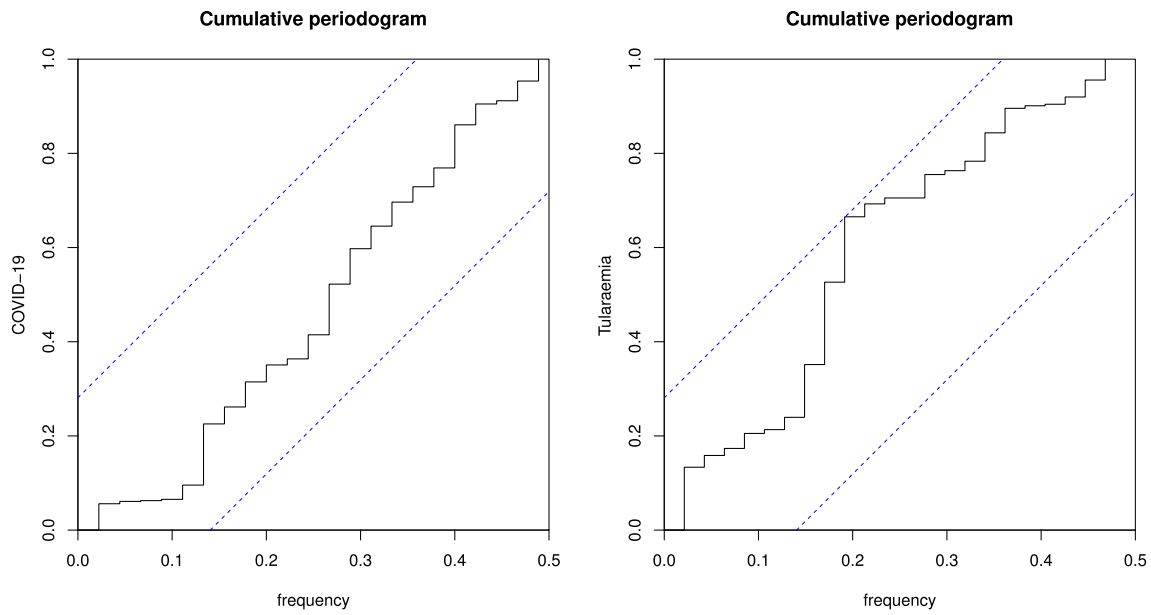
Cumulative periodogram

Fig. 6 The Pearson residuals cumulative periodogram for the two data sets
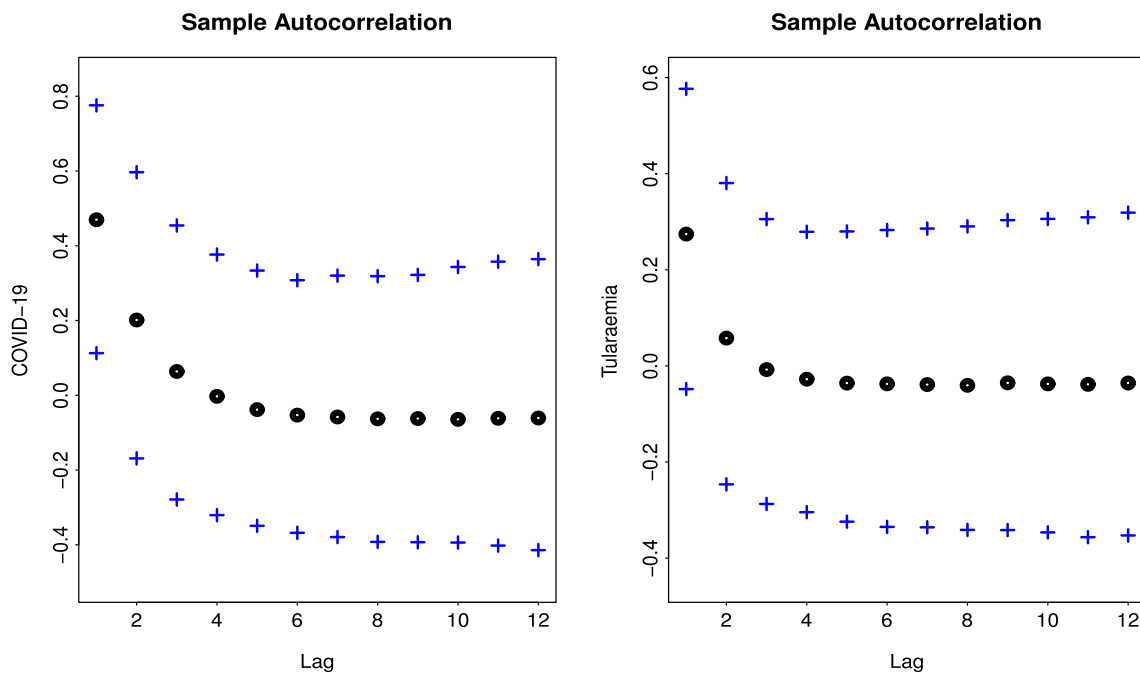
Sample Autocorrelation

Fig. 7 The acceptance areas and the bootstrap ACF

where $\quad m_t(\theta) = \hat{\alpha}_{mmel} X_{t-1}(X_{t-1}+1)\Big[V_t - \hat{\alpha}_{mmel}(\theta - \hat{\alpha}_{mmel})X_{t-1}^2 - \hat{\alpha}_{mmel}(\theta+1)X_{t-1} - \hat{\sigma}_{Z,mmel}^2\Big]$, $\quad$ and $\quad V_t = (X_t - \hat{\alpha}_{mmel}X_{t-1} - \hat{\mu}_{Z,mmel})^2$. The MMELE for the parameter θ will be obtained by minimizing log MEL function.

3.2.4 Yule–Walker Estimation

The Yule–Walker (YW) estimators of the unknown vector $\boldsymbol{\delta}$ are obtained as follows. Using the fact that $E(X_t) = \dfrac{\mu_z}{1 - \alpha}$ and $Corr(X_t, X_{t-1}) = \alpha$, the YW estimations of the parameters (α, μ_z) are generated using the sample mean and sample autocorrelation function as follows:

$$\hat{\alpha}_{YW} = \frac{\sum_{t=2}^{n}(X_t - \overline{X})(X_{t-1} - \overline{X})}{\sum_{t=1}^{n}(X_t - \overline{X})^2},$$
$$\hat{\mu}_{z,yw} = \overline{X}(1 - \hat{\alpha}_{YW}).$$

Similarly, the YW estimation of the parameters (λ, γ) is performed using the $\hat{\mu}_{Z,mmel}$, and finding the equation's root (1).

We utilize the second moment of the procedure to estimate the parameter θ as follows:

$$\begin{aligned}E(X_t^2) &= \alpha\theta E(X_t^2) + \alpha(1+\theta)E(X_t)\\ &\quad + E(Z_t^2) + 2\alpha\mu_z E(X_t)\\ &= \frac{\alpha(1+\theta)}{1-\alpha\theta}E(X_t) + \frac{E(Z_t^2)}{1-\alpha\theta} + \frac{2\alpha\mu_z E(X_t)}{1-\alpha\theta},\end{aligned}$$

which is obtained based on Remark 1. Let $\overline{X^2} = \frac{1}{n}\sum_{t=1}^{n}X_t^2$, then

$$\overline{X^2} = \frac{\hat{\alpha}_{YW}(1+\theta)}{1-\hat{\alpha}_{YW}\theta}\overline{X} + \frac{\hat{\mu}_{z,yw}^2 + \hat{\sigma}_{z,yw}}{1-\hat{\alpha}_{YW}\theta} + \frac{2\hat{\alpha}_{YW}\hat{\mu}_{z,yw}\overline{X}}{1-\hat{\alpha}_{YW}\theta}, \qquad (9)$$

as a result, estimate of the parameter θ is determined by computing the root of the Eq. (9) numerically.

4 Simulation Approach

We examine the efficiency of the parameter estimate approaches for the TRBH-INAR(1) model using Monte Carlo simulation, under different sample sizes $n = (100, 200, 500, 1000)$ over $h = 1000$ iterations. Two distinct parameter combinations are evaluated as $(\alpha, \theta, \lambda) = (0.4, 0.8, 0.7, 0.3)$ and $(0.2, 0.4, 0.9, 0.6)$. We use the mean squared error (MSE) metric to assess the estimators'

performance. The results are summarized in Tables 2 and 3, where represent that all estimates of the parameters are convergent to their actual values. Furthermore, when the sample size grows larger, the MSE decreases. Among different kinds of estimation methods, the CML and MMEL provide better performance than MCLs and YW estimations, since they have small MSE for all parameters. In comparison among the CML and MMEL methods, we provide the computer running time (R.time), which indicates that the MMEL method is faster than CML and as well as CML in MSE measure. As a result, the nonparametric MMEL technique outperforms other methods of estimation.

5 Application of Real-World Data

In this section, we investigate the application of the TRBH-INAR(1) process by using two types of clinical count data.

The first data set is devoted to daily counts of death from the COVID-19 disease, reported from Netherland and consists of 46 observations, from second July until 16-th August at 2021, by the World Health Organization (https://covid19.who.int).

The second data set represents the weekly counts of Tularemia disease, reported from Bavaria, and it consists of 48 observations, from first week until 48-th week on 2020, from the Robert Koch Institute: SurvStat@RKI 2.0 (https://survstat.rki.de) site.

Figure 4 depicts the sample path, autocorrelation function (ACF), and partial autocorrelation function (PACF) of the two data series, indicating that the data sets should be modeled using a first-order autoregressive model.

Table 6 The k-step ahead predictions of clinical data series

k	COVID-19			Tularemia		
	Actual data	Bootstrap	Classical	Actual data	Bootstrap	Classical
1	8	10	7.75761	1	1	2.28751
2	9	10	5.99219	2	2	1.61879
3	2	0	6.58066	3	1	1.95315
4	2	2	2.46137	2	1	2.28751
5	7	7	2.46137	3	7	1.95315
6	9	7	5.40372	2	1	2.28751
7	7	6	6.58066	12	13	1.95315
8	6	4	5.40372	3	2	5.29675
9	6	5	4.81525	4	4	2.28751
10	3	3	4.81525	5	3	2.62187
SMAPE		0.33131	0.43131		0.41133	0.52701

Furthermore, the augmented Dickey–Fuller test is used to justify stationarity of the two clinical data sets, where the p-value of augmented Dickey–Fuller test for COVID-19 data is less than 0.01 and for Tularemia data is equal 0.022, which confirm the stationarity of both data sets.

The mean, variance and autocorrelation of the two data sets are (3.565, 7.717, 0.557) and (2.125, 5.047, 0.310), respectively. Both data series are empirically over-dispersed with dispersion indices $\hat{I}_X = (2.164, 2.375)$, respectively.

We compare the TRBH-INAR(1) model to some competitive INAR(1) models as:

PINAR(1) (Al-Osh and Alzaid 1987), GINAR(1) (Alzaid and Al-Osh 1988), NBIINAR(1) (Al-Osh and Aly 1992), GPQINAR(1) (Alzaid and Al-Osh 1993), NBRCI-NAR(1) (Weiß 2008), NGINAR(1) (Ristić et al. 2009), DCGINAR(1) (Ristić et al. 2013), NDCINAR(1) (Miletić Ilić 2016), ρ-NGINAR(1) (Borges et al. 2017), GADCI-NAR(1) (Nastić et al. 2017) and GNBINAR(1) (Shamma et al. 2020).

We reported the CML estimates, the information criterion (IC) statistics as AIC, BIC, HQIC and CAIC, and the root mean squares of differences of observations and predicted values (RMS) for each INAR model. Tables 4 and 5 show the results for two different data series. Regarding Tables 4 and 5, the values of the IC and RMS are the smallest for the TRBH-INAR(1) model. Therefore, we can conclude that the TRBH-INAR(1) model provides the best loss information among other competitive INAR(1) models.

5.1 The Clinical Data Sets' Residual Analysis

We provide the results of a residual analysis of clinical data sets, which confirmed the suitability of the proposed model. The Pearson residuals are defined as

$$e_t = \frac{X_t - E(X_t \mid X_{t-1})}{\sqrt{Var(X_t \mid X_{t-1})}},$$

where $E(X_t \mid X_{t-1})$ and $Var(X_t \mid X_{t-1})$ are defined in (5) and (6), respectively. Note that, estimation of the parameters of the TRBH-INAR(1) model are substituted in each $E(X_t \mid X_{t-1})$ and $Var(X_t \mid X_{t-1})$ to compute the Pearson residuals.

The Pearson residuals ACF of both data sets is shown in Fig. 5. The residuals are non-correlated, as shown in Fig. 5, and the results are supported by the Ljung-Box test p-values (0.679, 0.935). Figure 6 shows the Pearson residuals cumulative periodogram, which shows how residuals are distributed randomly and without trend.

Figure 7 shows the result of the parametric re-sampling method. First, 5000 data sets with bootstrap sample size $n = (46, 48)$ are obtained using the fitted TRBH-INAR(1) model (with CML estimates of the parameters of each data set). Second, using the bootstrap samples, the ACF of each specific lag is calculated. The acceptance bounds $100(0.975)\%$ and $100(0.025)\%$ quantiles are shown as "+", and the samples ACF are presented by "•" symbols, in Fig. 7. According to Fig. 7, all of the sample autocorrelations were assigned between the acceptance boundaries, indicating that the model was adequate.

5.2 Methods of Forecasting

To test the TRBH-INAR(1) model's appropriateness and predictability, we present forecasts of the specified data sets using both the traditional and modified Sieve bootstrap approaches.

The k-step ahead classical predictor of the TRBH-INAR(1) model is represented as

$$\hat{X}_t = E(X_t \mid X_{t-k}) = \alpha^k X_{t-k} + \frac{1 - \alpha^k}{1 - \alpha} \mu_z,$$

where unknown parameters α and μ_z are substituted by the related CML estimates.

5.2.1 Modified Sieve Bootstrap Approach

The integer nature of the count data is not preserved by the classical predictor, despite the fact that the count time series is an integer. The Sieve bootstrap technique is a distribution-free predictor that preserves the integer nature of the count data. Hence, we modified the bootstrap approach proposed by Pascual et al. (2004) to apply for the TRBH-INAR(1) model via the following steps. Since $\alpha *_\theta (\alpha *_\theta X) \overset{d}{\neq} \alpha^2 *_\theta X$, we can only provide the one-step modified Sieve bootstrap prediction.

1. The thinning parameters (α, θ) are estimated based on the YW estimation approach.
2. Compute residuals $\hat{Z}_t = X_t - \hat{\alpha} X_{t-1}$, for $t = 2, ..., n$.
3. The empirical distribution of the modified residuals $\tilde{Z}_t$ is provided, where $\tilde{Z}_t = \left[\hat{Z}_t\right]$, and $[\cdot]$ shows the nearest integer value.
4. The bootstrap series X_t^b is given by

$$X_t^b = \hat{\alpha} *_{\hat{\theta}} X_{t-1}^b + Z_t^b, \qquad b = 1, ..., B,$$

   where B is the bootstrap sample size that was chosen to be $B = 500$, and Z_t^b is generated from the empirical distribution in step 3, for $t = 1, 2, ..., n$.
5. The YW estimation of the parameters $(\hat{\alpha}_{YW}, \hat{\theta}_{YW})$ is obtained by inserting the sample mean, variance, and solving the following equations

$$E(X_t)(1-\alpha) = E(Z_t)$$

$$Var(X_t) = \frac{\alpha(\theta-\alpha)\mu_z^2}{(1-\alpha)^2(1-\alpha\theta)}$$
$$+ \frac{\alpha(\theta+1)\mu_z}{(1-\alpha)(1-\alpha\theta)} + \frac{\sigma_Z^2}{1-\alpha\theta}.$$

6. Based on the sample means $\hat{\alpha} = \frac{1}{B}\sum_{i=1}^{B}\hat{\alpha}_{i,yw}$ and $\hat{\theta} = \frac{1}{B}\sum_{i=1}^{B}\hat{\theta}_{i,yw}$, the parameters (α, θ) are estimated.

7. The recursion method is used to acquire future bootstrap observations by the expression

$$\hat{X}_{t+1}^b = \hat{\alpha} *_{\hat{\theta}} X_t^b + Z_{t+1}^b.$$

The traditional and modified Sieve bootstrap predictions of the relevant data series, for which we know the observed values, are provided in Table 6 as a result of evaluating two prediction approaches. When there are zero or near-zero data demands, the symmetric mean absolute percent error (SMAPE) is applied to compare the forecast systems. The less SMAPE value leads to a better forecasting scheme. According to Table 6, the modified Sieve bootstrap predictors' SMAPE values are lower than classical, and the modified Sieve bootstrap predictors are integers, which are consistent with the nature of actual data.

6 Conclusions

We provide a first-order integer-valued autoregressive [INAR(1)] time series model based on the transmuted record type-discrete Burr–Hatke (TRT-DBH) distribution, which is a more flexible version of the discrete Burr–Hatke distribution. The TRT-DBH distribution is proved to over-dispersed, asymmetric and leptokurtic. The hazard rate function of the proposed distribution has different shapes as monotone and unimodal. The applicability of the TRT-DBH distribution is demonstrated in time series modeling based on an INAR(1) model with TRT-DBH distributed innovations. Properties of the model are studied as well as different estimation approaches for the model parameters. The assessment of the properties and estimation approaches is conducted via some simulation studies. The adequacy of fit of the proposed INAR(1) model is checked via two clinical data sets, including the Covid-19 series and is compared with other competitive models. For both clinical data sets, we perform the residual analysis (Pearson residuals), as well as traditional and modified Sieve bootstrap forecasting methods.

Appendix A: The Proof of Proposition 2

(i) The model's conditional expectation is calculated as follows:

$$E(X_t \mid X_{t-1}) = \alpha X_{t-1} + \mu_z,$$
$$E(X_t \mid X_{t-2}) = E(E(X_t \mid X_{t-1}) \mid X_{t-2})$$
$$= \alpha^2 X_{t-2} + (1+\alpha)\mu_z.$$

So, it is induced via the induction that

$$E(X_t \mid X_{t-k}) = \alpha^k X_{t-k} + \frac{(1-\alpha^k)\mu_z}{1-\alpha},$$

which is a linear function of X_t.

(ii) Based on Remark 1, the conditional variance of the TRBH-INAR(1) process is computed as

$$Var(X_t \mid X_{t-1}) = Var(\alpha *_\theta X_{t-1} \mid X_{t-1}) + Var(Z_t)$$
$$= \alpha(\theta-\alpha)X_{t-1}^2 + \alpha(\theta+1)X_{t-1} + \sigma_Z^2,$$

and

$$Var(X_t \mid X_{t-2}) = E(Var(X_t \mid X_{t-1}) \mid X_{t-2})$$
$$+ Var(E(X_t \mid X_{t-1}) \mid X_{t-2})$$
$$= \alpha(\theta-\alpha)E(X_{t-1}^2 \mid X_{t-2})$$
$$+ \alpha(\theta+1)E(X_{t-1} \mid X_{t-2})$$
$$+ \sigma_Z^2 + \alpha^2 Var(X_{t-1} \mid X_{t-2})$$
$$= \alpha\theta Var(X_{t-1} \mid X_{t-2})$$
$$+ \alpha(\theta-\alpha)E^2(X_{t-1} \mid X_{t-2})$$
$$+ \alpha(\theta+1)E(X_{t-1} \mid X_{t-2}) + \sigma_Z^2$$
$$= \alpha^2(\theta^2-\alpha^2)X_{t-2}^2 + \alpha^2(\theta+1)^2 X_{t-2}$$
$$+ 2\alpha^2(\theta-\alpha)\mu_z X_{t-2}$$
$$+ \alpha(\theta+1)\mu_z$$
$$+ \alpha(\theta-\alpha)\mu_z^2 + (1+\alpha\theta)\sigma_Z^2,$$

subsequently

$$Var(X_t \mid X_{t-3}) = \alpha^3(\theta^3-\alpha^3)X_{t-3}^2$$
$$+ \alpha^3(\theta+1)(1+\theta+\theta^2)X_{t-3}$$
$$+ 2\alpha^3\left((\theta-\alpha)+(\theta^2-\alpha^2)\right)\mu_z X_{t-3}$$
$$+ \alpha(\theta+1)\left(1+\alpha(1+\theta)\right)\mu_z$$
$$+ \left[\alpha(\theta-\alpha)+\alpha^2(\theta^2-\alpha^2)\right]\mu_z^2$$
$$+ 2\alpha^2(\theta-\alpha)\mu_z^2 + (1+\alpha\theta+\alpha^2\theta^2)\sigma_Z^2.$$

By induction, we can conclude that

$$Var(X_t \mid X_{t-k}) = \alpha^k \left(\theta^k - \alpha^k \right) X_{t-k}^2$$
$$+ \alpha^k (1 + \theta) \sum_{i=0}^{k-1} \theta^i X_{t-k}$$
$$+ 2\mu_z \alpha^k \sum_{i=0}^{k-1} \left(\theta^i - \alpha^i \right) X_{t-k}$$
$$+ \mu_z (1 + \theta) \sum_{i=1}^{k-1} \alpha^i \sum_{j=0}^{i-1} \theta^j$$
$$+ \mu_z^2 \sum_{i=0}^{k-1} \alpha^i \left(\theta^i - \alpha^i \right)$$
$$+ 2\alpha\mu_Z^2 \sum_{i=1}^{k-1} \alpha^i \sum_{j=1}^{i} \left(\theta^j - \alpha^j \right)$$
$$+ \sigma_Z^2 \sum_{i=0}^{k-1} (\alpha\theta)^i.$$

After some elementary calculations, the proof is complete.

(iii)   The proof is unimportant and may be ignored.

Declarations

Conflict of Interest The authors have no relevant financial or non-financial interests to disclose.

References

Al-Osh MA, Aly EEAA (1992) First order autoregressive time series with negative binomial and geometric marginals. Commun Stat-Theory Methods 21:2483–2492

Al-Osh MA, Alzaid AA (1987) First-order integer-valued autoregressive (INAR(1)) process. J Time Ser Anal 8:261–275

Altun E (2020) A new generalization of geometric distribution with properties and applications. Commun Stat-Simul Comput 49(3):793–807

Altun E, Mamode Khan N (2021) Modelling with the Novel INAR(1)-PTE Process. Methodol Comput Appl Probab. https://doi.org/10.1007/s11009-021-09878-2

Alzaid AA, Al-Osh MA (1988) First-order integer-valued autoregressive (INAR(1)) process: distributional and regression properties. Stat Neerl 42:53–61

Alzaid AA, Al-Osh MA (1993) Some autoregressive moving average processes with generalized Poisson marginal distributions. Ann Inst Stat Math 45:223–232

Bahti D, Bakouch HS (2019) A new infinitely divisible discrete distribution with applications to count data modeling. Commun Stat-Theory Methods 48:1401–1416

Bakouch HS, Jazi AM, Nadarajah S (2014) A new discrete distribution. Statistics 48:200–240

Borges P, Bourguignon M, Molinares FF (2017) A generalized NGINAR(1) process with inflated parameter geometric counting series. Aust N Z J Stat 59:137–150

Bourguignon M, Vasconcellos KLP (2015) First order non-negative integer valued autoregressive processes with power series innovations. Braz J Probab Stat 29(1):71–93

Chakraborty S, Chakravarty D (2012) Discrete gamma distribution: properties and parameter estimation. Commun Stat-Theory Methods 41:3301–3324

Eliwa MS, El-Morshedy M (2021) A one-parameter discrete distribution for over-dispersed data: statistical and reliability properties with applications. J Appl Stat. https://doi.org/10.1080/02664763.2021.1905787

El-Morshedy M, Eliwa MS, Altun E (2020) Discrete Burr–Hatke distribution with properties, estimation methods and regression model. IEEE Access 8:74359–74370

Gómez-Déniz E, Calderín-Ojeda E (2011) The discrete Lindley distribution: properties and applications. J Stat Comput Simul 81(11):1405–1416

Huang J, Zhu F (2021) A new first-order integer-valued autoregressive model with Bell innovations. Entropy 323:713

Hussain T, Aslam M, Ahmad M (2016) A two parameter discrete Lindley distribution. Rev Colomb Estadíst 39(1):45–61

Jazi MA, Jones G, Lai CD (2012) Integer valued AR(1) with geometric innovations. J Iran Stat Soc 11:173–190

Karlis D, Xekalaki E (2005) Mixed Poisson distributions. Int Stat Rev 73(1):35–58

Karlsen H, Tjøstheim D (1988) Consistent estimates for the NEAR(2) and NLAR(2) time series models. J R Stat Soc. Ser B (Methodol), **50**(2), 313–320

Kozubowski TJ, Podgórski K (2016) Transmuted distributions and random extrema. Stat Probab Lett 116:6–8

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

Lívio T, Mamode Khan N, Bourguignon M, Bakouch SH (2018) An INAR(1) model with Poisson–Lindley innovations. Econ Bull 38(3):1505–1513

McKenzie E (1986) Autoregressive moving-average processes with negative binomial and geometric distributions. Adv Appl Probab 18:679–705

McKenzie E (1988) Some ARMA models for dependent sequences of Poisson counts. Adv Appl Probab 20:822–835

Miletić Ilić AV (2016) A geometric time series model with a new dependent Bernoulli counting series. Commun Stat-Theory Methods 45:6400–6415

Nastić AS, Ristić MM, Miletić Ilić AV (2017) A geometric time series model with an alternative dependent Bernoulli counting series. Commun Stat-Theory Methods 46(2):770–785

Nekoukhou V, Alamatsaz MH, Bidram H (2013) Discrete generalized exponential distribution of a second type. Statistics 47(4):876–887

Pascual L, Romo J, Ruiz E (2004) Bootstrap predictive inference for ARIMA processes. J Time Ser Anal 25(4):449–65

Qin J, Lawless J (1994) Empirical likelihood and general estimating equations. Ann Stat 22(1):300–325

Ristić MM, Bakouch HS, Nastić AS (2009) A new geometric first-order integer-valued autoregressive (NGINAR(1)) process. J Stat Plan Inference 139:2218–2226

Ristić MM, Nastić SA, Miletić Ilić VA (2013) A geometric time series model with dependent Bernoulli counting series. J Time Ser Anal 34(4):423–516

Sellers KF, Shmueli G (2010) A flexible regression model for count data. Ann Appl Stat 4(2):943–961

Sellers KF, Borle S, Shmueli G (2012) The COM-Poisson model for count data: a survey of methods and applications. Stat Qual, Ind Technol 28(2):104–116

Shakil M, Ahsanullah M (2011) Record values of the ratio of Rayleigh random variables. Pak J Stat 27:307–325

Shamma N, Mohammadpour M, Shirozhan M (2020) A time series model based on dependent zero inflated counting series. Comput Stat 35:1737–1757

Shirozhan M, Mohammadpour M, Bakouch HS (2019) A new geometric INAR(1) model with mixing Pegram and generalized binomial thinning operators. Iran J Sci Technol, Trans A: Sci 43:1011–1020

Shmueli G, Minka TP, Kadane JB, Borle S, Boatwright P (2005) A useful distribution for fitting discrete data: revival of the Conway–Maxwell–Poisson distribution. J Appl Stat 54:127–142

Steutel FW, van Harn K (1979) Discrete analogues of self-decomposability and stability. Ann Probab 7:893–899

Weiß CH (2008) Thinning operations for modeling time series of count—a survey. AStA Adv Stat Anal 92:319–341