ORIGINAL PAPER

Analysis of zero-and-one inflated bounded count time series with applications to climate and crime data

Yao Kang[1] · Shuhui Wang[2] · Dehui Wang[3] · Fukang Zhu[2]

Abstract

This article introduces a new version of first-order binomial autoregressive (BAR(1)) process with zero-and-one inflated binomial marginals using the idea of hidden Markov models, which contains the BAR(1) and other existing processes as special cases. Stochastic properties of the new model are investigated and model parameters are estimated by the probability-based, quasi-maximum likelihood, maximum likelihood and Bayesian methods. A binomial one-inflation index is constructed and further utilized to develop a method to test whether zero and/or one inflation with respect to a BAR(1) model. We also give the asymptotic distribution of the corresponding test statistics under the null hypothesis. Applications to rainy-days and assaults-on-officers counts are conducted, which shows that the proposed model can accurately capture zero-inflation, one-inflation and overdispersion characteristics of the data. The predictive distributions are employed to identify the occurrence of anomalies and then establish early warning system of risk.

Keywords BAR(1) process · Bayesian inference · Bounded count time series · One-inflation · Test · Zero-inflation

Mathematics Subject Classification 62M10 · 62F10 · 62F12

1 Introduction

During the last decade, modeling and analysis of count time series with a bounded support have become a popular topic with a large quantity of articles in fields like epidemiology, social sciences, economics, life sciences and others. The most commonly

✉ Fukang Zhu
   zfk8010@163.com

1   School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

2   School of Mathematics, Jilin University, Changchun, China

3   School of Mathematics and Statistics, Liaoning University, Shenyang, China

used approach to model this kind of data is using the first-order binomial autoregressive (BAR(1)) process proposed by McKenzie (1985). The BAR(1) model was constructed by using the binomial thinning operator, which was proposed by Steutel and Van Harn (1979) and defined as $\alpha \circ X = \sum_{i=1}^{X} B_i$, where $\{B_i\}$ is an independent and identically distributed (i.i.d.) Bernoulli(α) random sequence independent of X. Based on this operator, the definition of the BAR(1) model is given below.

Definition 1 The BAR(1) process $\{X_t\}$ is defined by the recursion

$$X_t = \alpha \circ X_{t-1} + \beta \circ (n - X_{t-1}) \text{ with } X_0 \sim B(n, p), \tag{1.1}$$

where $n \in \mathbb{N}$, $\beta := p(1 - \rho)$, $\alpha := \beta + \rho$, $p \in (0, 1)$ and $\rho \in \left(\max\left\{ -\dfrac{p}{1-p}, -\dfrac{1-p}{p} \right\}, 1 \right)$. All thinnings are performed independently of each other, and the thinnings at time t are independent of $(X_s)_{s<t}$.

The BAR(1) model is a homogeneous Markov chain with h-step transition probabilities

$$
\begin{aligned}
&P(X_t = j | X_{t-h} = i) \\
&= \sum_{m=\max\{0,i+j-n\}}^{\min\{i,j\}} \binom{i}{m}\binom{n-i}{j-m} \alpha_h^m (1-\alpha_h)^{i-m} \beta_h^{j-m} (1-\beta_h)^{n-i-j+m},
\end{aligned}
\tag{1.2}
$$

where $\beta_h = p(1 - \rho^h)$ and $\alpha_h = \beta_h + \rho^h$.

During the past ten years, the interest in the BAR(1) process has significantly increased and research on this model has gained plentiful and substantial harvest. For example, Scotto et al. (2014) and Ristić and Popović (2019) investigated different types of bivariate BAR(1) processes. Weiß and Pollett (2014) introduced a class of density dependent BAR(1) process, where the thinning probabilities depend on the current observations. Kim and Weiß (2015) considered modeling of bounded count time series with overdispersion via proposing a random coefficient BAR(1) process. Möller et al. (2016) proposed types of self-exciting threshold BAR(1) models. Kang et al. (2020; 2021) introduced an extended version of BAR(1) model and a mixture BAR(1) model based on the generalized binomial thinning operator and mixing Pegram and binomial thinning operators, respectively. For some other meaningful contributions on the BAR(1) process, we refer to Cui and Lund (2010), Weiß (2009a, b), (2013), Weiß and Pollett (2012), Weiß and Kim (2013a, b, 2015), Kim and Weiß (2015), Yang et al. (2017) and Chen et al. (2021), among others.

In the context of count data with an unbounded support, the concept of zero inflation (excess zeros) is that the proportion of 0's of a model is greater than the proportion of 0's of the corresponding Poisson model. Zero-inflation phenomenon is frequently encountered in a great number of fields, such as econometrics, manufacturing defects, medical consultations, sexual behavior and so on (see Ridout et al. 1998). Research on modeling zero inflation is vitally necessary and important. As pointed out by Zuur et al. (2009, Chapter 11), ignoring zero inflation can have at least two consequences: first, the

estimated parameters and standard errors may be biased; second, the excessive number of zeros can cause overdispersion. Moreover, Perumean-Chaney et al. (2013) pointed out that when zero inflation in the data was ignored, the estimation results were poor and some potentially significant statistical findings were missing, the misspecifications caused by the ignorance of zero inflation may even lead to erroneous conclusions about the data and bring uncertainty to research and applications. In the context of count time series, modeling of zero inflation has also attracted the attention of researchers. On one hand, Zhu (2012) proposed zero-inflated Poisson and negative binomial integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) models. Jazi et al. (2012) introduced a first-order integer-valued autoregressive (INAR(1)) process based on binomial thinning operator with zero-inflated Poisson innovations. Li et al. (2015) proposed a mixed INAR(1) process with zero-inflated generalized power series innovations. These articles give a solution to modeling of zero-inflated count time series with an unbounded support. On the other hand, there are only a few articles about modeling and testing bounded count time series containing a large number of zeros. Möller et al. (2018) developed four extensions of the BAR(1) model, which can accommodate a broad variety of zero patterns. Kim et al. (2018) proposed statistics to evaluate the number of zeros and the dispersion with respect to a binomial model. Besides, Möller et al. (2020) introduced a state-dependent zero-inflation mechanism for count distributions with an unbounded or bounded support.

However, a count time series data set that contains a great number of zeros along with a large number of ones can also arise in practice. For example, most car owners will claim zero or one time in an insurance period since more than one claims can lead to higher premiums in the next insurance period; one may be infected by seasonal flu for at most one time in a quarter due to the existence of antibodies over a period of time. Research on modeling of zero-and-one inflated count time series is still extremely infrequent and there are only a few articles to concentrate on the relevant issue till now. Maiti et al. (2018) constructed a new mixture INAR(1) process for modeling count time series data, in particular data consisting of many zeros and ones. Qi et al. (2019) and Mohammadi et al. (2022) presented binomial thinning INAR(1) processes with zero-and-one inflated Poisson and zero-and-one inflated Poisson–Lindley innovations, respectively. The above three articles considered modeling of zero-and-one inflated unbounded count time series.

While it is also a vitally necessary and significant issue to come up with a solution to modeling of bounded count time series with a large number of zeros and ones, to our best knowledge, there is only one article that studies the relevant issue till now. Liu et al. (2022) proposed a zero-one-inflated bounded Poisson model with an autoregressive feedback mechanism in intensity to fit the normalcy-dominant ordinal time series. This article concentrates on the establishment of statistical model from a new perspective, so as to achieve the purpose of enriching the method for handling zero-and-one inflated bounded count time series. For this, the concept of hidden Markov model is utilized to help us construct a new version of extended BAR(1) model with zero-and-one inflated binomial marginal distribution. Furthermore, a binomial one-inflation index is presented and further employed to construct test statistics to evaluate the number of ones with respect to a binomial model. To illustrate the significance, it is necessary

to explain the contribution of this article, especially when our work is compared with Liu et al. (2022). This issue can be discussed based on the following perspectives:

(i) The model proposed by Liu et al. (2022) is constructed based on the zero-one-inflated bounded Poisson distribution and INGARCH framework. But it is well known that the binomial distribution and BAR(1) process are almost the most commonly used models to deal with bounded count data. Hence, constructing the extended BAR(1) model to fit the zero-and-one inflated bounded count time series is a natural issue.

(ii) Liu et al. (2022) proposed the zero-one-inflated bounded Poisson INGARCH model to better analyze normalcy-dominant ordinal data, rather than the zero-and-one inflated bounded count time series. As we mentioned before, this article concentrates on modeling of the zero-and-one inflated bounded count time series via a different modeling framework.

(iii) Another motivation is that the proposed model is flexible and it is a novel extension of the binomial modeling framework, and the BAR(1) and RZ-BAR(1) (see Möller et al. 2018) models are special cases of our model. Moreover, the proposed model not only has an attractive potential for modeling and analyzing bounded count time series with excess zeros and ones, but also can be practical in other situations when the observed process indicates some characteristics such as overdispersion and bimodal.

The rest contents of this article are organized as follows. In Sect. 2, the new model is introduced and some probabilistic and statistical properties are investigated. In Sect. 3, probability-based, quasi-maximum likelihood, maximum likelihood and Bayesian methods are employed to estimate the model parameters. Section 4 studies the binomial one-inflation index and the relevant test problem is addressed. In Sect. 5, we apply the proposed model to two rainy-days counts and an assaults-on-officers counts. The article ends with a conclusion section and has online supplementary materials.

2 A BAR-hidden Markov model

To model zero-and-one inflated count time series with a bounded support, a direct approach is to construct a modified and generalized version of the BAR(1) process. Suppose that $\{X_t\}$ is the underlying but unobservable BAR(1) process, it is natural to generate additional zeros and ones via randomly replacing some of the X_t by zeros or ones. Formally, at each time t, a uniform random variable S_t is realized, independently of the previous $\{X_m\}_{m<t}$ and $\{S_m\}_{m<t}$. We define the observable process $\{Y_t\}$, satisfying $Y_t := 0$ if $0 \le S_t < \phi_0$, $Y_t := 1$ if $\phi_0 \le S_t < \phi_0 + \phi_1$ and $Y_t := X_t$ if $\phi_2 < S_t \le 1$, with $0 \le \phi_0, \phi_1, \phi_2 \le 1$, $\phi_0 + \phi_1 + \phi_2 = 1$. It can be seen that $\{X_t, Y_t\}$ constitutes a hidden Markov model (see Zucchini et al. 2009) with BAR(1) kernel $\{X_t\}$ being the Markov chain of hidden states. From another point of view, some of the original counts $\{X_t\}$ are masked by a zero or a one, but the underlying BAR(1) process is not affected by them. The masking by zeros and ones is accomplished totally at random by the i.i.d. uniform random sequence $\{S_t\}$. The above approach for generating zeros and ones is based on the idea of generating missing values $\{S_t\}$

completely at random. Hence, the BAR-hidden Markov model can also be viewed as the BAR(1) process with zeros and ones at random. Now, we are ready to give the definition of our model based on the above analysis.

Definition 2 The BAR-hidden Markov model $\{Y_t\}_{t \geq 1}$ can be written as follows:

$$Y_t = \begin{cases} 0, & \text{with probability } \phi_0, \\ 1, & \text{with probability } \phi_1, \\ X_t, & \text{with probability } \phi_2, \end{cases} \tag{2.1}$$

where $0 \leq \phi_0, \phi_1, \phi_2 \leq 1$, $\phi_0 + \phi_1 + \phi_2 = 1$ and $\{X_t\}$ is the BAR(1) model defined in Eq. (1.1). We call the new process as the first-order zero-and-one inflated binomial autoregressive (ZOIBAR(1)) process.

Remark 1 The ZOIBAR(1) model given in Eq. (2.1) can be rewritten as

$$Y_t = 0 \cdot \mathbb{I}_{\{0 \leq S_t < \phi_0\}} + 1 \cdot \mathbb{I}_{\{\phi_0 \leq S_t < \phi_0 + \phi_1\}} + X_t \cdot \mathbb{I}_{\{1 - \phi_2 \leq S_t \leq 1\}}, \ t = 1, 2, \ldots,$$

where $\{S_t\}$ is an i.i.d. uniform(0,1) random sequence and $\{X_t\}$ is the BAR(1) model.

Remark 2 Some special cases of the ZOIBAR(1) process defined in Eq. (2.1) are:

(i) the classical BAR(1) process proposed by McKenzie (1985) if $\phi_0 = \phi_1 = 0$;
(ii) the BAR(1) process with zeros at random (RZ-BAR(1) process) proposed by Möller et al. (2018) if $\phi_1 = 0$;
(iii) the i.i.d. Bernoulli random sequence with $P(Y_t = 0) = 1 - P(Y_t = 1) = 1 - \phi_1 = \phi_0$ if $\phi_2 = 0$.

From Definition 2, the BAR(1) kernel $\{X_t\}$ obviously is a Markov chain. However, the same conclusion cannot be generalized to the process $\{Y_t\}$ since some of the Y_t are randomly masked by zeros and ones. Given the tth kernel count X_t, the corresponding observation Y_t is not only generated by X_t, but it also depends on the generating of S_t. We give an example to illustrate this statement in supplementary materials. Furthermore, the distribution of Y_t given X_t can be given as follows:

$$P(Y_t = y_t | X_t = x_t) = \phi_0 \mathbb{I}_{\{y_t = 0\}} + \phi_1 \mathbb{I}_{\{y_t = 1\}} + \phi_2 \mathbb{I}_{\{y_t = x_t\}}.$$

Now, we turn to concentrate on the statistical properties of the ZOIBAR(1) model. We first give the probability mass function (p.m.f.) of the marginal distribution of the ZOIBAR(1) model as follows:

$$P(Y_t = k) = \begin{cases} \phi_0 + \phi_2(1-p)^n, & k = 0, \\ \phi_1 + \phi_2 np(1-p)^{n-1}, & k = 1, \\ \phi_2 \binom{n}{k} p^k (1-p)^{n-k}, & k = 2, 3, \ldots, n. \end{cases}$$

It can be easily seen that the marginal distribution of the ZOIBAR(1) model is a zero-and-one inflated binomial distribution. From the marginal distribution of the ZOIBAR(1) model, the mean and variance can be easily obtained by

$$E(Y_t) = \phi_1 + \phi_2 np, \ \text{Var}(Y_t) = (\phi_1 + \phi_2 np)[1 - (\phi_1 + \phi_2 np)] + n(n-1)\phi_2 p^2.$$

The following proposition gives the autocovariance function and autocorrelation function at lag h for the ZOIBAR(1) process. The corresponding proof is provided by Supplement S5.

Proposition 1 *The autocovariance function and autocorrelation function at lag h for the ZOIBAR(1) model are given by*

$$\text{Cov}(Y_t, Y_{t+h}) = \begin{cases} (\phi_1 + \phi_2 np)[1 - (\phi_1 + \phi_2 np)] + n(n-1)\phi_2 p^2, & h = 0, \\ \rho^h \phi_2^2 np(1-p), & h = 1, 2, \ldots, \end{cases}$$

and

$$\rho(h) := \text{Corr}(Y_t, Y_{t+h})$$
$$= \begin{cases} 1, & h = 0, \\ \rho^h \dfrac{\phi_2^2 np(1-p)}{(\phi_1 + \phi_2 np)[1 - (\phi_1 + \phi_2 np)] + n(n-1)\phi_2 p^2}, & h = 1, 2, \ldots. \end{cases}$$

This proposition implies that the autocorrelation function of the process decays exponentially to 0 as $h \to \infty$. This characteristic is similar to that of the BAR(1) process. Furthermore, we have shown that the marginal mean of $\{Y_t\}$ and the autocovariance function between Y_{t+h} and Y_t do not rely on the time index t. So it can be concluded that the ZOIBAR(1) model is covariance (weakly) stationary. Furthermore, we can give the partial autocorrelation function based on the autocorrelation function in Proposition 1.

Remark 3 Denote $\rho_{\text{part}}(h) := a_{hh}$ as the hth-order partial autocorrelation for the ZOIBAR(1) model. According to Box et al. (1994), the a_{hj} follow the recursive scheme

$$a_{h+1,h+1} = \frac{\rho(h+1) - \sum_{i=1}^{h} a_{hi} \cdot \rho(h+1-i)}{1 - \sum_{i=1}^{h} a_{hi} \cdot \rho(i)}, \ h \in \mathbb{N}_0$$
$$a_{h+1,j} = a_{hj} - a_{h+1,h+1} \cdot a_{h,h-j+1}, \ j = 1, \ldots, h,$$

where $\rho(\cdot)$ is given in Proposition 1. Based on the above recursive scheme, the first-order and second-order partial autocorrelations for the ZOIBAR(1) model are

$$\rho_{\text{part}}(1) = \rho(1), \ \rho_{\text{part}}(2) = \frac{\rho(2) - \rho^2(1)}{1 - \rho^2(1)}.$$

The high-order partial autocorrelation for the ZOIBAR(1) model can also be given in the similar way.

Now, we are ready to introduce the binomial index of dispersion, I_d, which is a useful metric to quantify the dispersion behavior of a count random variable X with a finite range $\{0, 1, \ldots, n\}$. The binomial index of dispersion I_d is defined as

$$I_d = \frac{n\sigma^2}{\mu(n-\mu)} = 1 + \frac{n(\sigma^2 - \mu) + \mu^2}{\mu(n-\mu)} \in (0, \infty),$$

where μ and σ^2 are the mean and variance of the random variable X, respectively. A finite range count random variable is said to have overdispersion if $I_d > 1$, it is equidispersed if $I_d = 1$, and it is underdispersed if $I_d < 1$. From the mean and variance formulae, we can calculate the binomial dispersion index of the ZOIBAR(1) process as follows:

$$I_d = 1 + \frac{(n-1)[\phi_2 n^2 p^2 - (\phi_1 + \phi_2 np)^2]}{(\phi_1 + \phi_2 np)[n - (\phi_1 + \phi_2 np)]}.$$

The ability of the ZOIBAR(1) model to explain zero inflation is an important topic in our work. It is worth emphasizing that relying entirely on the 0's proportion, we may determine the extent of zero inflation in a misleading way, especially if we ignore the mean of data. To illustrate this statement, we consider a count time series data set, which represents monthly counts of car beats in Pittsburgh (among $n = 42$ such car beats) that had at least one police offense report of prostitution in that month. The data consist of $T = 96$ observations, starting from January 1990 and ending in December 1997. Figure S1 and Table S1, respectively, show the sample path, histogram and descriptive statistics for the prostitution series. A zero frequency of 63.5% may lead practitioners to firmly believe that the empirical degree of zero inflation is quite moderate. However, we can also observe the mean of the series is 0.5313, which is relatively small. By a simple calculation, a binomial random variable X with mean $\bar{X} = 0.5313$ and $n = 42$ gives a zero probability of 0.5846, which is a little smaller than 63.5%. Hence, declaring the degree of zero inflation of this series is "quite moderate" may be unconvincing. To quantify and assess the zero-inflation behavior more precisely, Kim et al. (2018) introduced binomial zero-inflation index and sample binomial zero-inflation index for a count random variable X with a finite range $\{0, 1, \ldots, n\}$ as follows:

$$z_0 = p_0 \left(1 - \frac{\mu}{n}\right)^{-n} \in (0, \infty), \quad \hat{z}_0 = \hat{p}_0 \left(1 - \frac{\bar{X}}{n}\right)^{-n},$$

where p_0 and μ are the zero probability and mean of the random variable X, $\hat{p}_0 = \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}_{\{X_t=0\}}$ and $\bar{X} = \frac{1}{T} \sum_{t=1}^{T} X_t$. If X is binomially distributed, then $z_0 = 1$ and $\hat{z}_0$ should be close to 1. If $z_0 > 1$, naturally, X is said to show zero inflation with respect to the binomial distribution. Table S1 shows that the sample zero-inflation index for the prostitution series is $\hat{z}_0 = 1.0845$, which is close to 1. So we cannot conclude that the concerned series is zero-inflated based on this aspect. As a summary, zero-inflation index is the more reasonable assessment criterion rather than 0's proportion.

The binomial zero-inflation index gives us a powerful tool to measure the zero pattern departure from the binomial model. Hence, we can calculate the binomial zero-inflation index of the ZOIBAR(1) process as follows:

$$z_0 = [\phi_0 + \phi_2 (1-p)^n]\left(1 - \frac{\phi_1 + \phi_2 np}{n}\right)^{-n}.$$

From the above equation, we find that z_0 increases as ϕ_0 increases when ϕ_1, ϕ_2, n and p are fixed.

The following proposition gives the h-step transition probabilities for the ZOIBAR(1) process and the corresponding proof is provided by Supplement S5.

Proposition 2 *Let $P_{ij}^{(h)} := P(Y_t = j | Y_{t-h} = i)$, then the h-step transition probabilities for the ZOIBAR(1) model are given by*

$$
P_{ij}^{(h)} = \begin{cases}
\phi_2 P(X_t = j | X_{t-h} = i), & i, j \geq 2, \\
\phi_i \dfrac{P(Y_t = j)}{P(Y_{t-h} = i)} + \phi_2 \dfrac{P(X_{t-h} = i)[\phi_j + \phi_2 P(X_t = j | X_{t-h} = i)]}{P(Y_{t-h} = i)}, & 0 \leq i, j \leq 1, \\
\dfrac{\phi_i P(Y_t = j) + \phi_2^2 P(X_{t-h} = i) P(X_t = j | X_{t-h} = i)}{P(Y_{t-h} = i)}, & i \leq 1, j \geq 2, \\
\phi_j + \phi_2 P(X_t = j | X_{t-h} = i), & i \geq 2, j \leq 1,
\end{cases}
$$

where $\{X_t\}$ is the BAR(1) process and $P(X_t = j | X_{t-h} = i)$ is given by Eq. (1.2).

Remark 4 As suggested by a referee, it should be pointed out that the maximum likelihood estimation for the ZOIBAR(1) model can not be implemented based on Proposition 2. The reason is that the ZOIBAR(1) process $\{Y_t\}$ is not a Markov chain. To be specific, we have

$$
P(Y_t = y_t | Y_{t-h} = y_{t-h}) \neq P(Y_t = y_t | Y_{t-h} = y_{t-h}, \cdots, Y_1 = y_1),
$$

i.e., Y_t will not only depend on Y_{t-h}, but also depends on Y_{t-k} ($h \leq k \leq t-1$). However, we still investigate the h-step transition probabilities in Proposition 2 based on the following considerations:

(i) The h-step transition probabilities can help us implement the quasi-likelihood estimation for the ZOIBAR(1) model (see Sect. 3.2).
(ii) The h-step transition probability is an important statistical property for the BAR(1) model. Since the ZOIBAR(1) model is a generalization of the BAR(1) model, it is a straightforward idea to investigate the h-step transition probability for the ZOIBAR(1) model.

From the h-step transition probabilities, we can obtain the conditional h-step moments of the ZOIBAR(1) model

$$
E(Y_t | Y_{t-h}) = \begin{cases}
\phi_1 + \phi_2 \dfrac{\phi_0 np + \phi_2(1-p)^n n\beta_h}{\phi_0 + \phi_2(1-p)^n}, & Y_{t-h} = 0, \\
\phi_1 + \phi_2 \dfrac{\phi_1 np + \phi_2 np(1-p)^{n-1}(n\beta_h + \rho^h)}{\phi_1 + \phi_2 np(1-p)^{n-1}}, & Y_{t-h} = 1, \\
\phi_1 + \phi_2(\rho^h Y_{t-h} + n\beta_h), & Y_{t-h} \geq 2,
\end{cases}
\tag{2.2}
$$

and $E(Y_t^2|Y_{t-h})$

$$
= \begin{cases}
\phi_1 + \phi_2 \dfrac{\phi_0[np(1-p) + n^2 p^2] + \phi_2(1-p)^n[n\beta_h(1-\beta_h) + n^2\beta_h^2]}{\phi_0 + \phi_2(1-p)^n}, \; Y_{t-h} = 0, \\[2ex]
\phi_1 + \phi_2 \dfrac{\phi_1[np(1-p) + n^2 p^2] + \phi_2 np(1-p)^{n-1}[(n-1)\beta_h(1 + n\beta_h + 2\rho^h) + \alpha_h]}{\phi_1 + \phi_2 np(1-p)^{n-1}}, \\[1ex]
\qquad Y_{t-h} = 1, \\[1ex]
\phi_1 + \phi_2[\rho^h(1-\rho^h)(1-2p)Y_{t-h} + n\beta_h(1-\beta_h) + (\rho^h Y_{t-h} + n\beta_h)^2], \; Y_{t-h} \geq 2.
\end{cases}
$$

It can be shown that the joint probability generating function of Y_{t-1} and Y_t is

$$
\begin{aligned}
\Phi_{Y_{t-1}, Y_t}&(u, v) \\
&= \phi_0^2 + \phi_0\phi_1(u + v) + \phi_1^2 uv + \phi_0\phi_2[(1 - p + pu)^n + (1 - p + pv)^n] \\
&\quad + \phi_1\phi_2[u(1 - p + pv)^n + v(1 - p + pu)^n] \\
&\quad + \phi_2^2[(1-p)(1-\beta) + p(1-p)(1-\rho)(u+v) + p(p - p\rho + \rho)uv]^n,
\end{aligned}
$$

which shows that $\Phi_{Y_{t-1}, Y_t}(u, v)$ is symmetric in u and v, so the proposed process is time reversible (see Schweer 2015).

3 Estimation procedure

Suppose that $\{Y_t\}$ are observations from the ZOIBAR(1) model defined in Definition 2. Our task is to estimate the unknown parameters $\boldsymbol{\lambda} = (p, \rho, \phi_0, \phi_1)^\top$ from a sample $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_T)$. The parameter n is assumed to be known. Four different estimators, namely probability-based estimator (PBE), quasi-maximum likelihood estimator (QMLE), maximum likelihood estimator (MLE) and Bayesian estimator (BE), are considered.

3.1 Probability-based estimation

The PBE of the parameters p, ϕ_0, ϕ_1 can be obtained from the fact that

$$
P(Y_t = 0) = \phi_0 + \phi_2(1-p)^n \;\; (\text{or } P(Y_t = 1) = \phi_1 + \phi_2 np(1-p)^{n-1}),
$$
$$
P(Y_t = k) = \phi_2 \binom{n}{k} p^k (1-p)^{n-k}, \; k = 2, 3, \ldots.
$$

The probability $p_k := P(Y_t = k)$ can be estimated by using the statistics

$$
\hat{p}_k := \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}_{\{Y_t = k\}}, \; k = 0, 1, 2, \ldots, n.
$$

Therefore, the PBE of the parameters p, ϕ_0, ϕ_1 can be given by

$$\hat{p}_{PB} = \frac{(k+1)\hat{p}_{k+1}}{(k+1)\hat{p}_{k+1} + (n-k)\hat{p}_k}, \quad \hat{\phi}_{2,PB} = \frac{\hat{p}_k \Gamma(k+1)\Gamma(n-k+1)}{\hat{p}_{PB}^k (1 - \hat{p}_{PB})^{n-k} \Gamma(n+1)},$$

$$\hat{\phi}_{0,PB} = \hat{p}_0 - \hat{\phi}_{2,PB}(1 - \hat{p}_{PB})^n \ (\text{or } \hat{\phi}_{1,PB} = \hat{p}_1 - \hat{\phi}_{2,PB} n \hat{p}_{PB}(1 - \hat{p}_{PB})^{n-1}),$$

$$(3.1)$$

where $k = 2, 3, \ldots, n - 1$. We can also obtain the PBE for ρ by taking advantage of $P_{ij}^{(h)}$ given in Proposition 2. However, it is too complex to derived the closed-form estimator $\hat{\rho}_{PB}$. So we turn to moment (MM) estimation and obtain the following corresponding estimator:

$$\hat{\rho}_{MM} = \frac{\sum_{t=1}^{T-1} (Y_t - \bar{Y})(Y_{t+1} - \bar{Y})}{(T-1)\hat{\phi}_{2,PB}^2 n \hat{p}_{PB}(1 - \hat{p}_{PB})}, \quad \text{with } \bar{Y} = \frac{\sum_{t=1}^{T} Y_t}{T}.$$

It is worth mentioning that the selection of k value in Eq. (3.1) is an unavoidable and crucial problem. The reason is that an unappropriate selection of k value may lead to the result that the estimators tend to infinity under some parameter combinations. For example, when we set large values of n, ϕ_0 and ϕ_1 and small value of p in simulation, it is very likely to happen that big numbers are absent in random numbers. Under this circumstance, assuming $k = n - 1$ may fail to produce available estimators. For the convenience of practitioners in practice, we propose a simple method to select k based on the following two steps: (i) let $\mathbf{S} = \{s : \hat{p}_s > 0, \hat{p}_{s+1} > 0, s = 2, 3, \ldots, n-1\}$; (ii) then $k = \underset{s \in \mathbf{S}}{\arg \max}\ \hat{p}_s$.

3.2 Quasi-maximum likelihood estimation

In order to obtain the maximum likelihood estimates of the ZOIBAR(1) process, we must maximize the log-likelihood function

$$\ell(\boldsymbol{\lambda}) = \log P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_T = y_T)$$
$$= \log[P(Y_1 = y_1)P(Y_2 = y_2|Y_1 = y_1)P(Y_3 = y_3|Y_1 = y_1, Y_2 = y_2)$$
$$\cdots P(Y_T = y_T|Y_1 = y_1, \cdots Y_{T-1} = y_{T-1})],$$

where $0 < p < 1$, $\max\left\{-\dfrac{1-p}{p}, -\dfrac{p}{1-p}\right\} < \rho < 1, 0 \leq \phi_0, \phi_1, \phi_2 \leq 1$ and $\phi_0 + \phi_1 + \phi_2 = 1$. However, it is not straightforward to give the closed-form expression for

$$P(Y_t = y_t|Y_1 = y_1, \cdots, Y_{t-1} = y_{t-1}), \ t \in \{2, 3, \ldots, T\}.$$

For this, we follow Maiti et al. (2018) and employ one-step quasi-maximum likelihood approach, which relies on the approximation

$$P(Y_t = y_t | Y_{t-1} = y_{t-1}) \approx P(Y_t = y_t | Y_1 = y_1, \ldots, Y_{t-1} = y_{t-1}), t \in \{2, 3, \ldots, T\}$$

to simplify the computation. Then, the QMLE can be obtained by maximizing the following logarithmic quasi-likelihood function:

$$\ell^*(\lambda) = \log P(Y_1 = y_1) + \sum_{t=2}^{T} \log P(Y_t = y_t | Y_{t-1} = y_{t-1}),$$

rather than maximizing the actual logarithmic likelihood function. The transition probabilities $P(Y_t = y_t | Y_{t-1} = y_{t-1})$, $t \in \{2, \ldots, T\}$ are given by Proposition 2. It is natural to notice that no closed-form expressions for the QMLE can be found, so the use of numerical procedure is inevitable.

3.3 Maximum likelihood estimation

The key to implement the MLE method is to give the likelihood function. The ZOIBAR(1) process is a stationary HMM, where the hidden process $\{X_t\}$ is the BAR(1) model. The BAR(1) model has the stationary binomial marginal distribution B(n, p) and its transition probabilities are given by Eq. (1.2). Following Zucchini et al. (2009), we denote the corresponding transition matrix by $\boldsymbol{\Gamma} = (\gamma_{i+1,j+1})_{i,j=0,\ldots,n}$ with $\gamma_{i+1,j+1} = P(X_{t+1} = j | X_t = i)$ and $\boldsymbol{\Gamma}^k$ is obtained from $\boldsymbol{\Gamma}$ by replacing ρ by ρ^k. We also denote the stationary marginal distribution of $\{X_t\}$ by $\boldsymbol{\delta} = (\delta_{1,k+1})_{k=0,\ldots,n}$ with $\delta_{1,k+1} = P(X_t = k)$. The diagonal matrices $\mathbf{P}(y) := \mathrm{diag}\big(P(Y_t = y | X_t = 0), \ldots, P(Y_t = y | X_t = n)\big) \in [0, 1]^{(n+1) \times (n+1)}$. Using the results and notations from Zucchini et al. (2009), we obtain

$$\boldsymbol{\alpha}_1 = \boldsymbol{\delta}\mathbf{P}(y_1), \ \ \boldsymbol{\alpha}_t = \boldsymbol{\alpha}_{t-1}\boldsymbol{\Gamma}\mathbf{P}(y_t), \ \text{for } t = 2, \ldots, T.$$

where $\boldsymbol{\alpha}_t$ is the vector of "forward probabilities" at time t, i.e., $\alpha_{t,x} = P(Y_1 = y_1, \ldots, Y_t = y_t, X_t = x)$ with $x \in \{0, 1, \ldots, n\}$. Based on the above formulae, the actual likelihood function can be given by

$$\ell(\lambda) = \boldsymbol{\alpha}_T \mathbf{1}^{\top}.$$

3.4 Bayesian estimation

The likelihood inference might be the most commonly used approach in count time series analysis to estimate the model parameters. However, this mainstream method still encounters some difficulties in practice: (i) it is highly affected by the outliers; (ii) the estimators rely on the appropriate selection of numerical optimization procedure and choice of initial values for numerically maximizing the (logarithmic) likelihood function; (iii) it sometimes fails to outperform the moment method and conditional least square method for short time series. As an extension to the classic likelihood method, the Bayesian analysis for time series of counts has attracted

a lot of attention and some articles have arisen in the literature. In this section, our interest is focused on estimating $\boldsymbol{\lambda} = (p, \rho, \phi_0, \phi_1)^\top$ under Bayesian paradigm. As we mentioned in Section 3.2, the actual likelihood function can be approximated by the quasi-likelihood function. Hence, the Bayesian estimator can be derived based on the quasi-likelihood function or actual likelihood function. Besides, the priors play an important role and represent strong pre-experimental assumptions for the possible values of model parameters. Now, we will pay attention to assign the appropriate priors for the model parameters.

3.4.1 Prior for parameter p

Beta distribution is commonly used to play a role as the priors of the autoregressive coefficients. Here, due to the constrain $p \in (0, 1)$ for the ZOIBAR(1) model, we assume that the prior of parameter p is Beta distribution given by

$$f_1(p) = \frac{1}{B(a, b)} p^{a-1}(1 - p)^{b-1}, \ 0 < p < 1,$$

where $B(a, b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)}$ and $a, b > 0$.

3.4.2 Prior for parameter ρ

The Kumaraswamy distribution was introduced by Kumaraswamy (1980) for modeling double bounded random processes with a wide variety of applications, specially in hydrology. Since the parameter ρ for our model is double bounded with $\rho \in \left(\max\left\{-\dfrac{p}{1 - p}, -\dfrac{1 - p}{p}\right\}, 1\right)$, so we assume that the prior of parameter ρ is Kumaraswamy distribution given by

$$f_2(\rho) = \frac{\varphi\delta}{d - c}\left(\frac{\rho - c}{d - c}\right)^{\varphi-1}\left[1 - \left(\frac{\rho - c}{d - c}\right)^\varphi\right]^{\delta-1}, \ c < \rho < d,$$

where $c = \max\left\{-\dfrac{p}{1 - p}, -\dfrac{1 - p}{p}\right\}, d = 1$ and $\delta, \varphi > 0$.

3.4.3 Prior for parameters ϕ_0 and ϕ_1

Dirichlet distribution is a direct extension of Beta distribution and serves as a bridge between distributions. Due to the characteristic of the values of ϕ_0, ϕ_1, ϕ_2, a suitable prior for ϕ_0, ϕ_1 is Dirichlet distribution, which is given below:

$$f_3(\phi_0, \phi_1) = \frac{\Gamma(\theta_0 + \theta_1 + \theta_2)}{\Gamma(\theta_0)\Gamma(\theta_1)\Gamma(\theta_2)}\phi_0^{\theta_0-1}\phi_1^{\theta_1-1}(1 - \phi_0 - \phi_1)^{\theta_2-1},$$
$$\phi_0, \phi_1 \geq 0, \phi_0 + \phi_1 < 1,$$

where $\theta_0, \theta_1, \theta_2 > 0$.

Based on the priors of parameters p, ρ, ϕ_0, ϕ_1 given above, the prior $\pi(\lambda)$ of our model can be given as

$$\pi(\lambda) = \frac{\varphi\delta\Gamma(\theta_0 + \theta_1 + \theta_2)(\rho - c)^{\varphi-1}\phi_0^{\theta_0-1}\phi_1^{\theta_1-1}p^{a-1}(1-p)^{b-1}}{B(a,b)\Gamma(\theta_0)\Gamma(\theta_1)\Gamma(\theta_2)(1-c)^{\varphi}(1-\phi_0-\phi_1)^{1-\theta_2}}$$
$$\times \left[1 - \left(\frac{\rho - c}{1 - c}\right)^{\varphi}\right]^{\delta-1},$$

where $0 < p < 1, c < \rho < 1, \phi_0, \phi_1 \geq 0, \phi_0 + \phi_1 < 1, a, b, \delta, \varphi, \theta_0, \theta_1, \theta_2 > 0$ and $c = \max\left\{-\dfrac{1-p}{p}, -\dfrac{p}{1-p}\right\}$. Hence, the corresponding posterior distribution can be written as

$$\pi(\lambda|Y) = \frac{\pi(\lambda)L(Y|\lambda)}{\int_0^1 \int_c^1 \int_0^1 \int_0^{1-\phi_1} \pi(\lambda)L(Y|\lambda)d\phi_0 d\phi_1 d\rho dp}, \tag{3.2}$$

where $L(Y|\lambda)$ is the quasi-likelihood function or actual likelihood function.

Finally, we consider the Bayesian inference under the square error loss function. It is well known that the Bayesian estimation under the square error loss function is the expectation of the posterior distribution. Therefore, the BE for parameter p can be given by

$$\hat{p} = \int_0^1 \int_c^1 \int_0^1 \int_0^{1-\phi_1} p\pi(\lambda|Y)d\phi_0 d\phi_1 d\rho dp,$$

where $\pi(\lambda|Y)$ is the posterior distribution given in Eq. (3.2). The BE for parameter ρ, ϕ_0, ϕ_1 can be given in a similar way.

3.5 Simulation study

The goal of the simulation study presented in this section is to examine the performances of the PBE, QMLE, MLE and BE previously described and compare their behaviors from the perspective of clean and contaminated data. The initial value $Y_0 \sim \text{ZOIB}(n, p, \phi_0, \phi_1)$. We generate data from the ZOIBAR(1) model and set the sample sizes $T = 50, 300, 1000$ to reflect small ($T = 50$), moderate ($T = 300$) and large ($T = 1000$) sample sizes. The true values of the parameters are selected as:

Model (A): $(n, p, \rho, \phi_0, \phi_1) = (30, 0.5, -0.4, 0.2, 0.1)$;
Model (B): $(n, p, \rho, \phi_0, \phi_1) = (30, 0.4, -0.1, 0.1, 0.2)$;
Model (C): $(n, p, \rho, \phi_0, \phi_1) = (50, 0.5, 0.2, 0.1, 0.1)$;
Model (D): $(n, p, \rho, \phi_0, \phi_1) = (50, 0.7, 0.1, 0.2, 0.2)$.

Note that we assume n is known. In simulations, we apply the mean absolute deviation error (MADE) and standard error (SD) based on $m = 1000$ replications for each

parameter combination to evaluate the performances of the proposed estimators. These criteria are defined as follows:

$$\text{MADE} = \frac{1}{m} \sum_{k=1}^{m} |\hat{\rho}_k - \rho|, \ \ \text{SD} = \frac{1}{m-1} \sqrt{\sum_{k=1}^{m} (\hat{\rho}_k - \bar{\rho})^2},$$

where $\hat{\rho}_k$ is the estimator of ρ in the kth replication and $\bar{\rho} = \frac{1}{m} \sum_{k=1}^{m} \hat{\rho}_k$. The hyperparameters in the prior distributions are assumed to be $a = 3, b = 5, \theta_0 = 2, \theta_1 = 4,$ $\theta_2 = 6, \varphi = 1$ and $\delta = 2$. Figure S2 shows the sample paths and marginal distributions of Models (A)–(D). From Figure S2, we can see that the ZOIBAR(1) model not only has the ability to capture zero-and-one inflation, but also can well explain the bimodal characteristic. Table S2 lists some statistics of Models (A)-(D) including the mean, variance, binomial dispersion index I_d, zero probability p_0 and one probability p_1.

3.5.1 Performance without outliers

In this section, we compare performances of the estimators in two settings by using the clean data, i.e., data in absence of outliers. On the one hand, as can be seen from Table 1, the MADEs and SDs values of the four estimators are reduced if the sample size T increases and this situation is identical to our expectation. Broadly speaking, the MADEs and SDs for the BE are smaller than those for the QMLE and MLE when the sample size is small ($T = 50$). As the sample size increases, the performances of the BE, QMLE and MLE become competitive. The emergence of this phenomenon is understandable since the effect of priors on parameters is negligible when the sample size is prettily large. In addition, although the PBE fails to give satisfactory results compared with the QMLE, MLE and BE, it still has an advantage in calculation. To be specific, the PBE has the closed-form estimators and can help practitioners to avoid the time-consuming approach since it does not need any numerical optimization procedure.

On the other hand, we can always compute the estimates for $(p, \rho, \phi_0, \phi_1)$ according to the discussed methods in Sects. 3.1–3.4, but the obtained estimates did not always satisfy the restrictions on $(p, \rho, \phi_0, \phi_1)$. A pair of estimates $(\hat{p}, \hat{\rho}, \hat{\phi}_0, \hat{\phi}_1)$ is classified as being admissible if it satisfies the conditions on $(p, \rho, \phi_0, \phi_1)$ based on the definition of the ZOIBAR(1) process. Thus, we further evaluate the proposed approaches via referring the percentage of inadmissible estimates of each method with different sample sizes. The corresponding results have been reported in Table 2. For the PBE, the percentages of inadmissible estimates are significantly affected by the parameter combinations and sample sizes. An intuitional phenomenon is that the percentages of inadmissible estimates for the PBE increase if the parameters are close to the boundary (see the cases for $\rho = 0.1$ and $\phi_0 = \phi_1 = 0.1$). Differently, it is clear that the QMLE, MLE and BE almost always produces admissible estimates for all parameter combinations even though the sample size is fairly small and parameters are close to the boundary. To illustrate this statement, we further consider the extremely small sample size and find that the above three methods can always produce admissible estimates even though $T = 30$, especially the Bayesian method. Based on

Table 1 The MADEs and SDs of the estimates

Sample size	Parameter	PBE MADE	PBE SD	QMLE MADE	QMLE SD	MLE MADE	MLE SD	BE MADE	BE SD
$T = 50$	$p = 0.5$	0.1596	0.1190	0.0092	0.0117	0.0092	0.0117	0.0073	0.0111
	$\rho = -0.4$	0.3934	0.2902	0.1437	0.1834	0.1361	0.1747	0.1353	0.1643
	$\phi_0 = 0.2$	0.0497	0.0647	0.0454	0.0567	0.0454	0.0567	0.0377	0.0457
	$\phi_1 = 0.1$	0.0773	0.1995	0.0338	0.0433	0.0338	0.0433	0.0477	0.0349
$T = 300$	$p = 0.5$	0.0621	0.0463	0.0037	0.0046	0.0037	0.0046	0.0076	0.0076
	$\rho = -0.4$	0.3800	0.3399	0.0526	0.0653	0.0495	0.0615	0.0528	0.0644
	$\phi_0 = 0.2$	0.0194	0.0243	0.0194	0.0243	0.0194	0.0243	0.0206	0.0242
	$\phi_1 = 0.1$	0.0137	0.0174	0.0137	0.0174	0.0137	0.0174	0.0148	0.0165
$T = 1000$	$p = 0.5$	0.0332	0.0280	0.0019	0.0024	0.0019	0.0024	0.0078	0.0077
	$\rho = -0.4$	0.3102	0.3268	0.0284	0.0355	0.0262	0.0331	0.0389	0.0350
	$\phi_0 = 0.2$	0.0104	0.0130	0.0104	0.0130	0.0104	0.0130	0.0198	0.0214
	$\phi_1 = 0.1$	0.0076	0.0095	0.0076	0.0095	0.0076	0.0095	0.0074	0.0092
$T = 50$	$p = 0.4$	0.1424	0.0969	0.0112	0.0143	0.0113	0.0144	0.0083	0.0097
	$\rho = -0.1$	0.1482	0.2084	0.1937	0.2397	0.1817	0.2262	0.1480	0.1838
	$\phi_0 = 0.1$	0.0369	0.0459	0.0330	0.0418	0.0330	0.0418	0.0281	0.0337
	$\phi_1 = 0.2$	0.0759	0.1572	0.0457	0.0580	0.0457	0.0580	0.0432	0.0468
$T = 300$	$p = 0.4$	0.0563	0.0420	0.0045	0.0056	0.0045	0.0056	0.0093	0.0038
	$\rho = -0.1$	0.1758	0.2317	0.0666	0.0850	0.0662	0.0846	0.0712	0.0909
	$\phi_0 = 0.1$	0.0140	0.0174	0.0140	0.0174	0.0140	0.0174	0.0155	0.0178
	$\phi_1 = 0.2$	0.0183	0.0228	0.0183	0.0228	0.0183	0.0228	0.0158	0.0207
$T = 1000$	$p = 0.4$	0.0292	0.0256	0.0025	0.0031	0.0025	0.0031	0.0101	0.0013
	$\rho = -0.1$	0.1698	0.2160	0.0366	0.0457	0.0362	0.0452	0.0426	0.0527
	$\phi_0 = 0.1$	0.0077	0.0096	0.0077	0.0096	0.0077	0.0096	0.0147	0.0155
	$\phi_1 = 0.2$	0.0100	0.0124	0.0100	0.0124	0.0100	0.0124	0.0094	0.0100

Table 1 continued

Sample size	Parameter	PBE		QMLE		MLE		BE	
		MADE	SD	MADE	SD	MADE	SD	MADE	SD
$T = 50$	$p = 0.5$	0.1740	0.1163	0.0110	0.0137	0.0110	0.0137	0.0098	0.0134
	$\rho = 0.2$	0.2165	0.1520	0.1565	0.1977	0.1526	0.1929	0.1373	0.1513
	$\phi_0 = 0.1$	0.0350	0.0453	0.0335	0.0433	0.0335	0.0433	0.0289	0.0349
	$\phi_1 = 0.1$	0.0555	0.1540	0.0315	0.0407	0.0315	0.0407	0.0460	0.0328
$T = 300$	$p = 0.5$	0.0710	0.0517	0.0042	0.0053	0.0042	0.0053	0.0076	0.0081
	$\rho = 0.2$	0.2345	0.2318	0.0576	0.0715	0.0566	0.0703	0.0495	0.0635
	$\phi_0 = 0.1$	0.0138	0.0172	0.0138	0.0172	0.0138	0.0172	0.0152	0.0175
	$\phi_1 = 0.1$	0.0136	0.0172	0.0136	0.0172	0.0136	0.0172	0.0153	0.0168
$T = 1000$	$p = 0.5$	0.0353	0.0280	0.0024	0.0029	0.0024	0.0029	0.0078	0.0078
	$\rho = 0.2$	0.2161	0.2335	0.0303	0.0384	0.0302	0.0382	0.0341	0.0398
	$\phi_0 = 0.1$	0.0074	0.0093	0.0074	0.0093	0.0074	0.0093	0.0144	0.0152
	$\phi_1 = 0.1$	0.0077	0.0096	0.0077	0.0096	0.0077	0.0096	0.0075	0.0087
$T = 50$	$p = 0.7$	0.2094	0.1707	0.0106	0.0133	0.0105	0.0133	0.0067	0.0070
	$\rho = 0.1$	0.1615	0.2243	0.2189	0.2683	0.2099	0.2573	0.1333	0.1696
	$\phi_0 = 0.1$	0.0513	0.0691	0.0433	0.0553	0.0433	0.0553	0.0363	0.0447
	$\phi_1 = 0.2$	0.0804	0.1783	0.0445	0.0561	0.0445	0.0561	0.0420	0.0456
$T = 300$	$p = 0.7$	0.0737	0.0593	0.0041	0.0051	0.0041	0.0051	0.0082	0.0051
	$\rho = 0.1$	0.3156	0.4352	0.0793	0.1008	0.0781	0.0990	0.0799	0.1054
	$\phi_0 = 0.1$	0.0180	0.0227	0.0180	0.0227	0.0180	0.0227	0.0195	0.0229
	$\phi_1 = 0.2$	0.0180	0.0226	0.0180	0.0226	0.0180	0.0226	0.0192	0.0236
$T = 1000$	$p = 0.7$	0.0339	0.0285	0.0023	0.0029	0.0023	0.0029	0.0086	0.0028
	$\rho = 0.1$	0.4423	0.5335	0.0418	0.0524	0.0414	0.0519	0.0435	0.0553
	$\phi_0 = 0.1$	0.0102	0.0126	0.0102	0.0126	0.0102	0.0126	0.0200	0.0215
	$\phi_1 = 0.2$	0.0104	0.0130	0.0104	0.0130	0.0104	0.0130	0.0152	0.0181

Table 2 Percentages of admissible estimates

$(n, p, \rho, \phi_0, \phi_1)$	Sample size	PBE	QMLE	MLE	BE
$(30, 0.5, -0.4, 0.2, 0.1)$	30	0.2860	0.9450	0.9500	0.9980
	50	0.4560	0.9920	0.9920	1.0000
	100	0.6630	1.0000	1.0000	1.0000
	300	0.8700	1.0000	1.0000	1.0000
	500	0.9320	1.0000	1.0000	1.0000
	1000	0.9510	1.0000	1.0000	1.0000
$(30, 0.4, -0.1, 0.1, 0.2)$	30	0.3550	0.9140	0.9290	1.0000
	50	0.5100	0.9900	0.9930	0.9970
	100	0.7210	1.0000	1.0000	0.9990
	300	0.9320	1.0000	1.0000	1.0000
	500	0.9620	1.0000	1.0000	0.9990
	1000	0.9860	1.0000	1.0000	1.0000
$(50, 0.5, 0.2, 0.1, 0.1)$	30	0.1590	0.9080	0.9080	1.0000
	50	0.2790	0.9900	0.9900	1.0000
	100	0.4510	1.0000	1.0000	1.0000
	300	0.7730	1.0000	1.0000	1.0000
	500	0.8770	1.0000	1.0000	1.0000
	1000	0.9690	1.0000	1.0000	1.0000
$(50, 0.7, 0.1, 0.2, 0.2)$	30	0.0870	0.8720	0.8940	1.0000
	50	0.1650	0.9810	0.9820	1.0000
	100	0.2510	0.9900	0.9930	1.0000
	300	0.3970	1.0000	1.0000	1.0000
	500	0.4520	1.0000	1.0000	1.0000
	1000	0.4980	1.0000	1.0000	1.0000

these discussions, we conclude that the BE is more reliable than the other methods based on this aspect.

3.5.2 Performance with outliers

In this section, we compare performances of the estimators by using the contaminated data. We consider additive outlier generating mechanisms with positive and negative outliers, which can be given in details as follows:

$$Z_t = \begin{cases} \min\{Y_t + \zeta, n\}, & t = \tau_1, \tau_2, \ldots, \tau_k, \\ Y_t, & \text{otherwise,} \end{cases} \quad (3.3)$$

and

$$Z_t = \begin{cases} \max\{Y_t - \zeta, 0\}, & t = \tau_1, \tau_2, \ldots, \tau_k, \\ Y_t, & \text{otherwise,} \end{cases} \quad (3.4)$$

where Z_t is the contaminated observed value at specific time τ_i, $i = 1, \ldots, k$, and ζ is the size of outlier.

In this article, we consider cases of consecutive and isolated outliers and assume the size and number of outliers are $\zeta = 2$ and $k = 3, 5$, respectively. For the consecutive outliers, we set $\tau_1 = T/2 + 1$ and illustrate this outliers generating mechanism as follows: when the sample size is $T = 50$ and the number of outliers is $k = 3$, then we will observe 3 consecutive outliers to occur at time $t = 26, 27$ and 28. Differently, the outliers are added to the observed data in randomly chosen positions for the isolated outliers.

Tables S3–S6 report the MADEs and SDs of the estimators with sample size $T = 50$ and the contaminated data generated from Eqs. (3.3)–(3.4), respectively. Comparing Tables S3–S4 with Tables S5–S6, it can be seen that the consecutive and isolated outliers, positive and negative outliers have similar effect on the estimators. Meanwhile, we find that the performances of the estimators become more worse as the number of outliers increases, as expected. Additive outliers have been known to be specifically harmful for the estimation of dependence parameters and this point has been verified by the phenomenon that the accuracy of the QMLE, MLE and BE for parameter ρ has been significantly reduced. Comparing the four estimation methods, we can see that the BE still yields the smallest values of the MADEs and SDs in most situations.

3.5.3 Discussion

We have compared the proposed estimators based on different aspects. For the clean data, the BE yields the smallest values of MADEs and SDs when the simple size is small. Moreover, the BE can also give the biggest percentages of admissible estimators in most cases. For the contaminated data, the BE also gives more robust performances than other methods. In summary, we recommend the use of the Bayesian approach for our model.

It must be pointed out that the quasi-likelihood function is used in Bayesian estimation rather than the actual likelihood function. From Tables 1 and S3–S6, we can see that the MLE and QMLE always give similar performances. To be specific, the quasi-maximum likelihood and maximum likelihood methods yield the same estimates when we consider parameters p, ϕ_0 and ϕ_1. Based on the parameter ρ, the MADEs and SDs values of the MLE are a little smaller than the QMLE when the sample size is small. Hence, we conclude that the improvements, which can be accomplished by the MLE, to the QMLE are extremely limited. Moreover, compared to the QMLE, the MLE need more calculation and time costs. To illustrate this point, we give a simulation study in Table S7, which shows the durations, MADEs and SDs of the MLE and QMLE. From Table S7, we can see that the MLE need much more running time than the QMLE, especially in the cases of large n values ($n = 50$ in Models (C) and (D)). Based on the above discussions, the employment of the quasi-likelihood function in the Bayesian estimation is a more reasonable option than the actual likelihood function. Besides, it is worth noting that the time cost of the MLE is not sensitive about the sample size. However, by comparing Models (A) and (B) with Models (C) and (D), the increase of the parameter n value will greatly increase the calculation pressure and running time.

As suggested by a referee, the scenario $\phi_0 = \phi_1 = 0$ is considered. We pointed out in Remark 2 that the ZOIBAR(1) model will reduce to the BAR(1) model when $\phi_0 = \phi_1 = 0$. Furthermore, $\phi_0 = \phi_1 = 0$ is the boundary of the parameter space for the unknown parameters $\boldsymbol{\lambda} = (p, \rho, \phi_0, \phi_1)^\top$. To investigate the performances of the proposed in the case of $\phi_0 = \phi_1 = 0$, an addition simulation study is given in Table S8. The parameter combinations are selected as:

Model (E1): $(n, p, \rho, \phi_0, \phi_1) = (10, 0.5, -0.4, 0, 0)$;
Model (E2): $(n, p, \rho, \phi_0, \phi_1) = (10, 0.7, 0.1, 0, 0)$.

From Table S8, we can see that the simulation results of the QMLE, MLE and BE are still satisfactory. This phenomenon implies that the above three methods are still effective even though we consider degradation forms of the ZOIBAR(1) model.

4 Testing for one inflation in BAR(1) model

Weiß and Pollett (2014) developed a test for detecting overdispersion via taking advantage of the sample binomial dispersion index (without bias correction). Kim et al. (2018) improved and generalized their results by deriving a novel test concerning overdispersion and further proposed a zero-inflation index, which was employed to test for detecting zero inflation. Moreover, the authors introduced a joint test for zero inflation/deflation and overdispersion/underdispersion with respect to a BAR(1) process. In this article, our goal is to focus on the following two perspectives:

 (i) The first aim is to test the null hypothesis of a BAR(1) process against the alternative that the data exhibit more ones than implied by a BAR(1) model. To achieve this objective, a binomial one-inflation index is constructed and further applied to develop a method to test whether one inflated with respect to a BAR(1) model.
 (ii) The second target is to generalize the interesting results in Kim et al. (2018) via putting forward two new joint tests for zero inflation/deflation, one inflation/deflation and overdispersion/underdispersion with respect to a BAR(1) process.

4.1 Binomial one-inflation index

As we discussed in Sect. 2, in our opinion, it is unreasonable to determine one inflation or one deflation entirely relying on the 1's proportion since some other important statistical properties should be necessarily considered, in particular the mean of the data (similar discussion has been given in Sect. 2). Hence, we propose a (sample) binomial one-inflation index to measure the departure from binomial model in the following. Let X be an integer-valued random variable with a finite range $\{0, 1, \ldots, n\}$, then the binomial one-inflation index z_1 and sample binomial one-inflation index $\hat{z}_1$ can be defined as follows:

$$z_1 = p_1 \mu^{-1} \left(1 - \frac{\mu}{n}\right)^{1-n} \in (0, \infty), \quad \hat{z}_1 = \hat{p}_1 \bar{X}^{-1} \left(1 - \frac{\bar{X}}{n}\right)^{1-n},$$

where p_1 and μ are the one probability and mean of the random variable X, $\hat{p}_1 = \frac{1}{T}\sum_{t=1}^{T}\mathbb{I}_{\{X_t=1\}}$ and $\bar{X} = \frac{1}{T}\sum_{t=1}^{T}X_t$. If X is binomially distributed, then $z_1 = 1$ and $\hat{z}_1$ should be close to 1. If $z_0 > 1$, naturally, X is said to show one inflation with respect to the binomial distribution. Thus, we can compute the binomial one-inflation index for the ZOIBAR(1) process as follows:

$$z_1 = \frac{\phi_1 + \phi_2 np(1-p)^{n-1}}{\phi_1 + \phi_2 np}\left(1 - \frac{\phi_1 + \phi_2 np}{n}\right)^{1-n}.$$

From the above equation, we find that z_1 increases as the parameter ϕ_1 increases when the parameters ϕ_2, n and p are fixed.

4.2 Asymptotic distribution of binomial indices

Our objective is to investigate the limiting behavior of the trivariate statistic, which is composed of the sample binomial zero index, sample binomial index of dispersion and sample binomial one index when the data-generating process (DGP) is a BAR(1) process. For this purpose, we first consider the vector-valued process $\{\mathbf{Z}_t\}$ defined by

$$\mathbf{Z}_t := \begin{pmatrix} \mathbb{I}_{\{X_t=0\}} - (1-p)^n \\ X_t - np \\ X_t^2 - np(np+1-p) \\ \mathbb{I}_{\{X_t=1\}} - np(1-p)^{n-1} \end{pmatrix} \text{ with } E(\mathbf{Z}_t) = \mathbf{0}, \tag{4.1}$$

and we derive the asymptotic distribution of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{Z}_t$ in the following theorem.

Theorem 1 *Let $\{X_t\}$ be a BAR(1) process with $n \geq 2$, with $p_0 = (1-p)^n$ and $p_1 = np(1-p)^{n-1}$, define $\{\mathbf{Z}_t\}$ as in Eq. (4.1). Then,*

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\mathbf{Z}_t \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}),$$

where the covariance matrix $\boldsymbol{\Sigma}$ has entries $\sigma_{ij}(= \sigma_{ji})$ with $i, j = 1, 2, 3, 4$, which will be given in the proof.

Next theorem establishes the joint asymptotic normality of the indices by using Theorem 1.

Theorem 2 *Let $\{X_t\}$ be a BAR(1) process with $n \geq 2$. Then, we have*

$$\begin{pmatrix} \hat{z}_0 - 1 \\ \hat{I}_d - 1 \\ \hat{z}_1 - 1 \end{pmatrix} \xrightarrow{d} N(\mathbf{0}, \boldsymbol{\Sigma}'),$$

where the covariance matrix $\boldsymbol{\Sigma}'$ has entries $\sigma_{ij}'(= \sigma_{ji}')$ with $i, j = 1, 2, 3$, which will be given in the proof.

The limiting behavior of $\hat{z}_0, \hat{I}_d, \hat{z}_1$ shows that they are asymptotically unbiased estimators of z_0, I_d, z_1, respectively. However, if we compute $\hat{z}_0, \hat{I}_d, \hat{z}_1$ from a time series of finite length T, these estimators would be obviously biased. Therefore, it is necessary and meaningful to propose a bias-corrected approximation to the mean of z_0, I_d, z_1 in the following theorem.

Theorem 3 *Let $\{X_t\}$ be a BAR(1) process with $n \geq 2$. Then, the means of $\hat{z}_0, \hat{I}_d, \hat{z}_1$ are asymptotically given by*

$$E(\hat{z}_0) \approx 1 - \frac{(n-1)p(1+\rho)}{2T(1-p)(1-\rho)}, \quad E(\hat{I}_d) \approx 1 - \frac{(n-1)(1+\rho)}{Tn(1-\rho)},$$

$$E(\hat{z}_1) \approx 1 - \frac{(n-1)(np-2)(1+\rho)}{2Tn(1-p)(1-\rho)}.$$

Theorem 3 can be applied to define (approximately) bias-corrected indices $\hat{z}_{0;\mathrm{corr}}$, $\hat{I}_{d;\mathrm{corr}}$ and $\hat{z}_{1;\mathrm{corr}}$ as

$$\hat{z}_{0;\mathrm{corr}} = \hat{z}_0 + \frac{(n-1)\hat{p}(1+\hat{\rho})}{2T(1-\hat{p})(1-\hat{\rho})}, \qquad \hat{I}_{d;\mathrm{corr}} = \hat{I}_d + \frac{(n-1)(1+\hat{\rho})}{Tn(1-\hat{\rho})},$$

$$\hat{z}_{1;\mathrm{corr}} = \hat{z}_1 + \frac{(n-1)(n\hat{p}-2)(1+\hat{\rho})}{2Tn(1-\hat{p})(1-\hat{\rho})},$$

where the estimates for p and ρ are recommended to use a plug-in approach, i.e.,

$$\hat{p} = \frac{\bar{X}}{n}, \quad \hat{\rho} = \frac{\sum_{t=1}^{T-1}(X_t - \bar{X})(X_{t+1} - \bar{X})}{\sum_{t=1}^{T}(X_t - \bar{X})^2}, \quad \text{with } \bar{X} = \frac{\sum_{t=1}^{T} X_t}{T}.$$

4.3 Tests based on binomial indices

As we discussed before, our first task is to test the null hypothesis of a BAR(1) process against the alternative that the data exhibit more ones than implied by a BAR(1) model. In other words, we are interested in testing $H_0 : x_1, \ldots, x_T$ stem from BAR(1) with $p_1 = \mu(1 - \mu/n)^{n-1}$ against the alternative hypothesis $H_1 : x_1, \ldots, x_T$ with $p_1 \neq np(1-p)^{n-1}$. We can take advantage of the sample one-inflation index and its corresponding bias-corrected version to propose the following two test statistics:

$$\mathrm{U}_1 = \frac{\hat{z}_1 - 1}{\sqrt{\sigma_{33}'/T}}, \quad \mathrm{U}_{1;\mathrm{corr}} = \frac{\hat{z}_{1;\mathrm{corr}} - 1}{\sqrt{\sigma_{33}'/T}}.$$

Let q_γ be the γ-quantile of the standard normal distribution, i.e., $\Phi(q_\gamma) = \gamma$, where $\gamma \in (0, 1)$ and $\Phi(\cdot)$ is the distribution function of the standard normal distribution. We reject H_0 at a significance level γ if the test statistics U_1 and $\mathrm{U}_{1;\mathrm{corr}}$ violate the two-sided critical values, i.e., $\mathrm{U}_1, \mathrm{U}_{1;\mathrm{corr}} \notin (q_{\frac{\gamma}{2}}, q_{1-\frac{\gamma}{2}})$. Alternatively, we can check if the p value $2[1 - \Phi(|\mathrm{U}_1|)]$ and $2[1 - \Phi(|\mathrm{U}_{1;\mathrm{corr}}|)]$ fall below the significance level

γ. If a hypothetical value for $\tilde{\sigma}_{33}$ is not available, we recommend to use a plug-in approach, i.e., to replace p and ρ by their moment estimates in the formula for $\tilde{\sigma}_{33}$ given by Theorem 2.

Now, we pay attention to accomplish the second target, which is testing H_0 : $x_1, \ldots, x_T$ stem from BAR(1) with $n\sigma^2 = \mu(n - \mu)$ and/or $p_0 = (1 - \mu/n)^n$ and/or $p_1 = \mu(1 - \mu/n)^{n-1}$ against the alternative hypothesis H_1 : $x_1, \ldots, x_T$ with $n\sigma^2 \neq \mu(n - \mu)$ and/or $p_0 \neq (1 - p)^n$ and/or $p_1 \neq np(1 - p)^{n-1}$. Under the null hypothesis, $U_{\text{joint}} := T(\hat{\boldsymbol{I}}_{\text{joint}} - \boldsymbol{I}_0)^\top (\hat{\boldsymbol{\Sigma}}')^{-1}(\hat{\boldsymbol{I}}_{\text{joint}} - \boldsymbol{I}_0)$ and $U_{\text{joint;corr}} := T(\hat{\boldsymbol{I}}_{\text{joint;corr}} - \boldsymbol{I}_0)^\top (\hat{\boldsymbol{\Sigma}}')^{-1}(\hat{\boldsymbol{I}}_{\text{joint;corr}} - \boldsymbol{I}_0)$ follow an asymptotic χ^2-distribution with 3 degrees of freedom, with

$$\hat{\boldsymbol{I}}_{\text{joint}} = (\hat{z}_0, \hat{I}_d, \hat{z}_1)^\top, \ \hat{\boldsymbol{I}}_{\text{joint;corr}} = (\hat{z}_{0;\text{corr}}, \hat{I}_{d;\text{corr}}, \hat{z}_{1;\text{corr}})^\top, \ \boldsymbol{I}_0 = (1, 1, 1)^\top,$$

and $\hat{\boldsymbol{\Sigma}}'$ being the appropriate (plug-in) covariance matrix from Theorem 2. Based on the two test statistics, we reject H_0 at significance level γ if $U_{\text{joint}} \geq \chi^2_{3,1-\gamma}$ and $U_{\text{joint;corr}} \geq \chi^2_{3,1-\gamma}$, respectively, where $\chi^2_{3,1-\gamma}$ denotes the $(1 - \gamma)$-quantile of the χ^2_3-distribution.

4.4 Simulation study

In this section, some simulations are conducted to investigate the performances of the four test methods. We select the significance level $\gamma = 0.05, 0.1$, sample size $T = 50, 100, 300, 1000$ and each experiment is based on 10,000 replications.

4.4.1 Size study

For analyzing the empirical size, the following parameter combinations are considered: (E1) $(n, p, \rho) = (10, 0.15, 0.25)$, (E2) $(n, p, \rho) = (10, 0.15, 0.5)$, (E3) $(n, p, \rho) = (10, 0.15, 0.75)$. As can be seen from Table 3, for the empirical sizes, all the test statistics give the satisfactory performances and the empirical sizes for Models (E1)–(E3) become more closer to the significant levels $\gamma = 0.05, 0.1$ as the sample size increases. Comparing with the test statistics U_1 and U_{joint}, the bias-corrected test statistics $U_{1;\text{corr}}$ and $U_{\text{joint;corr}}$ give better performances, especially for the small sample size $T = 50$.

4.4.2 Power study

We consider the ZOIBAR(1) model with following parameter combinations to further investigate the power of four test statistics under the alternative hypothesis:

(F1) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.25, 0, 0.1)$,

(F2) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.25, 0, 0.2)$,

(F3) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.25, 0, 0.3)$;

Table 3 Empirical sizes for test statistics

(n, p, ρ)	Sample size	U_1		$U_{1;corr}$		U_{joint}		$U_{joint;corr}$	
		$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$
(10, 0.15, 0.25)	$T = 50$	0.0944	0.0444	0.0950	0.0459	0.0909	0.0473	0.0944	0.0512
	$T = 100$	0.0951	0.0489	0.0964	0.0489	0.0946	0.0519	0.0979	0.0529
	$T = 300$	0.1022	0.0488	0.1020	0.0484	0.0969	0.0506	0.0967	0.0504
	$T = 1000$	0.1009	0.0510	0.1006	0.0511	0.0990	0.0503	0.0990	0.0506
(10, 0.15, 0.5)	$T = 50$	0.0936	0.0427	0.0946	0.0429	0.0851	0.0436	0.0870	0.0465
	$T = 100$	0.0956	0.0461	0.0949	0.0454	0.0925	0.0462	0.0920	0.0484
	$T = 300$	0.0981	0.0515	0.0981	0.0517	0.0954	0.0493	0.0972	0.0509
	$T = 1000$	0.1006	0.0515	0.0996	0.0519	0.0972	0.0509	0.0981	0.0513
(10, 0.15, 0.75)	$T = 50$	0.1196	0.0612	0.1160	0.0591	0.1154	0.0657	0.1100	0.0640
	$T = 100$	0.1036	0.0530	0.1040	0.0531	0.0987	0.0527	0.0966	0.0523
	$T = 300$	0.1009	0.0515	0.0995	0.0512	0.0990	0.0479	0.0988	0.0486
	$T = 1000$	0.1001	0.0510	0.1005	0.0512	0.0980	0.0490	0.0985	0.0489

Table 4 Empirical powers for test statistics

$(n, p, \rho, \phi_0, \phi_1)$	Sample size	U_1		$U_{1;corr}$		U_{joint}		$U_{joint;corr}$	
		$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$
(10, 0.15, 0.25, 0, 0.1)	$T = 50$	0.2830	0.1857	0.2870	0.1880	0.1915	0.1159	0.1971	0.1204
	$T = 100$	0.4566	0.3379	0.4596	0.3394	0.3205	0.2120	0.3241	0.2149
	$T = 300$	0.8323	0.7466	0.8340	0.7474	0.6930	0.5738	0.6957	0.5768
	$T = 1000$	0.9993	0.9976	0.9993	0.9976	0.9944	0.9858	0.9945	0.9859
(10, 0.15, 0.25, 0, 0.2)	$T = 50$	0.6539	0.5366	0.6551	0.5392	0.4810	0.3572	0.4849	0.3601
	$T = 100$	0.8976	0.8314	0.8972	0.8317	0.7854	0.6817	0.7855	0.6826
	$T = 300$	0.9999	0.9993	0.9999	0.9993	0.9979	0.9939	0.9979	0.9938
	$T = 1000$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
(10, 0.15, 0.25, 0, 0.3)	$T = 50$	0.9156	0.8558	0.9151	0.8545	0.8021	0.7021	0.8021	0.6991
	$T = 100$	0.9956	0.9888	0.9956	0.9887	0.9802	0.9612	0.9804	0.9608
	$T = 300$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	$T = 1000$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
(10, 0.15, 0.5, 0, 0.1)	$T = 50$	0.2788	0.1847	0.2844	0.1886	0.1971	0.1196	0.2031	0.1244
	$T = 100$	0.4285	0.3109	0.4316	0.3157	0.3108	0.2028	0.3097	0.2044
	$T = 300$	0.8095	0.7171	0.8113	0.7191	0.6782	0.5564	0.6792	0.5579
	$T = 1000$	0.9986	0.9964	0.9986	0.9964	0.9931	0.9850	0.9932	0.9852
(10, 0.15, 0.5, 0, 0.2)	$T = 50$	0.6489	0.5282	0.6472	0.5277	0.4936	0.3646	0.4903	0.3645
	$T = 100$	0.8831	0.8146	0.8828	0.8150	0.7736	0.6707	0.7734	0.6726
	$T = 300$	0.9998	0.9986	0.9998	0.9986	0.9967	0.9919	0.9967	0.9920
	$T = 1000$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 4 continued

$(n, p, \rho, \phi_0, \phi_1)$	Sample size	U_1		$U_{1;corr}$		U_{joint}		$U_{joint;corr}$	
		$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$	$\gamma = 0.1$	$\gamma = 0.05$
(10, 0.15, 0.5, 0, 0.3)	$T = 50$	0.9054	0.8416	0.9025	0.8384	0.7965	0.6928	0.7937	0.6909
	$T = 100$	0.9944	0.9865	0.9943	0.9863	0.9775	0.9562	0.9765	0.9552
	$T = 300$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	$T = 1000$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
(10, 0.15, 0.75, 0, 0.1)	$T = 50$	0.2710	0.1858	0.2714	0.1875	0.2341	0.1518	0.2294	0.1476
	$T = 100$	0.3887	0.2885	0.3897	0.2919	0.3440	0.2362	0.3401	0.2357
	$T = 300$	0.7180	0.6178	0.7200	0.6204	0.6338	0.5194	0.6323	0.5207
	$T = 1000$	0.9896	0.9768	0.9901	0.9771	0.9731	0.9485	0.9730	0.9487
(10, 0.15, 0.75, 0, 0.2)	$T = 50$	0.6039	0.4934	0.6009	0.4913	0.5022	0.3813	0.4921	0.3757
	$T = 100$	0.8289	0.7496	0.8264	0.7481	0.7543	0.6556	0.7496	0.6526
	$T = 300$	0.9960	0.9918	0.9959	0.9918	0.9882	0.9776	0.9884	0.9769
	$T = 1000$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
(10, 0.15, 0.75, 0, 0.3)	$T = 50$	0.8685	0.7978	0.8616	0.7898	0.7776	0.6874	0.7694	0.6794
	$T = 100$	0.9829	0.9671	0.9820	0.9654	0.9605	0.9315	0.9575	0.9284
	$T = 300$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	$T = 1000$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

(G1) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.5, 0, 0.1)$,

(G2) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.5, 0, 0.2)$,

(G3) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.5, 0, 0.3)$;

(H1) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.75, 0, 0.1)$,

(H2) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.75, 0, 0.2)$,

(H3) $(n, p, \rho, \phi_0, \phi_1) = (10, 0.15, 0.75, 0, 0.3)$.

The rejection rates of the tests are reported in Table 4. It can be observed that the rejection rates are increasing with the sample size and ϕ_1. In each case, the rejection rates of U_1 and $U_{1;corr}$ are larger than U_{joint} and $U_{joint;corr}$. Reason for this phenomenon is that the DGP is ZOIBAR(1) model with $\phi_0 = 0$ and different parameter values of ρ and ϕ_1 and the distinction between DGP and the corresponding BAR(1) model is the 1's proportion. U_1 and $U_{1;corr}$ are constructed based on the binomial one-inflation index and thus can give higher rejection rates. U_{joint} and $U_{joint;corr}$ take the information of I_d and z_0 into account but these two indices of DGP and the corresponding BAR(1) model are close. Hence, compared with U_1 and $U_{1;corr}$, the powers of U_{joint} and $U_{joint;corr}$ are reduced. These findings can also be corroborated by another power analysis, which is given by power graphs in Figs. 1 and 2. There is an abnormal occurrence in Figs. 1 and 2. To be specific, when $\phi_0 = 0.2, 0.3$, the powers of $U_{1;corr}$ and $U_{joint;corr}$ decrease as ϕ_1 increases under the condition $0 < \phi_1 \leq 0.2$. This circumstance may be explained by that $\phi_0 = 0.2, 0.3$ and $0 < \phi_1 \leq 0.2$ may effect the mean of the model and thus lead to the result that the $\hat{z}_0$, $\hat{I}_d$ and $\hat{z}_1$ cannot well reflect the deviation between DGP and the corresponding BAR(1) model.

5 Real data examples

In this section, three applications are conducted to illustrate the performance of the ZOIBAR(1) process in explaining zero-inflation, one-inflation and overdispersion phenomena. We compare our process with four extended BAR(1) models, which were proposed by Möller et al. (2018) to handle zero inflation:

- BAR(1) model with zeros at random (RZ-BAR(1) model);
- BAR(1) model with innovational zeros (IZ-BAR(1) model);
- zero-inflated binomial thinning AR(1) model (ZIB-AR(1) model);
- BAR(1) model with zero threshold (ZT-BAR(1) model).

Moreover, we take the classical BAR(1) process as a benchmark due to its usefulness and popularity in real-world applications. Although we pointed out the BE is the most appropriate method for our model, for the purpose of fairness, the (quasi) conditional maximum likelihood approach is employed for all the alternative models. We also compute the following statistics of the fitted models: Akaike information criterion (AIC), Bayesian information criterion (BIC), binomial dispersion index I_d, binomial zero-inflation index z_0 and binomial one-inflation index z_1.
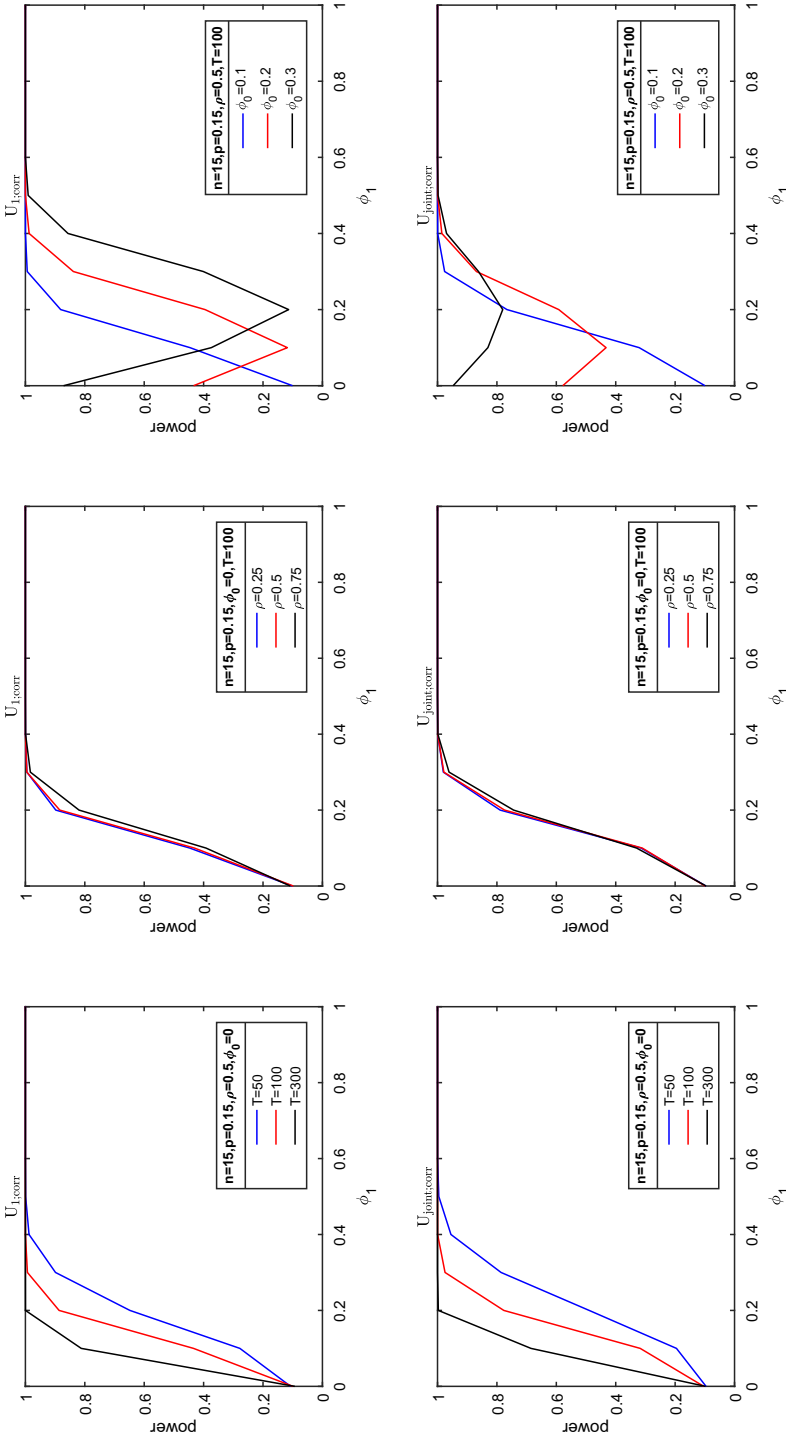
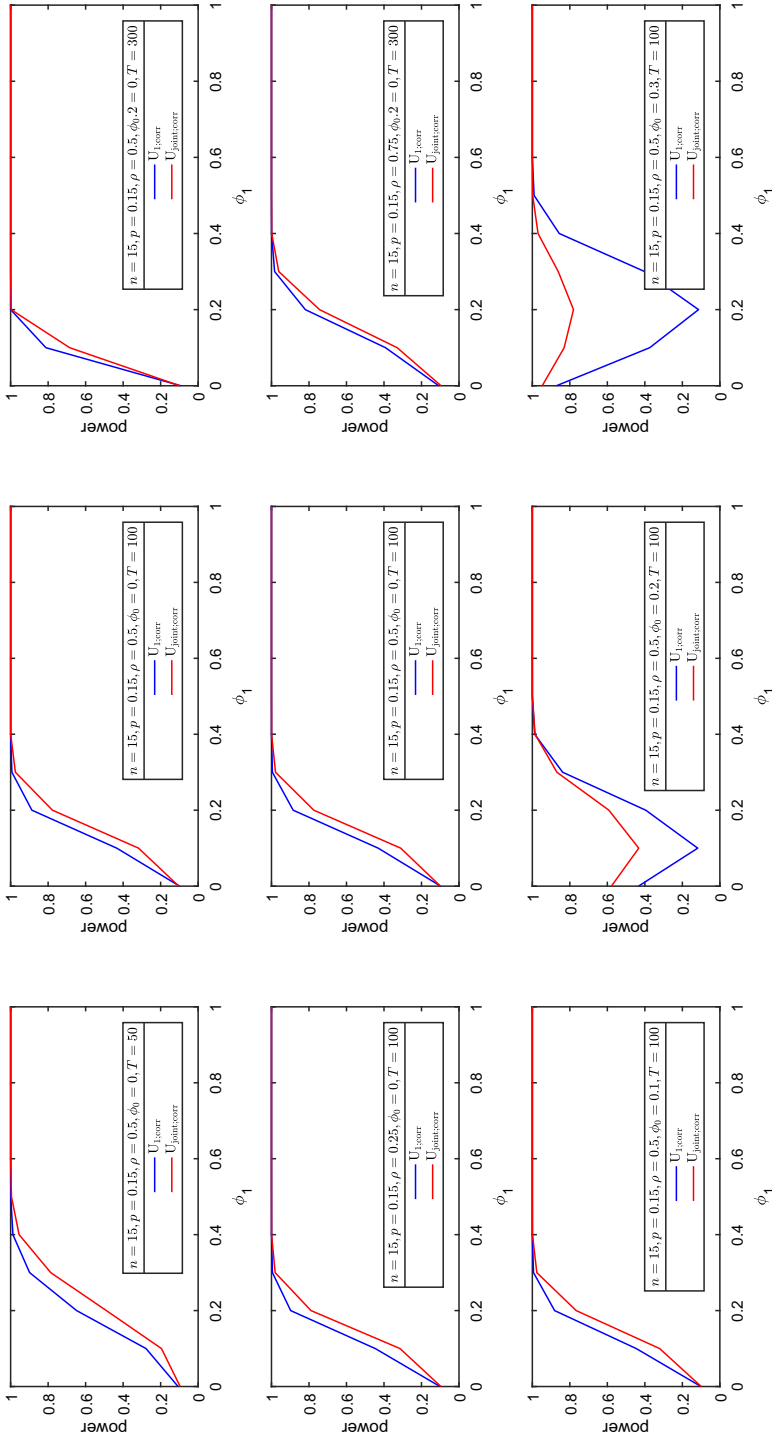**Fig. 1** Power analysis for test statistics with significance level $\gamma = 0.1$

Fig. 2 Power analysis for test statistics with significance level $\gamma = 0.1$

Table 5 Descriptive statistics for rainy-days counts in Germany

Mean	Variance	I_d	p_0	z_0	p_1	z_1
3.8371	3.6753	2.1198	0.0479	12.4623	0.0831	2.5437

5.1 Rainy-days counts in Germany

In this section, we consider the number of rainy-days per week at Hamburg-Neuwiedenthal in Germany. The data were collected from January 1st, 2005 till December 31th, 2010 by the German Weather Service (DWD ="Deutscher Wetter-Dienst," http://www.dwd.de/), where weeks are defined from Saturday to Friday. The length of the data is 313 and the fixed upper limit is $n = 7$. This data set was also investigated by Chen et al. (2020).

The sample path, histogram, ACF, PACF and summary statistics of the observations are given in Figure S3 and Table 5. The binomial dispersion index I_d of the data is 1.7084, which indicates that the data set is overdispersed. The zero frequency and one frequency are, respectively, 0.0479 and 0.0831, which are close to zero. However, as we discussed before, relying entirely on 0's or 1's proportion may mislead one to determine zero inflation or one inflation and the corresponding indices are more reasonable assessment criteria rather than 0's or 1's proportion. The sample binomial zero-inflation and one-inflation indices are, respectively, 12.4623 and 2.5437, which make us confirm that this data set is zero-inflated and one-inflated.

As can be seen from Table 6, the ZT-BAR(1) and BAR(1) models are not suitable for this data set since they give the biggest values of AIC and BIC among alternative models. Furthermore, the ZT-BAR(1) and BAR(1) processes are the only two models, which fail to reflect the overdispersion feature of the counts. Although the IZ-BAR(1) model yields the smaller values of AIC and BIC and captures the overdispersion characteristic of the data, it cannot give the information about the zero and one patterns of the data since it is complicated to derive a closed-form expression for the stationary marginal distribution of the IZ-BAR(1) process. As a special case of our model, the ZR-BAR(1) model can accurately explain zero inflation, but it gives the wrong information that the data set is zero-deflated. Based on AIC and BIC, the ZIB-AR(1) and ZOIBAR(1) processes are the most appropriate model to fit the data. However, the ZIB-AR(1) model also encounters the difficulty that the information about zero and one patterns are not easily obtained. Differently, the ZOIBAR(1) process can accurately capture zero-inflation, one-inflation and overdispersion characteristics and gives the smallest values of AIC and BIC. Hence, we recommend the use of the ZOIBAR(1) process to fit the data set.

To further compare the above models, we consider their corresponding Pearson residual analysis. The standardized Pearson residual is defined as

$$e_t = \frac{X_t - E(Y_t|Y_{t-1})}{\sqrt{\text{Var}(Y_t|Y_{t-1})}}, \ t = 2, \ldots, T.$$

Table 6 Estimates of the parameters and statistics for rainy-days counts in Germany

Model	Estimates	AIC	BIC	I_d	z_0	z_1
ZOIBAR(1)	$\hat{p} = 0.6054$	1310.4419	1325.4267	1.6420	11.9936	2.4473
	$\hat{\rho} = 0.0977$					
	$\hat{\phi}_0 = 0.0463$					
	$\hat{\phi}_1 = 0.0676$					
BAR(1)	$\hat{p} = 0.5483$	1390.1823	1397.6747	1.0000	1.0000	1.0000
	$\hat{\rho} = 0.1323$					
ZR-BAR(1)	$\hat{p} = 0.0898$	1346.245	1357.4836	1.3376	11.9574	0.7065
	$\hat{\rho} = 0.5721$					
	$\hat{\omega} = 0.0857$					
IZ-BAR(1)	$\hat{p} = .04460$	1341.0390	1352.2776	1.3430	–	–
	$\hat{\rho} = 0.1075$					
	$\hat{\phi} = 0.0447$					
ZI-BAR(1)	$\hat{\alpha} = 0.6762$	1311.6871	1326.6719	1.6390	–	–
	$\hat{\beta} = 0.5768$					
	$\hat{\omega}_\alpha = 0.1139$					
	$\hat{\omega}_\beta = 0.1731$					
ZT-BAR(1)	$\hat{\alpha} = 0.5987$	1382.3806	1393.6192	–	–	–
	$\hat{\beta} = 0.5006$					
	$\hat{\beta}_0 = 0.3619$					

Table 7 Properties of the standardized Pearson residuals for rainy-days counts in Germany

Model	Standardized Pearson residuals			
	Mean	Variance	$\hat{\rho}_{e_t}(1)$	$\hat{\rho}_{e_t^2}(1)$
ZOIBAR(1)	0.0043	1.2832	0.1294	0.0673
BAR(1)	-0.0027	2.0917	0.0458	0.0721
ZR-BAR(1)	0.0037	1.5721	0.1170	0.0813
IZ-BAR(1)	-0.0007	1.4243	0.0722	0.0597
ZIB-AR(1)	0.0120	0.4451	0.0447	-0.1464

If the model is correctly specified, then the residuals should have zero mean, unit variance, and no significant serial correlation in e_t and e_t^2. Table 7 shows the mean, variance, ACF(1) of e_t ($\hat{\rho}_{e_t}$) and ACF(1) of e_t^2 ($\hat{\rho}_{e_t^2}$). The above models all have satisfactory performances when we consider the mean of the residuals. Based on the variance of the residuals, although all alternative models give bigger variance values than 1, the ZOIBAR(1) process gives the significantly better result than the other models. So we conclude that the standardized Pearson residual supports the above conclusion that the ZOIBAR(1) process is the most suitable one to fit this data set. One may ask if there is correlation within the residuals since $\hat{\rho}_{e_t} = 0.1294$, which is not close to 0. For this, Ljung–Box Q-test for the residuals and square of residuals

shows that p values are equal to 0.3993 and 0.4077 based on 15 lags at the 5% level. This situation suggests that there is no significant autocorrelation in the residuals and square of residuals.

As pointed out by a referee, it is an important topic to discuss the actual meaning of the parameter estimates for the ZOIBAR(1) model. In this case, on the one hand, the hidden process $\{X_t\}$ in Eq. (2.1) can be viewed as the number of rainy-days per week without special circumstances. $\hat{p} = 0.6054$ shows that the probability of rain for each day is 0.6054, which means that this area is rainy. $\hat{\rho} = 0.0977$ indicates that the correlations of the number of rainy-days per week between two consecutive weeks are weak, which reveals that the local weather is changeable and difficult to predict. On the other hand, the observed process $\{Y_t\}$ in Eq. (2.1) is the real data of the number of rainy-days per week. The difference between X_t and Y_t is that X_t is masked by the zero or one with probability $\hat{\phi}_0 + \hat{\phi}_1 = 0.1139$ due to some unusual changes in the climate and these unusual changes can lead to the less rainfall. $\hat{\phi}_0 = 0.0463$ and $\hat{\phi}_1 = 0.0676$ represent the probabilities of the zero and one times rainy-days per week caused by unusual changes in the climate.

5.2 Rainy-days counts in Sweden

The second data set is the number of rainy-days per week at Stockholm in Sweden. The data were obtained from the ECA&D website, http://www.ecad.eu and were collected from January 1st, 2000 till March 21th, 2008 by the German Weather Service (DWD ="Deutscher WetterDienst", http://www.dwd.de/), where weeks are defined from Saturday to Friday. The length of the data is 480 and the fixed upper limit is $n = 7$. The sample path, histogram, ACF, PACF and summary statistics of the observations are given in Figure S4 and Table S9. The binomial dispersion index I_d of the data is 1.8746, which indicates that the data set is overdispersed. The sample binomial zero index and one index are, respectively, 3.3525 and 1.7472, which seem that this data set is zero-and-one inflated.

As can be seen from Table S10, the BAR(1) and ZR-BAR(1) processes are not suitable for the data since they give the bigger AIC and BIC values than other models. Moreover, the ZR-BAR(1) process overestimates the degree of zero inflation and gives the wrong information that the data set is one-deflated. The IZ-BAR(1), ZIB-AR(1) and ZT-BAR(1) models have the similar performances when we consider AIC and BIC. The IZ-BAR(1) and ZIB-AR(1) models can also capture the overdispersion, but the unobtained marginal distributions constrain their applications in explaining zero inflation and one inflation. Similar to the first data set, the ZOIBAR(1) process still can accurately capture zero-inflation, one-inflation and overdispersion characteristics and gives the smallest values of AIC and BIC. The Pearson residual analysis also gives the information that the ZOIBAR(1) process is the most appropriate model to fit this data set (see Table S11).

The discussion about the actual meaning of the parameter estimates for the ZOIBAR(1) model can be proceeded in the similar way as we did in Sect. 5.1. Again, the hidden process $\{X_t\}$ is the number of rainy-days per week without special circumstances. $\hat{p} = 0.5551$ shows that the probability of rain for each day is 0.5551, which

means that this area is also rainy. $\hat{\rho} = 0.2493$ indicates that the correlations of the number of rainy-days per week between two consecutive weeks are significant, which implies that the local weather is easier to predict. We observe the real data of the number of rainy-days per week $\{Y_t\}$ by masking $\{X_t\}$ with zero or one and the masking probability is $\hat{\phi}_0 + \hat{\phi}_1 = 0.1873$, which reveals that the unusual climate changes in Sweden are more commonly encountered than in Germany.

5.3 Assaults-on-officers counts

The third data set is collected from the file PghCarBeat.csv, which was downloaded from The Forecasting Principles site (http://www.forecastingprinciples.com). The data set is collected for 42 different car beats and from January 1990 to June 2001. For each month t, the value y_t counts the number of car beats reported at least one case of assaults-on-officers. Hence, the considered data set has finite range with fixed upper limit $n = 42$ and the series contains 138 observations. The sample path, histogram, ACF, PACF and summary statistics of the observations are given in Figure S5 and Table S12. The binomial dispersion index I_d of the data is 1.1490, which indicates that the data set is overdispersed. The sample binomial zero index and one index are, respectively, 0.9507 and 1.2008, which indicate that the considered data set is zero-deflated but one-inflated.

The performances of the alternative models are reported in Table S13. Based on AIC and BIC, the models are competitive and the BAR(1) process yields the smallest BIC value. The reason is that BIC gives more severe penalty for the number of model parameters. It can be observed that the parameter ω for the ZR-BAR(1) model equals to 0, which indicates that the ZR-BAR(1) model reduces to the BAR(1) model in this case. This is also well understood since this data set is zero-deflated, and thus the zero-inflated binomial marginal distribution is just the opposite to what one wishes. Except for the ZOIBAR(1) model, all the alternatives fail to capture zero deflation and one inflation. Again, the ZOIBAR(1) model successfully capture zero deflation, one inflation and overdispersion and the corresponding indices are prettily close to the empirical values. During our study, we also tried to apply the ZIB-AR(1) model to fit this data set but the workable results are unavailable. Thus, the relevant information is not reported here. The Pearson residual analysis also supports the ZOIBAR(1) model since its variance value of residuals is the closest to 1 among the alternatives. One may ask if there is correlation within the square of residuals since $\hat{\rho}_{e_t^2} = -0.1218$, which is not close to 0. For this, Ljung–Box Q-test for the residuals and square of residuals shows that p values are equal to 0.2097 and 0.9796 based on 15 lags at the 5% level. This situation suggests that there is no significant autocorrelation in the residuals and square of residuals.

Now we turn to discuss the actual meaning of the parameter estimates for the ZOIBAR(1) model in the assaults-on-officers case. On the one hand, the hidden process $\{X_t\}$ is the number of car beats reported at least one case of assaults-on-officers per month during the period of normal state. $\hat{p} = 0.0484$ shows that the probability of the appearance of assaults-on-officers for each car beat in a month is 0.0484, which is close to zero and reflects that this area is in a state of harmony and stability. $\hat{\rho} = 0.3013$

indicates that the correlations of the number of car beats reported at least one case of assaults-on-officers per month between two consecutive months are visibly significant. This phenomenon can be interpreted by the fact that the improvement or deterioration of social order changes gradually over time. Hence, the assaults-on-officers cases in a month will certainly affect the assaults-on-officers cases in next month. On the other hand, the observed process $\{Y_t\}$ is the real data of the number of car beats reported at least one case of assaults-on-officers per month. The difference between X_t and Y_t is that X_t is masked by the zero or one with probability $\hat{\phi}_0 + \hat{\phi}_1 = 0.4116$ due to the government management measures and these measures can lead to the social stability. Based on $\hat{\phi}_0 + \hat{\phi}_1 = 0.4116$, we can guess that the local government attached great importance to the security situation and the corresponding measures were frequently adopted. $\hat{\phi}_0 = 0.1606$ and $\hat{\phi}_1 = 0.2510$ represent the probabilities of the zero and one car beat reported at least one case of assaults-on-officers under the circumstances of government management.

5.4 Prediction

Prediction problem is always a popular issue in time series analysis to check the adequacy and predictability of the selected model. Three different prediction methods will be discussed in the following:

First, we discuss the classical method for predicting the model. Conditional expectation method is the most common approach for constructing classical prediction in time series models due to its optimal property of the mean squared error. The h-step ahead predictor for Y_{t+h} is $\hat{Y}_{t+h}|Y_t = E(Y_{t+h}|Y_t)$, where $E(Y_{t+h}|Y_t)$ is given by Eq. (2.2). In practice, the parameters p, ρ, ϕ_0 and ϕ_1 are replaced by their corresponding QMLE.

However, the conditional expectation method goes against the data coherence and prediction coherence since the integer-valued predictors can hardly be produced. Therefore, we call for an adaptive procedure which has the ability to produce integer-valued predictors. Freeland and McCabe (2004) proposed a feasible approach which uses the h-step-ahead predictive conditional distributions to predict the future value. One can obtain the point prediction from the median or the mode of the predictive distribution. This method has been employed by some researchers, see Möller et al. (2016), Maiti and Biswas (2017) and Kang et al. (2021), among others. Following this path, we compute the h-step-ahead conditional distribution of Y_{t+h} given $Y_1 = y_1, \ldots, Y_t = y_t$, i.e., $P(Y_{t+h} = y_{t+h}|Y_1 = y_1, \ldots, Y_t = y_t)$ of the ZOIBAR(1) process based on the HMM theory (see Zucchini et al. 2009, Section 5.3). Motivated by the idea of the quasi-likelihood estimation, another alternative way is to calculate the h-step-ahead quasi-conditional distribution, i.e., $P(Y_{t+h} = y_{t+h}|Y_t = y_t)$ based on Proposition 2.

The Bayesian prediction method is another approach to produce coherent predictors. The future value is derived via using the h-step-ahead Bayesian predictive probability function, which is based on the assumption that both Y_{t+h} and unknown parameters are random and given by

Table 8 The MADEs of the h-step-ahead predictor for rainy-days counts in Germany

Predict step	Conditional expectation	Quasi-conditional distribution	Conditional distribution	Bayesian prediction
$h = 1$	0.8934	1.0000	1.0000	1.0000
$h = 2$	1.8303	2.0000	2.0000	2.0000
$h = 3$	2.1764	2.0000	2.0000	2.0000
$h = 4$	0.8229	1.0000	1.0000	1.0000
$h = 5$	1.1772	1.0000	1.0000	1.0000
$h = 6$	1.1772	1.0000	1.0000	1.0000

$$
P(Y_{t+h} = y_{t+h}|\mathbf{Y})
$$
$$
= \int_0^1 \int_c^1 \int_0^1 \int_0^{1-\phi_1} P(Y_{t+h} = y_{t+h}|\mathbf{Y}, \boldsymbol{\lambda}) \boldsymbol{\pi}(\boldsymbol{\lambda}|\mathbf{Y}) \mathrm{d}\phi_0 \mathrm{d}\phi_1 \mathrm{d}\rho \mathrm{d}p,
$$

where $y_{t+h} \in \{0, 1, \ldots, n\}$, $\mathbf{Y} = (Y_1, \ldots, Y_T)^\top$ and the posterior distribution $\boldsymbol{\pi}(\boldsymbol{\lambda}|\mathbf{Y})$ is given by Eq. (3.2). Also, we can obtain the point prediction from the median or the mode of the h-step-ahead Bayesian predictive probability function.

To compare the above four prediction approaches, we evaluate the performances in the above three data sets via MADE of out-of-sample prediction, where the last h observations are excluded when estimating parameters for out-of-sample prediction. The corresponding results are shown in Tables 8 and S15–S16 and h-step-ahead predictive quasi-conditional distribution, h-step-ahead predictive conditional distribution and h-step-ahead Bayesian predictive conditional distribution of the considered data sets are given in Figs. 3, 4, 5 and S6–S11. We can see that the h-step-ahead predictive conditional distribution is similar to the h-step-ahead Bayesian predictive conditional distribution, so the predictors produced by the two methods are close. This phenomenon is well understood since the only difference between predictive conditional distribution and Bayesian predictive conditional distribution is the priors. It is also easily observed that the conditional expectation prediction can produce more stable results and conditional distribution prediction and Bayesian prediction succeed in giving better predictors than the conditional expectation prediction in some situations. In general, we are more inclined to conditional distribution prediction and Bayesian prediction. The reason is that these two methods not only can give us reliable predictors, but also have the ability to provide more information such as the predictive distribution. The predictive distribution can help practitioners establish early warning system of risk via utilizing the idea of Value-at-Risk (VaR, see Chen and Watanabe 2019). Take assaults-on-officers counts as an example, from Figures S9–S11, it is a small probability event that the observation is larger than 2. Hence, if we once observe the number of car beats reported at least one case of assaults-on-officers is larger than 2, it can be deduced that some events have caused unrest in social security and the corresponding treatment measure should be adopted. The similar discussion can be proceeded in the rainy-days counts case.
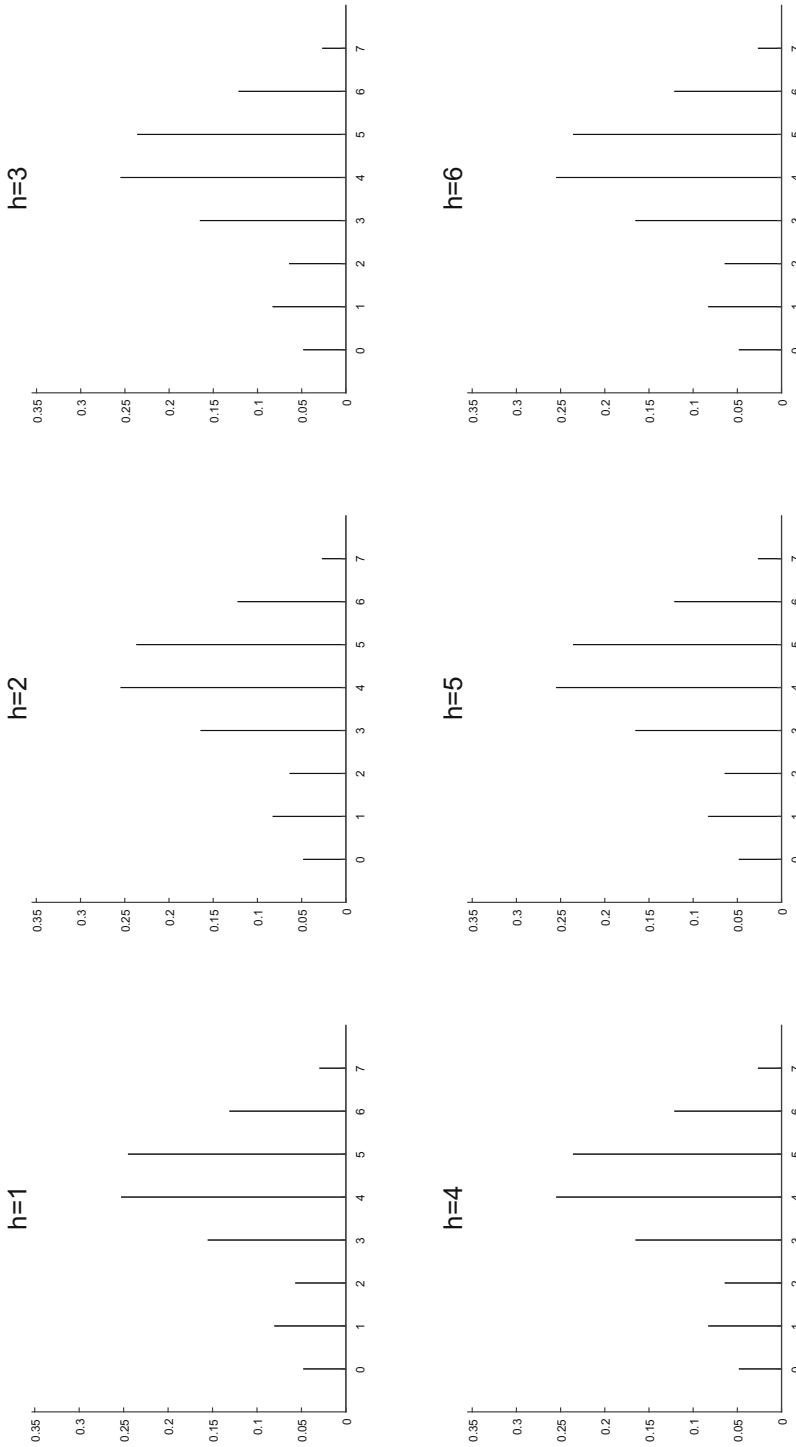
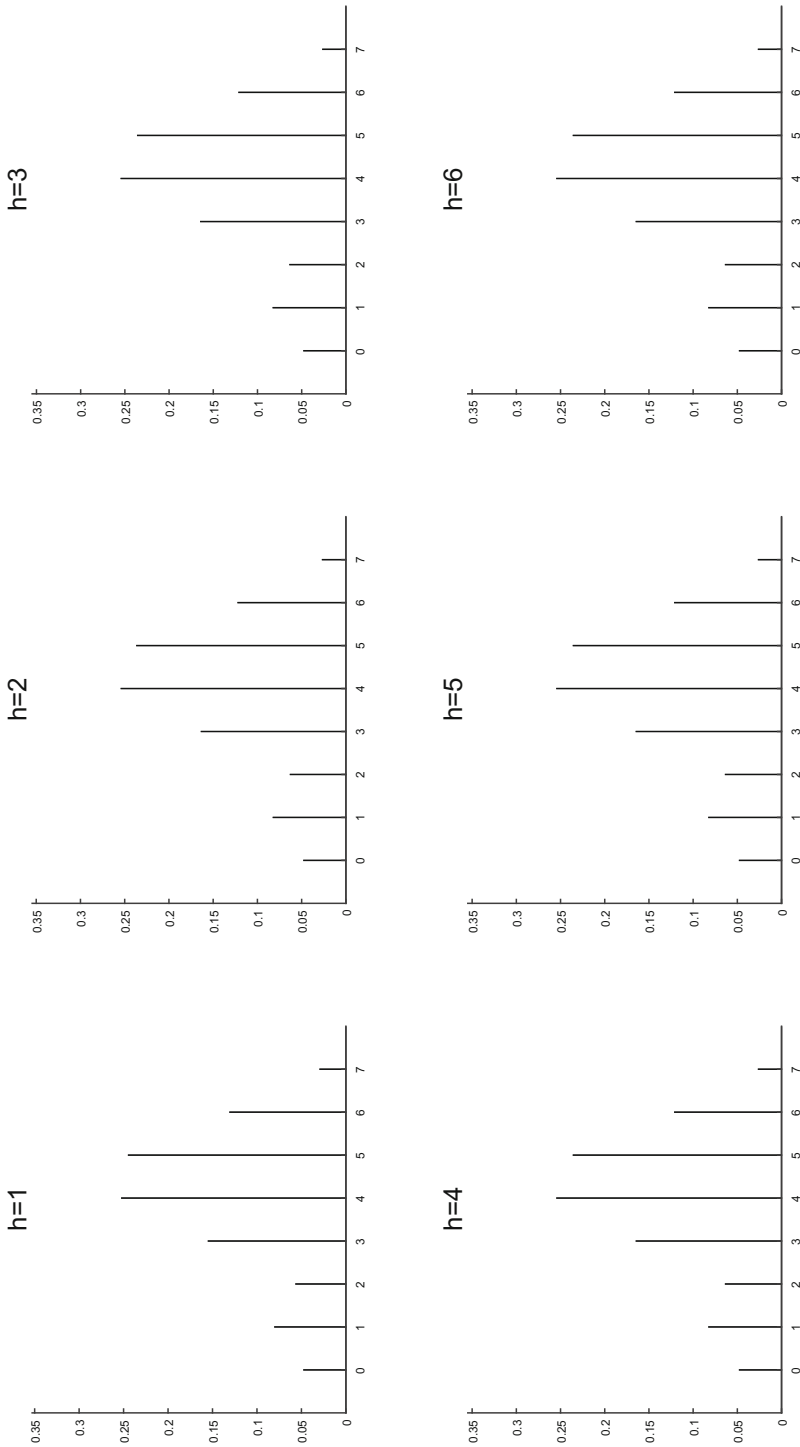**Fig. 3** *h*-step-ahead predictive quasi-conditional distribution of rainy-days counts in Germany

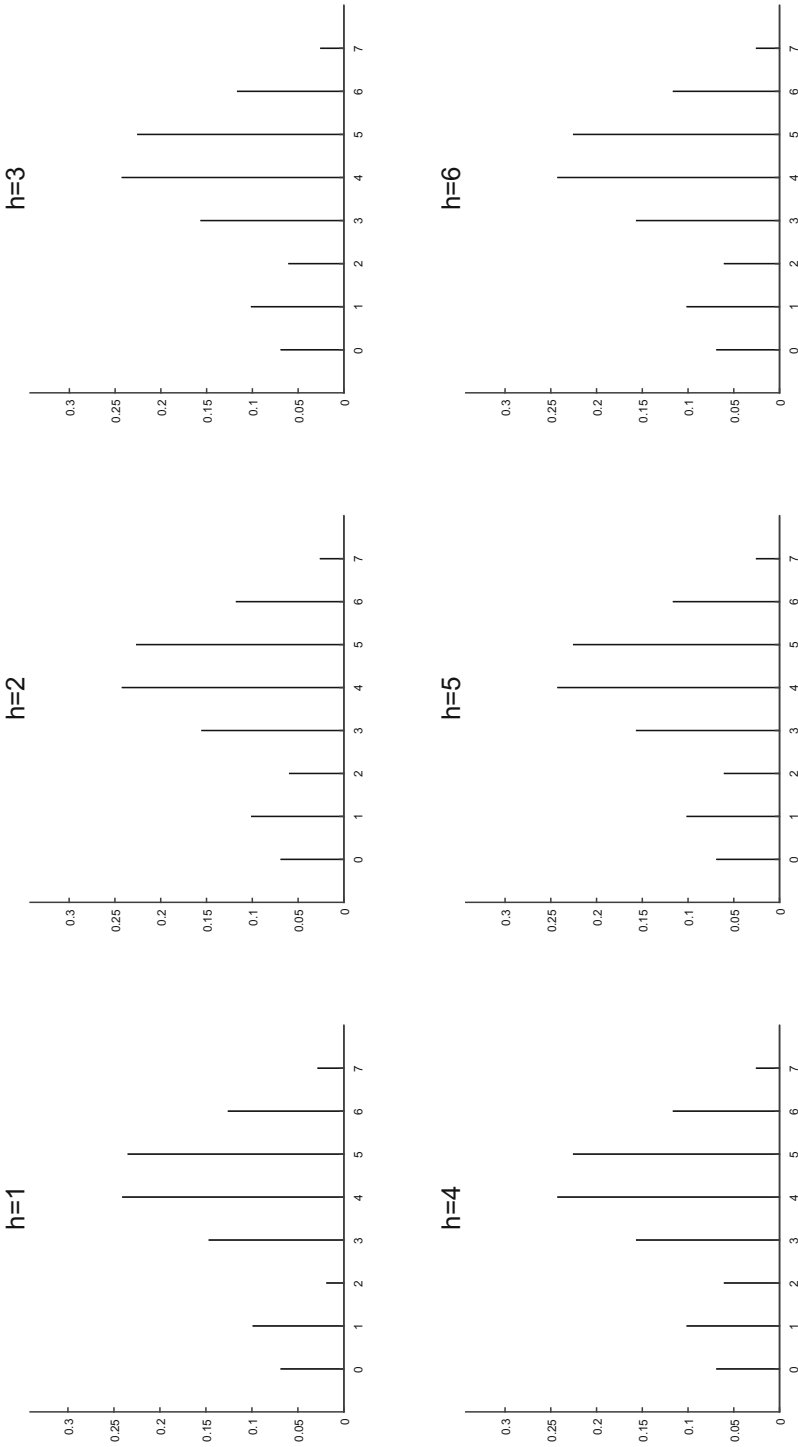**Fig. 4** *h*-step-ahead predictive conditional distribution of rainy-days counts in Germany

Fig. 5 *h*-step-ahead Bayesian predictive conditional distribution of rainy–days counts in Germany

6 Conclusion

This article proposes a hidden Markov model with zero-and-one inflated binomial marginals to better analyze bounded counts with excess zeros and ones. The stochastic properties of the new model are investigated and estimators of the model parameters are derived by the probability-based, quasi-maximum likelihood, maximum likelihood and Bayesian approaches. A binomial one-inflation index is constructed and further utilized to develop a method to test whether zero-and-one inflated with respect to a BAR(1) model. Three applications to the real-data examples are given to assess the performance of our model.

However, more research is still needed for some aspects of the new model. The first issue is that the estimation problem for the ZOIBAR(1) model should be treated in more detail. For example, it would be interesting in applying the empirical likelihood approach to the ZOIBAR(1) model and investigating the asymptotic behavior of the estimators. Extensions to the ZOIBAR(1) model are the second issue. The construction of high-order and multivariate ZOIBAR processes may be an interesting topic. The third issue is that modeling for bounded counts with excess zeros and ones is still rare and the corresponding Markov models are urgently needed. Relative research will be devoted as future work.

References

Box GEP, Jenkins GM, Reinsel GC (1994) Time Series Analysis-Forecasting and Control. Prentice Hall, New Jersey

Chen CWS, Watanabe T (2019) Bayesian modeling and forecasting of Value-at-Risk via threshold realized volatility. Appl Stoch Model Bus Ind 35:747–765

Chen H, Li Q, Zhu F (2020) Two classes of dynamic binomial integer-valued ARCH models. Braz J Probab Stat 34:685–711

Chen H, Li Q, Zhu F (2021) Binomial AR(1) processes with innovational outliers. Commun Stat Theory Methods 50:446–472

Cui Y, Lund R (2010) Inference in binomial AR(1) models. Stat Probab Lett 80:1985–1990

Freeland RK, McCabe BPM (2004) Forecasting discrete valued low count time series. Int J Forecast 20:427–434

Jazi MA, Jones G, Lai CD (2012) First-order integer valued AR processes with zero-inflated poisson innovations. J Time Ser Anal 33:954–963

Kang Y, Wang D, Yang K (2020) Extended binomial AR(1) processes with generalized binomial thinning operator. Commun Stat Theory Methods 49:3498–3520

Kang Y, Wang D, Yang K (2021) A new INAR(1) process with bounded support for counts showing equidispersion, underdispersion and overdispersion. Stat Pap 62:745–767

Kim HY, Weiß CH (2015) Goodness-of-fit tests for binomial AR(1) processes. Statistics 49:291–315

Kim HY, Weiß CH, Möller TA (2018) Testing for an excessive number of zeros in time series of bounded counts. Stat Methods Appl 27:689–714

Kumaraswamy P (1980) A generalized probability density function for double bounded random processes. J Hydrol 46:79–88

Li C, Wang D, Zhang H (2015) First-order mixed integer-valued autoregressive processes with zero-inflated generalized power series innovations. J Korean Stat Soc 44:232–246

Liu M, Zhu F, Zhu K (2022) Modeling normalcy-dominant ordinal time series: an application to air quality level. J Time Ser Anal 43:460–478

Maiti R, Biswas A (2017) Coherent forecasting for stationary time series of discrete data. AStA Adv Stat Anal 99:337–365

Maiti R, Biswas A, Chakraborty B (2018) Modelling of low count heavy tailed time series data consisting large number of zeros and ones. Stat Methods Appl 27:407–435

McKenzie E (1985) Some simple models for discrete variate time series. Water Resour Bull 21:645–650

Mohammadi Z, Sajjadnia Z, Bakouch HS, Sharafi M (2022) Zero-and-one inflated Poisson-Lindley INAR (1) process for modelling count time series with extra zeros and ones. J Stat Comput Simul 92:2018–2040

Möller TA, Silva ME, Weiß CH, Scotto MG, Pereira I (2016) Self-exciting threshold binomial autoregressive processes. AStA Adv Stat Anal 100:369–400

Möller TA, Weiß CH, Kim HY (2020) Modelling counts with state-dependent zero inflation. Stat Model 20:127–147

Möller TA, Weiß CH, Kim HY, Sirchenko A (2018) Modeling zero inflation in count data time series with bounded support. Methodol Comput Appl Probab 20:589–609

Perumean-Chaney SE, Morgan C, McDowall D, Aban I (2013) Zero-inflated and overdispersed: what's one to do? J Stat Comput Simul 83:1671–1683

Qi X, Li Q, Zhu F (2019) Modelling time series of count with excess zeros and ones based on INAR(1) model with zero-and-one inflated poisson innovations. J Comput Appl Math 346:572–590

Ridout M, Demétrio CGB, Hinde J (1998) Models for count data with many zeros. In: Proceedings of the 19th international biometric conference. Cape Town, South Africa, pp 179–190

Ristić MM, Popović B (2019) A new bivariate binomial time series model. Markov Process Relat Fields 25:1–26

Schweer S (2015) On the time-reversibility of integer-valued autoregressive processes of general order. In: Steland A et al (eds) Stochastic models, statistics and their applications. Springer, Berlin, pp 169–177

Scotto MG, Weiß CH, Silva ME, Pereira I (2014) Bivariate binomial autoregressive models. J Multivar Anal 125:233–251

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

Weiß CH (2009) Jumps in binomial AR(1) processes. Stat Probab Lett 79:2012–2019

Weiß CH (2009) Monitoring correlated processes with binomial marginals. J Appl Stat 36:399–414

Weiß CH, Kim HY (2013) Binomial AR(1) processes: moments, cumulants, and estimation. Statistics 47:494–510

Weiß CH, Kim HY (2013) Parameter estimation for binomial AR(1) models with applications in finance and industry. Stat Pap 54:563–590

Weiß CH, Kim HY (2015) Diagnosing and modeling extra-binomial variation for time-dependent counts. Appl Stoch Model Bus Ind 30:588–608

Weiß CH, Pollett PK (2012) Chain binomial models and binomial autoregressive processes. Biometrics 68:815–824

Weiß (2013) Integer-valued autoregressive models for counts showing underdispersion. J Appl Stat 40:1931–1948

Weiß CH, Pollett PK (2014) Binomial autoregressive processes with density-dependent thinning. J Time Ser Anal 35:115–132

Yang K, Wang D, Li H (2017) Threshold autoregression analysis for finite range time series of counts with an application on measles data. J Stat Comput Simul 88:597–614

Zhu F (2012) Zero-inflated Poisson and negative binomial integer-valued GARCH models. J Stat Plan Inference 142:826–839

Zucchini W, MacDonald IL, Langrock R (2009) Hidden Markov models for time series: an introduction using R. Chapman & Hall, London

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009) Mixed effects models and extensions in ecology with R. Springer, New York