Journal of the Korean Statistical Society (2024) 53:815–843
https://doi.org/10.1007/s42952-024-00269-0

Online ISSN 2005-2863
Print ISSN 1226-3192

RESEARCH ARTICLE

Multiple values-inflated bivariate INAR time series of counts: featuring zero–one inflated Poisson-Lindly case

Sangyeol Lee[1] · **Minyoung Jo**[1]

Received: 25 September 2023 / Accepted: 24 April 2024 / Published online: 1 June 2024

Abstract

This study considers multiple values-inflated bivariate integer-valued autoregressive (MV-inflated BINAR) models. It develops the inferential procedures for parameter estimation on this model, which apply to constructing a change point test and outlier detection rule. We first introduce the MV-inflated BINAR model with one parameter exponential family and Poisson-Lindley innovations. Then, we propose a quasi-maximum likelihood estimator (QMLE) and divergence-based estimator featuring minimum density power divergence estimator (MDPDE) for robust estimation. To evaluate the performance of these estimators, we conduct Monte Carlo simulations and demonstrate the adequacy of MDPDE in zero–one inflated models. Real data analysis is also carried out using the number of monthly earthquake cases in the United States.

Keywords MV-inflated BINAR model · Zero-one inflated time series of counts · Poisson-Lindley distribution · Divergence-based estimation · Minimum density power divergence estimator

1 Introduction

In this study, we introduce a multiple values-inflated bivariate integer-valued autoregressive (MV-inflated BINAR) model and investigate the asymptotic behavior of the quasi-maximum likelihood estimator (QMLE) and the divergence-based estimator for robust estimation that includes minimum density power divergence estimator (MDPDE). Time series of counts appear in diverse research fields related to social, physical, medical, and engineering sciences. Some specific examples are a weekly number of airline passengers, yearly deaths from lung diseases, and the volume of stocks transacted over a finite minute period. To model these kinds of time series, McKenzie (1985) and Al-Osh and Alzaid (1987) coined the integer-valued

✉ Sangyeol Lee
sylee@stats.snu.ac.kr

[1] Department of Statistics, Seoul National University, Seoul 08826, South Korea

autoregressive (INAR) models. As with the INAR models, the integer-valued generalized autoregressive conditional heteroscedastic (INGARCH) model was also proposed by Ferland et al. (2006) and Fokianos et al. (2009). For the past decades, these two models have gained great popularity among researchers in the integer-valued time series context. We refer to Weiß (2018) for a general overview.

Overdispersion is a frequently-observed phenomenon in time series of counts, and one of the reasons is zero inflation due to a high frequency of no incidences as seen in the manufacturing process, epidemiology, and insurance. When this phenomenon is not appropriately treated, it can cause a significant bias in inferences. Thus, to circumvent this, zero-inflated models have been taken into account on research problems, refer to Lambert (1992), Chen et al. (2019), and Jo and Lee (2023). Subsequently, as a task of extension of the zero-inflated models, multiple values-inflated INGARCH models have been developed to model time series of counts with excessive zeros and other values, refer to Lee and Kim (2023), who considered INGARCH models whose conditional probability mass function is a multiple values (MV)-inflated one parameter exponential family. In a similar spirit, here we extend the zero-inflated INAR models to more general multiple values-inflated BINAR models. Compared to the MV-inflated INGARCH model of Lee and Kim (2023), its INAR counterpart has the merits to have a simpler model structure and requires much less technically demanding conditions in deriving the asymptotic properties of parameter estimators, which renders the MV-inflated BINAR models to be more tractable and preferential in practical usages.

Historically, Poisson innovations have been utilized in modeling with an INAR scheme by many researchers because of its simplicity in applications, refer to Pedeli and Karlis (2013), Weiß (2015), and Lu and Wang (2022). However, the Poisson distribution is not always suitable for modeling as its mean and variance concur with each other and this property is not occasionally acceptable in real situations. Several alternatives to the Poisson innovations have been proposed in the literature, for example, geometric innovations (Jazi et al. 2012), negative binomial innovations (Pedeli and Karlis 2011), Poisson-weighted exponential innovations (Altun 2020), and Poisson-Lindley innovations (Mohammadpour et al. 2018; Livio et al. 2018). In particular, the Poisson-Lindley distribution belongs to the compound Poisson family and has good properties like unimodality and infinite divisibility. Also, unlike the Poisson distribution, the Poisson-Lindley distribution can capture well the overdispersion of datasets.

For parameter estimation for the proposed MV-inflated BINAR models, we consider the quasi-maximum likelihood estimator (QMLE) and the Brègman divergence-based estimator featuring minimum power density divergence estimator (MDPDE), and substantiate their consistency and asymptotic normality under regularity conditions. These estimators are applicable to identifying change points and outliers in time series of counts, as seen in Lee and Jo (2023b), who used MDPDE in handling those. Developing change point tests and outlier detection methods are critical issues in time series analysis as ignoring those may lead to false conclusions. A brief summary of those methods is provided for easy access to our empirical analysis. The change point tests based on the scores and residuals were established for AR and GARCH models (Gombay (2008), Berkes et al. (2004), Lee (2020)), as well as for INAR and INGARCH models (Lee and Lee 2019; Lee and Jo 2023a). We refer to Csörgő & Horváth, (1997) and Lee et al. (2003) for an overview. For assessing the performance of QMLE and MDPDE, we conduct Monte Carlo simulations and confirm the functionality of MDPDE in the presence of outliers. Real data analysis is performed for illustration using the monthly earthquake datasets in the United States, which validates our methods.

The remainder of this paper is organized as follows. Section 2 introduces the bivariate multiple values-inflated INAR model. Section 3 establishes the asymptotic results of QMLE and MDPDE for the proposed model. Sections 3.2 and 5 conduct a simulation study and real data analysis to demonstrate the validity of our methods. Finally, Sect. 6 provides the concluding remarks.

2 MV-inflated BINAR(1) model

Let $\{Y_t\}$ be a bivariate time series of counts following the bivariate INAR(1) (BINAR(1)) model:

$$\begin{aligned} Y_t = \begin{pmatrix} Y_{t1} \\ Y_{t2} \end{pmatrix} &= \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \circ \begin{pmatrix} Y_{t-1,1} \\ Y_{t-1,2} \end{pmatrix} + \begin{pmatrix} Z_{t1} \\ Z_{t2} \end{pmatrix} \\ &= \begin{pmatrix} \alpha_{11} \circ Y_{t-1,1} + \alpha_{12} \circ Y_{t-1,2} + Z_{t1} \\ \alpha_{21} \circ Y_{t-1,1} + \alpha_{22} \circ Y_{t-1,2} + Z_{t2} \end{pmatrix}, \end{aligned} \tag{2.1}$$

wherein $\alpha_{ij} \circ Y_{t-1,i} = \sum_{k=1}^{Y_{t-1,i}} w_{t-1,ijk}$ for $i,j = 1, 2$, $\{w_{tijk}\}$ are mutually independent iid Bernoulli random variables with success probability $\alpha_{ij} \in [0, 1)$, and $Z_t = (Z_{t1}, Z_{t2})^T$ are iid nonnegative integer-valued random vectors independent of all w_{tij} with $E\|Z_t\|^2 < \infty$. Here, $\|\cdot\|$ denotes the L_2-norm for vectors and matrices. We rewrite (2.1) as

$$Y_t = A \circ Y_{t-1} + Z_t, \tag{2.2}$$

with

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}.$$

Darolles et al. (2019) and Lee and Jo (2023a) verified that the BINAR(1) process in (2.1) has a unique strictly stationary and ergodic solution under certain conditions, such as $E||Z_1|| < \infty$ and

$$(1 - \alpha_{11})(1 - \alpha_{22}) - \alpha_{12}\alpha_{21} > 0. \tag{2.3}$$

Moreover, when (2.3) holds, Lee and Jo (2023a) showed that $E\|Z_1\|^r < \infty$ for some $r \geq 2$ if and only if $\|Y_1\|^r < \infty$. Also, Y_t is shown to be adapted to all the random variables comprising Y_s, $s \leq t$, which particularly indicates that Y_s, $s \leq t$, are independent of all Z_s, $s > t$.

As an estimate of A, Lee and Jo (2023a) employed conditional least squares estimator (CLSE), modified quasi-likelihood estimator (MQLE), and exponential family quasi-likelihood estimator (EQLE). However, instead of these estimators, we consider the quasi-maximum likelihood estimator (QMLE) as QMLE is in general more efficient than those, and the minimum density power divergence estimator (MDPDE) in Lee and Jo (2023b) as MDPDE is robust against possible outliers and model misspecifications. Results on QMLE and MDPDE will be appended to specify their consistency and asymptotic normality.

In what follows, $\{Y_t\}$ is a strictly stationary and ergodic process following Model (2.1) with $E||Z_1||^4 < \infty$. Also, we assume that the support of Z_t includes (i, j) with $i, j = 0, 1$. We set the conditional mean equation as

$$X_t(\theta) = E(Y_t|Y_{t-1}) = AY_{t-1} + \lambda, \tag{2.4}$$

where A is the matrix in (2.3) and $\lambda = (\lambda_1, \lambda_2)^T = EZ_t$. This equation enables practitioners to construct the CLSE of A and λ. However, CLSE is not so inefficient as QMLE, and in the presence of multiple inflation parameters, it is not feasible to construct, see Remark 1 below.

Modeling MV-inflated BINAR(1) time series merely requires a parametric modeling procedure for the innovational distributions in (2.1). For this task, one can consider Z_{ti} whose underlying distribution belongs to the MV-inflated one-parameter exponential family $\{f_{\varrho_i,\psi_i} : 0 < \phi_i := \sum_{j=1}^{M_i} \rho_{ij} < 1, \psi_i > 0\}$, $\varrho_i = (\rho_{i1}, \dots, \rho_{iM_i})^T$, with pmf:

$$f_{\varrho_i,\psi_i}(y) = \sum_{j=1}^{M_i} \rho_{ij} I(y = j-1) + \big(1 - \sum_{j=1}^{M_i} \rho_{ij}\big) f_{\psi_i}(y), \quad y \geq 0, \tag{2.5}$$

where M_i is a natural number and f_{ψ_i} is a discrete pmf with parameter $\psi_i > 0$. The number M_i denotes the number of the inflated values; for example, it is one if $\{Y_t\}$ is zero-inflated. For f_{ψ_i}, one can employ one-parameter exponential family distribution with pmf:

$$f_{\psi_i}(y) = \exp\{\eta_i y - \mathcal{A}_i(\eta_i)\} h_i(y), \quad y \geq 0. \tag{2.6}$$

wherein η_i is the natural parameter, $\mathcal{A}_i$ and h_i are known functions, both $\mathcal{A}_i$ and $\mathcal{B}_i := \mathcal{A}_i'$ are strictly increasing, where $\psi_i = \mathcal{B}_i(\eta_i)$ is the mean of the distribution. Also, one can alternatively use the Poisson-Lindley distribution with pmf:

$$f_{\psi_i}(y) = \frac{\psi_i^2(\psi_i + 2 + y)}{(1 + \psi_i)^{y+3}}, \quad y \geq 0. \tag{2.7}$$

Poisson-Lindley distribution merits to accommodate various types of distributions due to its skewness and kurtosis features, well describing all under-, equi-, and over-dispersion phenomena of time series, and is well known to have more desirable properties over Poisson and negative binomial distributions. For the basic properties of Poisson-Lindley INAR(1) models, we refer to Mohammadi et al. (2022).

In what follows, we assume that $\{f_{\psi_i}\}$ is either the one in (2.6) or (2.7). Furthermore, we assume that f_{ϱ_i,ψ_i} in (2.5) satisfies the identifiability condition: $f_{\varrho_i,\psi_i} = f_{\varrho_i',\psi_i'}$ implicates $\psi_i = \psi_i'$ and $\varrho_i = \varrho_i'$. Provided $\phi_i = \sum_{j=1}^{M_i} \rho_{ij} < 1$, this condition is fulfilled for frequently used distributions, for example, Poisson, negative binomial, and Poisson-Lindley distributions. Notice that the probability generating function (pgf) of the last distribution with parameter ψ is $\frac{\psi^2(2+\psi-s)}{(1+\psi)(1+\psi-s)^2}$. More generally, this identifiability condition holds if the pgf $\pi(s|\psi_i)$ of $f_{\psi_i}(s)$, $s \in (0,1]$, is (M_i+1) times differentiable and if $g(s|\psi_i) := \frac{\pi^{(M_i+1)}(s|\psi_i)}{\pi^{(M_i)}(s|\psi_i)}$ satisfies that $g(s|\psi_i) = g(s|\psi_i')$ for all $s \in (0,1]$ implies $\psi_i = \psi_i'$, because this will ultimately lead to $\phi = \phi_i'$ and so $\varrho_i = \varrho_i'$. See Proposition 2.1 of Lee and Kim (2023).

Herein, we do not intend to model the joint distribution of Z_t, only focusing on its components' marginal distributions as our inferential procedure will rely on a quasi-likelihood method comprising the sum of the two marginal likelihoods, as described below. This approach has been taken by Lee and Jo (2023a) and Lee et al. (2023) and has proven to remarkably lessen the efforts invested for modeling and inferences.

3 Inference for MV-inflated BINAR(1) model

3.1 QMLE

We set $\theta = (\alpha^T, \varrho^T, \psi^T)^T$ with $\alpha = (\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})^T$, $\varrho = (\varrho_1^T, \varrho_2^T)^T$, $\varrho_i = (\rho_{i1}, \dots, \rho_{iM_i})^T$, and $\psi = (\psi_1, \psi_2)^T$, and $f_{\theta,i}$ denotes the conditional pmf of Y_{ti} given Y_{t-1} when θ is a model parameter. We denote by $\theta_0 = (\alpha_0, \rho_0, \psi_0)^T = (\alpha_{110}, \alpha_{120}, \alpha_{210}, \alpha_{220}, \varrho_0^T, \psi_0^T)^T$ the true parameter, which is assumed to be an interior of a compact subset Θ of $\mathbb{R}^{M_1+M_2+6}$, wherein

$$\begin{aligned}\Theta = \big\{\theta{:}\kappa_1 \le \alpha_{ij} \le 1-\kappa_1, \kappa_2 \le \rho_{ij} \le 1-\kappa_2,\\ \kappa_3 \le \phi_i \le 1-\kappa_3, \kappa_4 \le \psi_i \le \kappa_5 \text{ for all } i,j = 1,2\big\}\end{aligned}$$

for some constants $0 < \kappa_i < 1, i = 1, \dots, 4$, and $\kappa_5 > 1$.

The conditional pmf of Y_{ti}, $i = 1, 2$, is then expressed as

$$\begin{aligned} & f_{\theta,i}(y|Y_{t-1}) \\ & = P(Y_{ti} = y|Y_{t-1}) \\ & = \sum_{j=0}^{y} J(Y_{t-1,1}, Y_{t-1,2}, \alpha_{i1}, \alpha_{i2}, j) f_{\varrho_i,\psi_i}(y-j), \end{aligned} \tag{3.8}$$

where $J(m,n,p_1,p_2,k) = P(B_1(m,p_1) + B_2(n,p_2) = k)$, $k \ge 0$, and $B_1(m,p_1)$ and $B_2(n,p_2)$ two independent binomial random variables with success probabilities p_1 and p_2 $(B_i(0,p_i) = 0)$.

Note that the model identifiability holds for Model (2.1), due to the following. We rewrite (2.1) as $Y_t = A \circ Y_{t-1} + Z_t(\varrho, \psi)$, which becomes $Y_t = A_0 \circ Y_{t-1} + Z_t(\varrho_0, \psi_0)$ when θ equals θ_0. Note that if both θ and θ_0 generate the same $\{Y_t\}$, we get $f_{\theta,i}(y|Y_{t-1}) = f_{\theta_0,i}(y|Y_{t-1})$ for all $y \geq 0$, $i = 1, 2$, so that $X_t(\theta) = X_t(\theta_0)$ holds true, namely, $AY_{t-1} + \lambda = A_0 Y_{t-1} + \lambda_0$, owing to (2.4). This implicates $A = A_0$ and $\lambda = \lambda_0$ as Y_t can take values of (i, j) with $i, j = 0, 1$ owing to our assumption. This in turn implies $Z_t(\varrho, \psi) \stackrel{d}{=} Z_t(\varrho_0, \psi_0)$, which leads to $\varrho = \varrho_0$ and $\psi = \psi_0$, due to the model identifiability of f_{ϱ_i, ψ_i}, so that we finally have $\theta = \theta_0$.

QMLE is defined by

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmax}} \frac{1}{n} \sum_{t=1}^{n} \ell_t(\theta),$$

where

$$\ell_t(\theta) := \sum_{i=1}^{2} \ell_{it}(\theta) = \sum_{i=1}^{2} \log f_{\theta,i}(Y_{ti}|Y_{t-1}).$$

In Section 7, we verify the consistency and asymptotic normality of QMLE, as this verification was not explicitly addressed in the precedent studies even for the univariate cases. The following theorem states the consistency and asymptotic normality of QMLE under the regular conditions.

Theorem 1 *Suppose that* (2.3) *holds*, θ_0 *is an interior of a compact parameter space* Θ, *and* Z_t *satisfies* $E||Z_1||^4 < \infty$ *and has a pmf of the form in* (2.5) *with* f_{ψ_i} *in* (2.6) *with* $E|\log h(Y_{ti})| < \infty,\ i = 1, 2$, *or in* (2.7). *Then*, $\hat{\theta}_n \longrightarrow \theta_0$ *a.s. as* $n \to \infty$. *Moreover, if*

$$\mathcal{J}_0 = -E\left(\frac{\partial^2 \ell_t(\theta_0)}{\partial\theta\partial\theta^T}\right),\ \mathcal{K}_0 = E\left(\frac{\partial \ell_t(\theta_0)}{\partial\theta}\frac{\partial \ell_t(\theta_0)}{\partial\theta^T}\right),$$

and $\mathcal{J}_0$ *is non-singular, then, as* $n \to \infty$,

$$\sqrt{n}(\hat{\theta}_n - \theta_0) \stackrel{d}{\longrightarrow} N(0, \mathcal{J}_0^{-1}\mathcal{K}_0\mathcal{J}_0^{-1}).$$

3.2 Divergence-based robust estimation

As QMLE is sensitive to outliers, we consider a robust estimator based on density power divergence (DPD), as DPD scheme is well known to supply an excellent estimating method in practice. Herein, we consider the DPD scheme within the Brègman divergence paradigm, aiming at providing a more general theory. Brègman divergence is defined for two densities (or probability mass functions) g and h as follows:

$$D_{\varphi}(h,g)=\int\Big\{\varphi(h(y))-\varphi(g(y))-(h(y)-g(y))\varphi'(g(y))\Big\}dy, \tag{3.9}$$

where $\varphi:\mathbb{R}\to\mathbb{R}$ is a strictly convex function, named the divergence generator, as this divergence family accommodates Kullback–Leibler divergence, Brègman-exponential divergence (BED), proposed by Mukherjee et al. (2019), and DPD, which respectively correspond to $\varphi(y)=y\log y-y$, $2(e^{ay}-ay-1)/a^2$, $a\in\mathbb{R}$, and $(y^{1+\gamma}-y)/\gamma$, $\gamma>0$. The consistency and asymptotic normality of the estimators based on those divergences are commonly obtainable within the general framework of Brègman divergences.

Let $\mathcal{G}=\{g_\theta:\theta\in\Theta^*\}$ be a family of densities, where Θ^* is a compact parameter space in an Euclidean space, and assume that $\mathcal{G}$ satisfies the identifiability condition, namely, $\theta=\theta'$ if and only if $g_\theta=g_{\theta'}$ for any θ and $\theta'\in\Theta^*$. Given an iid random sample $Y_1,\dots,Y_n$ following density $g_\theta\in\mathcal{G}$, the minimum Brègman divergence estimator $\hat{\theta}_n^B$ for true parameter θ_0 is attained as the minimizer of the objective function:

$$\frac{1}{n}\sum_{i=1}^{n}l_i(\theta):=\frac{1}{n}\Big(\sum_{i=1}^{n}\int\Big\{\varphi'(g_\theta(y))g_\theta(y)-\varphi(g_\theta(y))\Big\}dy-\varphi'(g_\theta(Y_i))\Big). \tag{3.10}$$

Notice that $l_i(\theta)$ has the property that

$$El_i(\theta)-El_i(\theta_0)=D(g_{\theta_0},g_\theta)>0\ \text{ for all }\theta\neq\theta_0, \tag{3.11}$$

which ensures the strong convergence of $\hat{\theta}_n^B$ to θ_0 if $E(\sup_{\theta\in\Theta^*}|l_i(\theta)|)<\infty$.

The aforementioned scheme can be directly applied to the BINAR(1) models. For Brègman divergence with generator φ_η, where η belongs to a space $\mathcal{N}\subset\mathbb{R}^m$, $m\geq 1$, similarly to (3.10), we define the minimum Brègman divergence estimator as follows:

$$\hat{\theta}_{\eta,n}=\operatorname*{argmin}_{\theta\in\Theta}\frac{1}{n}\sum_{t=1}^{n}\ell_{\eta,t}(\theta), \tag{3.12}$$

where $\ell_{\eta,t}=\sum_{i=1}^{2}\ell_{\eta,i,t}(\theta)$ with

$$\ell_{\eta,i,t}(\theta)=\sum_{y=0}^{\infty}\Big\{\varphi'_\eta(f_{\theta,i}(y|Y_{t-1}))f_{\theta,i}(y|Y_{t-1})-\varphi_\eta(f_{\theta,i}(y|Y_{t-1}))\Big\}-\varphi'_\eta(f_{\theta,i}(Y_t|Y_{t-1})).$$

Particularly, the minimum power density divergence estimator (MDPDE), with divergence generator $\varphi_\gamma(y)=(y^{1+\gamma}-y)/\gamma$, $\gamma\geq 0$, is given as

$$\hat{\theta}_{\gamma,n}=\ \text{argmin}\ \operatorname*{argmin}_{\theta\in\Theta}\frac{1}{n}\sum_{t=1}^{n}\ell_{\gamma,t}(\theta),$$

where $\ell_{\gamma,t}=\sum_{i=1}^{2}\ell_{\gamma,i,t}(\theta)$ with

$$\ell_{\gamma,i,t}(\theta) = \sum_{y=0}^{\infty} f_{i,\theta}^{1+\gamma}(y|Y_{t-1}) - \left(1 + \frac{1}{\gamma}\right) f_{i,\theta}^{\gamma}(Y_t|Y_{t-1}).$$

MDPDE is well known to be a competent robust estimator, controlling the degree of robustness with the tuning parameter γ. When $\gamma = 0$, it becomes QMLE, and when $\gamma = 1$, it is an L_2 estimator.

By virtue of (3.11) and the model identifiability of BINAR(1) models, provided

$$E\Big(\sup_{\theta \in \Theta} |\ell_{\eta,i,t}(\theta)| \Big) < \infty, \tag{3.13}$$

we can readily see that $\hat{\theta}_{\eta,n}$ in (3.12) is strongly consistent to θ_0, as (3.13) implicates that almost surely,

$$\sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{t=1}^{n} \ell_{\eta,i,t}(\theta) - E(\ell_{\eta,i,t}(\theta)) \right| = o(1),$$

due to the uniform strong law of large numbers for stationary and ergodic processes.

Moreover, if the following additionally holds,

$$E\left\| \frac{\partial \ell_{\eta,t}(\theta_0)}{\partial \theta} \right\|^2 < \infty, \quad E\left(\sup_{\theta \in \Theta} \left\| \frac{\partial^2 \ell_{\eta,t}(\theta)}{\partial \theta \partial \theta^T} \right\| \right) < \infty, \tag{3.14}$$

which lead to the following theorem, the proof of which is similar to that of Theorem 1 and is omitted for brevity. Particularly, it can be shown that MDPDE satisfies the conditions in (3.13) and (3.14), refer to Lee and Jo (2023b).

Theorem 2 *Suppose that* (2.3) *holds,* θ_0 *is an interior of a compact parameter space* Θ, *and* Z_t *has a pmf of the form either in* (2.5) *with* f_{ψ_i} *in* (2.6) *or* (2.7) *and satisfies* $E||Z_1||^4 < \infty$. *In addition, if* (3.13) *and* (3.14) *are satisfied, as* $n \to \infty$, $\hat{\theta}_{n,\eta} \longrightarrow \theta_0$ *a.s., and*

$$\sqrt{n}(\hat{\theta}_{\eta,n} - \theta_0) \xrightarrow{d} N(0, \mathcal{J}_\eta^{-1} \mathcal{K}_\eta \mathcal{J}_\eta^{-1}),$$

where

$$\mathcal{J}_\eta = -E\left(\frac{\partial^2 \ell_{\eta,t}(\theta_0)}{\partial \theta \partial \theta^T} \right), \; \mathcal{K}_\eta = E\left(\frac{\partial \ell_{\eta,t}(\theta_0)}{\partial \theta} \frac{\partial \ell_{\eta,t}(\theta_0)}{\partial \theta^T} \right),$$

and $\mathcal{J}_\eta$ *is assumed to be non-singular.*

Remark 1 Using a convex combination of the φ-functions of BED and DPD with tuning parameter a and γ, respectively, Singh et al. (2021) proposed the exponential power divergence (EPD) family with φ in (3.9) of the form:

$$\varphi_\eta(y) = b\frac{e^{ay} - ay - 1}{a^2} + (1-b)\frac{y^{1+\gamma} - y}{\gamma} \tag{3.15}$$

with $\eta = (a, b, \gamma)^T$, $a \in \mathbb{R}$, $b \in [0,1]$, and $\gamma \geq 0$. Kim and Lee (2024) recently applied this to INGARCH(1,1) models. From (3.15), the minimum exponential power divergence estimator (MEPDE) is obtained as the minimizer of $\ell_{\eta,t}(\theta) = \sum_{i=1}^{2} \ell_{\eta,i,t}(\theta)$ with

$$\begin{aligned}\ell_{\eta,i,t}(\theta) = \sum_{y=0}^{\infty}&\left[\frac{b}{a^2}\left\{e^{af_{i,\theta}(y|Y_{t-1})}(af_{i,\theta}(y|Y_{t-1}) - 1) + 1\right\} + (1-b)f_{i,\theta}^{1+\gamma}(y|Y_{t-1})\right]\\ &- \frac{b}{a}\left(e^{af_{i,\theta}(Y_t|Y_{t-1})} - 1\right) - \frac{1-b}{\gamma}\left((1+\gamma)f_{i,\theta}^{\gamma}(Y_t|Y_{t-1}) - 1\right).\end{aligned}$$

Any Brègman divergence estimators satisfying (3.13) and (3.14), including MDPDE and MEPDE, can be harnessed potentially for robust estimation. In our study, though, we employ MDPDE as its performance excels in general and compares well with others as reported in the simulation study of Kim and Lee (2024), while it has much simpler algorithms in computing.

Remark 2 The optimal choice of γ in MDPDE can be a crucial issue. A larger γ is recommended when the portion of outliers is large or the robustness takes precedence over efficiency. Hong and Kim (2001), Warwick (2005), and Warwick and Jones (2005), Fujisawa and Eguchi (2006), Toma and Broniatowski (2011), and Durio and Isaia (2011) developed a procedure for choosing an optimal γ. One can instead use Pearson's correlation as a criterion to choose an optimal γ by comparing the correlations of Y_{ti} and $\hat{Y}_{ti}$ obtained from MDPDE. Lee and Na (2005) conservatively proposed to use a small γ, say, 0.1 to 0.3, in practice, considering the presence of a possible structural change in the dataset.

Though Theorems 1 and 2 hold for general MV-inflated models, in our empirical studies of Sections 4 and 5, we focus on the Poisson-Lindley zero–one inflated BINAR(1) models with the innovational distribution in (3.16) because this kind of time series of counts with zero–one inflation frequently occurs in real situations. If f_{ψ_i} is Poisson-Lindley in (2.7), we can write

$$f_{\varrho_i,\psi_i}(y) = \rho_{i1}I(y=0) + \rho_{i2}I(y=1) + (1-\rho_{i1}-\rho_{i2})\frac{\psi_i^2(\psi_i+2+y)}{(1+\psi_i)^{y+3}}, \quad y \geq 0, \tag{3.16}$$

which composes the pmf in (3.8). Note that the corresponding mean value of (3.16) is as follows:

$$\mu_{\varrho_i,\psi_i} := \sum_{y=0}^{\infty} y f_{\varrho_i,\psi_i}(y) = \rho_{i1} + \rho_{i2}\frac{\psi_i+2}{\psi_i(1+\psi_i)}.$$

As $\mu_{\varrho_i,\psi_i} = \mu_{\varrho_i',\psi_i'}$ does not necessarily guarantee the coincidence of (ϱ_i, ψ_i) and (ϱ_i', ψ_i'), Mohammadi et al. (2022) proposed to use a two-step CLSE using the conditional variance formula of Y_t. However, such a method is not directly applicable to more general multiple inflation cases and requires more laborious efforts. This inconvenience also supports the use of QMLE.

Remark 3 The result of Theorems 1 and 2 can be applied to the change point test. The change point test is critical in the analysis of time series of counts, as the previous studies frequently reveal its existence in practical applications. For performing a test, we set up the hypotheses:

$$\mathcal{H}_0 : \theta_0 \text{ does not change over } Y_1, \ldots, Y_n \quad \text{vs.} \quad \mathcal{H}_1 : \text{not } \mathcal{H}_0,$$

and then perform a test using the score vector-based test statistic:

$$\hat{\mathcal{T}}_n^{\gamma} := \max_{1\le k\le n} \hat{\mathcal{T}}_n^{\gamma}(k) = \max_{1\le k\le n} \frac{1}{n}\left(\sum_{t=1}^{k} \frac{\partial \ell_{\gamma,t}(\hat{\theta}_{\gamma,n})}{\partial\theta}\right)^T \hat{\mathcal{K}}_{\gamma,n}^{-1}\left(\sum_{t=1}^{k} \frac{\partial \ell_{\gamma,t}(\hat{\theta}_{\gamma,n})}{\partial\theta}\right) \tag{3.17}$$

where $\hat{\theta}_{\gamma,n}$ is MDPDE and

$$\hat{\mathcal{K}}_{\gamma,n} = \frac{1}{n}\sum_{t=1}^{n} \frac{\partial \ell_{\gamma,t}(\hat{\theta}_{\gamma,n})}{\partial\theta} \frac{\partial \ell_{\gamma,t}(\hat{\theta}_{\gamma,n})}{\partial\theta^T},$$

which is consistent to $\mathcal{K}_\gamma$ in (3.15) under $\mathcal{H}_0$. Then, provided $\mathcal{K}_\gamma$ is nonsingular, under the conditions in Theorem 2, we can check that under $\mathcal{H}_0$, (3.13) and (3.14) hold true, and

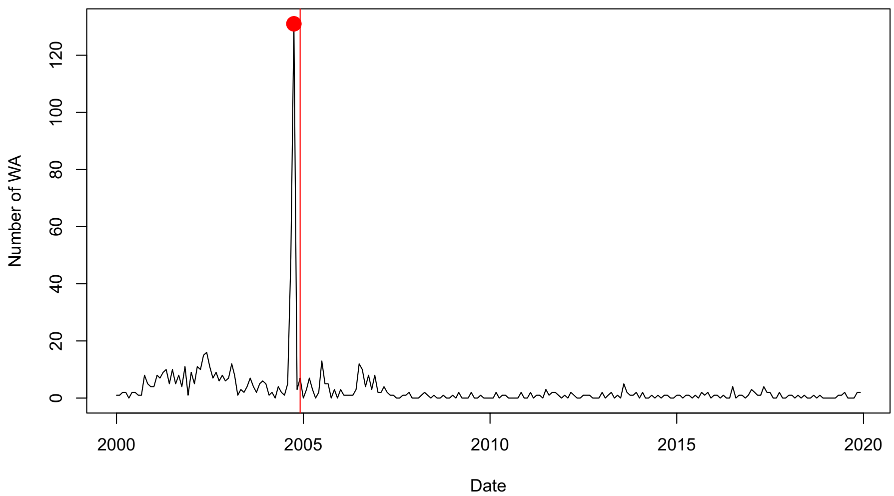

Fig. 1 Monthly time series of WA

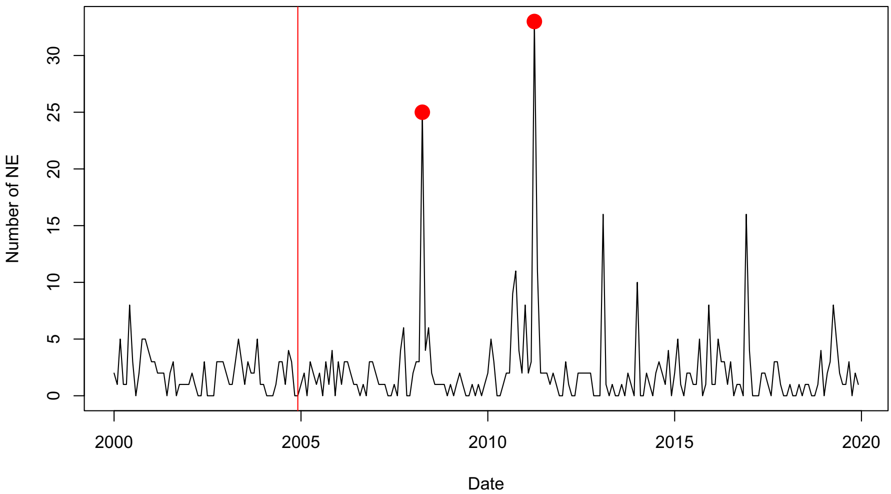

Fig. 2 Monthly time series of NE

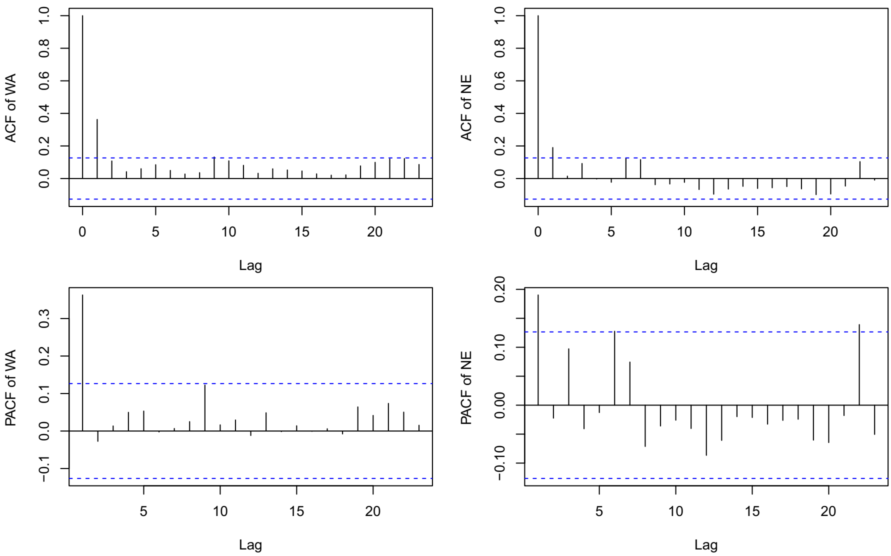

Fig. 3 ACF and PACF of WA (left), and NE (right)

$$\hat{\mathcal{T}}_n \xrightarrow{d} \sup_{0\le s\le 1} \|\mathbf{B}^{\circ}_{M_1+M_2+6}(s)\|^2, \quad n \to \infty, \tag{3.18}$$

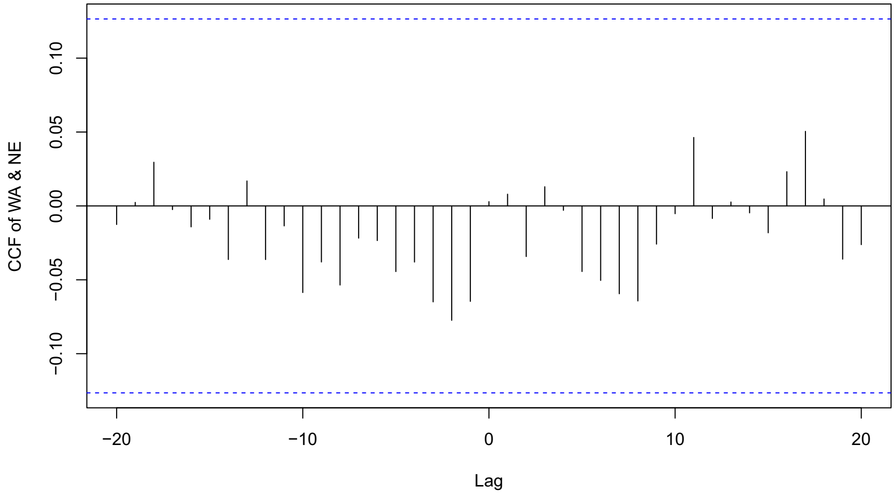

Fig. 4 CCF of WA and NE

where $\mathbf{B}_d^\circ(s)$ denotes a d-dimensional Brownian bridge. For the proof, we refer to Lee and Jo (2023b). The critical value of the test can be obtained through a Monte Carlo simulation, using the result in (3.18). For example, for $d = 10$, $c = 4.435$ is used as the critical value at the significance level 0.05. The location of the change point at the rejection of the null hypothesis is estimated as the point k that maximizes $\hat{\mathcal{T}}_n(k)$ in (3.17).

Remark 4 In practice, detecting outliers is important in modeling and inference because outliers can hamper their correct behaviors. Since Fox (1972), there have been many studies on this subject in time series analysis, refer to Chang et al. (1988) and Tsay et al. (2000). Lee and Jo (2023b) described the procedure for correcting the innovational outliers (IO) and additional outliers (AO) in bivariate random coefficient INAR(1) (BRCINAR(1)) models. Below we describe the AO case that will be used in our empirical study of Section 5. We reexpress Model (2.1) as $Y_t = AY_{t-1} + \mu + \epsilon_t$, where $\mu = EZ_t$ and $\{\epsilon_t\}$ forms a martingale difference sequence. To express the AO-contamination of $\{Y_t\}$ with magnitude ω at time τ, we write $Y_t^A = Y_t + \omega\xi_t^{(\tau)}$, with $\xi_t^{(\tau)} = I(t = \tau)$, which is an actual observation. Then, using $\hat{\epsilon}_t^A = Y_t^A - \hat{A}_n Y_{t-1}^A - \hat{\mu}_n$, where $\hat{A}_n$ and $\hat{\mu}_n$ are MDPDEs of A and μ obtained from observations $Y_1^A, \ldots, Y_n^A$, we obtain the estimates of ω and τ as follows:

$$(\hat{\omega}_n, \hat{\tau}_n) = \underset{\omega,\tau}{\operatorname{argmin}}\ \operatorname{argmin} \frac{1}{n} \sum_{t=1}^{n} \left\| \hat{\epsilon}_t^A - \omega\xi_t^{(\tau)} + \hat{A}_n\omega\xi_t^{(\tau-1)} \right\|^2 .$$

Subsequently, to test the existence of AO at time $\hat{\tau}_n$, we use the chi-square statistic $\hat{\omega}_n^T \hat{\Sigma}_n^{-1} \hat{\omega}_n$, where

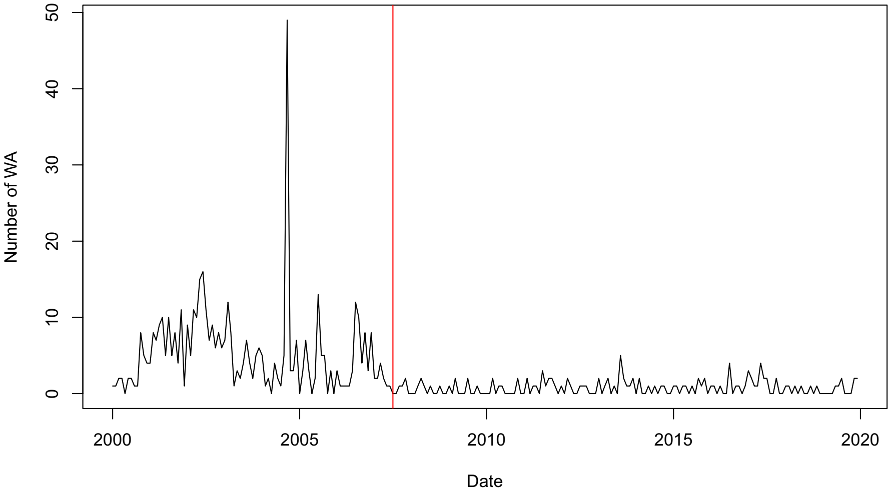

Fig. 5 Monthly time series of WA after the correction of outliers

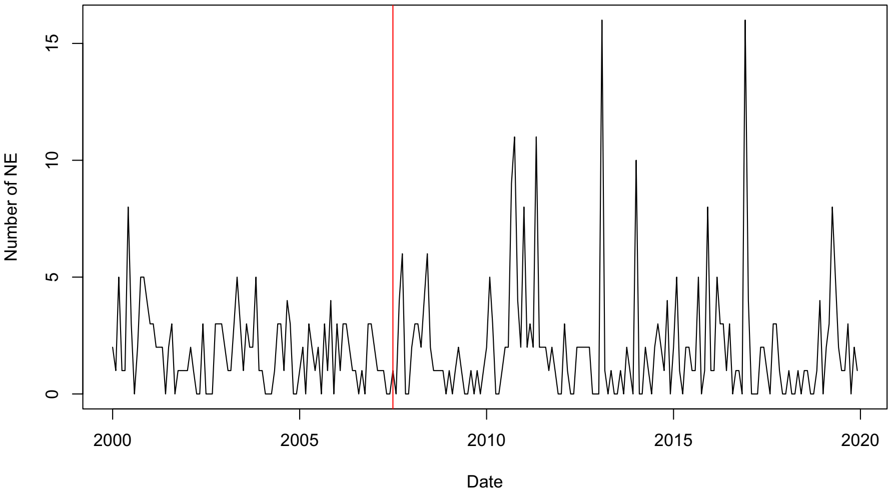

Fig. 6 Monthly time series of NE after the correction of outliers

$$\hat{\Sigma}_n = \frac{1}{n}\sum_{t=1}^{n}(I_2 + \hat{A}_n^T\hat{A}_n)^{-1}(\hat{\epsilon}^A_{\hat{\tau}_n} - \hat{A}_n^T\hat{\epsilon}^A_{\hat{\tau}_n+1})(\hat{\epsilon}^A_{\hat{\tau}_n} - \hat{A}_n^T\hat{\epsilon}^A_{\hat{\tau}_n+1})^T(I_2 + \hat{A}_n^T\hat{A}_n)^{-1},$$

and I_2 is a 2×2 identity matrix. In case an outlier is detected, the corrected Y_t is obtained by subtracting $\hat{w}_n$ from Y_t^A at $t = \hat{\tau}_n$.

Remark 5 In the analysis of count time series models, forecasting is a noteworthy topic to consider. To perform the forecasting through our model, one can employ the

Table 1 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1)$, and no outliers exist

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.50	0.30	0.20	0.59	0.20	0.31	0.30	0.21	0.98	0.96
	Variance $\times 10^2$	0.28	0.26	0.15	0.20	0.33	0.52	0.71	0.25	1.22	1.18
	MSE $\times 10^2$	**0.46**	**0.45**	**0.26**	**0.33**	**0.59**	**0.92**	**1.19**	**0.46**	**33.24**	**64.60**
0.1	Mean	0.53	0.34	0.22	0.62	0.21	0.25	0.32	0.18	1.04	1.06
	Variance $\times 10^2$	0.31	0.22	0.13	0.18	0.31	0.40	0.61	0.24	1.23	1.46
	MSE $\times 10^2$	0.56	0.54	0.27	0.34	0.62	0.93	1.20	0.49	37.40	70.07
0.2	Mean	0.53	0.34	0.22	0.62	0.21	0.26	0.32	0.18	1.06	1.07
	Variance $\times 10^2$	0.33	0.24	0.13	0.20	0.30	0.42	0.62	0.24	2.03	1.75
	MSE $\times 10^2$	0.61	0.56	0.27	0.37	0.61	0.93	1.22	0.49	41.83	74.14
0.3	Mean	0.53	0.34	0.22	0.63	0.21	0.26	0.32	0.19	1.06	1.07
	Variance $\times 10^2$	0.35	0.25	0.14	0.21	0.30	0.44	0.62	0.25	2.65	2.23
	MSE $\times 10^2$	0.65	0.57	0.27	0.39	0.60	0.93	1.23	0.48	45.34	74.69
0.5	Mean	0.53	0.34	0.22	0.63	0.21	0.27	0.32	0.19	1.14	1.18
	Variance $\times 10^2$	0.40	0.28	0.15	0.25	0.30	0.46	0.62	0.26	2.83	2.32
	MSE $\times 10^2$	0.73	0.60	0.28	0.46	0.66	0.96	1.21	0.50	52.58	77.96
1	Mean	0.53	0.33	0.21	0.62	0.21	0.28	0.31	0.20	1.23	1.40
	Variance $\times 10^2$	0.52	0.34	0.18	0.37	0.30	0.54	0.67	0.27	2.97	2.60
	MSE $\times 10^2$	0.94	0.68	0.33	0.64	0.67	1.00	1.31	0.50	58.50	86.27

conditional mean equation in (2.4). First, the target parameters θ are estimated using MDPDE or QMLE from the training data $Y_1, \ldots, Y_n$. Then, the observed values and the estimated parameters are input into the conditional mean equation in (2.4) and (2.5), and the obtained values are rounded to complete the prediction. To perform multi-step ahead forecasting, the predicted values obtained through the one step prediction procedure are recursively utilized for the subsequent forecasting as follows: for $i = 1, 2$,

$$\hat{Y}_{ti} = \left[\hat{\alpha}_{i1}\hat{Y}_{t-1,1} + \hat{\alpha}_{i2}\hat{Y}_{t-1,2} + \sum_{j=1}^{M_i}(j-1)\hat{\rho}_{ij} + (1 - \sum_{j=1}^{M_i}\hat{\rho}_{ij})\hat{\mu}_i\right], \quad t \geq n+1,$$

where $[x]$ denotes the greatest integer not exceeding x, and $\hat{\alpha}_{ij}$, $\hat{\rho}_{ij}$, and $\hat{\mu}_i$ are estimators like MDPDE or QMLE based on $Y_1, \ldots, Y_n$, and $\hat{Y}_{t-1,j}$ are obtained predicted values with initial values $\hat{Y}_{nj} = Y_{nj}$. Specifically, $\hat{\mu}_i$ is the corresponding estimate of $\mu_i = \sum_{y=0}^{\infty} y f_{\psi_i}(y)$, which equals ψ_i in the cases of the one-parameter exponential families. This method is referred to in Jung and Tremayne (2006), and the actual analysis for this forecasting method is referenced in Lee et al. (2023).

Table 2 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 2, 2)$, and no outliers exist

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.39	0.20	0.29	0.49	0.10	0.11	0.21	0.31	1.93	2.12
	Variance $\times 10^2$	0.29	0.13	0.27	0.17	0.02	0.05	0.11	0.38	4.58	6.57
	MSE $\times 10^2$	**0.53**	**0.24**	**0.49**	**0.29**	**0.10**	**0.11**	**0.40**	**0.82**	**26.40**	**32.09**
0.1	Mean	0.41	0.19	0.32	0.51	0.11	0.08	0.19	0.25	1.85	1.87
	Variance $\times 10^2$	0.32	0.15	0.25	0.17	0.06	0.05	0.28	0.35	4.64	7.49
	MSE $\times 10^2$	0.55	0.27	0.50	0.30	0.14	0.13	0.53	0.84	26.62	32.24
0.2	Mean	0.42	0.19	0.32	0.51	0.11	0.08	0.20	0.26	1.75	1.77
	Variance $\times 10^2$	0.30	0.14	0.32	0.18	0.06	0.05	0.29	0.40	7.68	7.84
	MSE $\times 10^2$	0.53	0.24	0.60	0.33	0.14	0.12	0.57	0.87	27.13	33.46
0.3	Mean	0.42	0.19	0.32	0.52	0.11	0.08	0.21	0.26	1.64	1.60
	Variance $\times 10^2$	0.33	0.17	0.28	0.18	0.06	0.28	0.44	0.50	8.46	10.11
	MSE $\times 10^2$	0.58	0.30	0.64	0.32	0.14	0.13	0.54	0.92	29.01	36.07
0.5	Mean	0.41	0.20	0.32	0.52	0.11	0.09	0.21	0.27	1.51	1.52
	Variance $\times 10^2$	0.37	0.17	0.29	0.23	0.06	0.06	0.28	0.52	12.35	16.26
	MSE $\times 10^2$	0.65	0.30	0.65	0.44	0.15	0.14	0.55	1.02	30.79	36.81
1	Mean	0.41	0.20	0.32	0.53	0.10	0.09	0.19	0.29	1.44	1.40
	Variance $\times 10^2$	0.43	0.18	0.37	0.31	0.08	0.07	0.32	0.55	15.15	20.23
	MSE $\times 10^2$	0.75	0.33	0.70	0.61	0.16	0.15	0.63	1.02	31.68	37.56

4 Simulation study

In this section, we evaluate the performance of QMLE and MDPDE for the zero–one inflated BINAR(1) model (namely, $M_1 = M_2 = 2$) with Poison-Lindley innovations. In this experiment, we use the samples of size $n = 200$ and the parameter settings as follows:

$$(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1),$$
$$(0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 2, 2).$$

The repetition number for calculating empirical sizes and powers in each simulation is 1000. We compare the performance of MDPDE ($\gamma = 0.1, 0.2, 0.3, 0.5, 1$) with that of QMLE ($\gamma = 0$). For this, we examine the estimators' sample mean, variance, and mean squared error (MSE). In Tables 1, 2, 3, 4, 5 and 6, the bold figures represent minimal MSEs for each parameter setting. To generate Z_t, we first generate $U_t = (\Phi(V_{t1}), \Phi(V_{t2}))^T$, where $V_t = (V_{t1}, V_{t2})^T$ are iid bivariate normal random variables with $EV_{ti} = 0$, $Var(V_{ti}) = 1$, $i = 1, 2$, and $Corr(V_{t1}, V_{t2}) = \rho$, then $Z_t = (F^{-1}_{\psi_1}(U_{t1}), F^{-1}_{\psi_2}(U_{t2}))^T$, where F_ψ denotes the Poisson-Lindley distribution function with parameter ψ. Here, we take account of $\rho = -0.5, 0, 0.5$.

Table 3 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1)$, and outliers exist

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.52	0.30	0.20	0.63	0.20	0.31	0.29	0.21	0.88	0.88
	Variance $\times 10^2$	0.18	0.21	0.15	0.12	0.31	0.60	0.60	0.30	2.49	2.28
	MSE $\times 10^2$	0.34	0.34	0.25	0.25	0.59	1.12	1.11	0.57	45.25	87.55
0.1	Mean	0.52	0.31	0.20	0.62	0.15	0.29	0.22	0.19	0.85	0.84
	Variance $\times 10^2$	0.16	0.18	0.12	0.11	0.08	0.6	0.18	0.28	1.46	1.27
	MSE $\times 10^2$	0.32	**0.32**	**0.24**	0.23	**0.40**	1.13	**0.89**	0.54	46.31	90.09
0.2	Mean	0.51	0.33	0.22	0.59	0.15	0.31	0.22	0.21	0.95	0.94
	Variance $\times 10^2$	0.16	0.18	0.12	0.12	0.08	0.59	0.21	0.29	2.02	1.75
	MSE $\times 10^2$	**0.29**	0.38	0.25	**0.19**	0.41	1.13	0.92	0.56	42.31	83.04
0.3	Mean	0.54	0.33	0.22	0.64	0.15	0.32	0.23	0.21	1.07	1.05
	Variance $\times 10^2$	0.17	0.19	0.13	0.12	0.11	0.59	0.27	0.29	2.19	2.07
	MSE $\times 10^2$	0.43	0.41	0.28	0.35	0.44	1.19	0.97	0.57	37.93	76.64
0.5	Mean	0.56	0.32	0.21	0.66	0.17	0.35	0.26	0.22	1.06	1.06
	Variance $\times 10^2$	0.22	0.25	0.16	0.18	0.23	0.41	0.53	0.23	1.70	1.68
	MSE $\times 10^2$	0.69	0.48	0.30	0.66	0.56	1.06	1.21	0.52	**30.88**	**65.87**
1	Mean	0.60	0.29	0.18	0.72	0.23	0.32	0.33	0.24	1.30	1.29
	Variance $\times 10^2$	0.24	0.29	0.17	0.24	0.27	0.18	0.60	0.09	2.87	2.92
	MSE $\times 10^2$	1.42	0.50	0.34	1.76	0.62	**0.92**	1.31	**0.43**	47.15	90.64

We first report the simulation result in the case of $\rho = 0$. Tables 1 and 2 show the results when the data is not contaminated by outliers, which shows that QMLE has the minimal MSE for all cases and the MSEs of MDPDE with small γ look similar to those of QMLE. As γ increases, the MSE of MDPDE also increases, which confirms that MDPDE with a large γ has a loss of efficiency as observed in Kim and Lee (2020).

Next, to evaluate the robustness of the estimators against outliers, we generate the contaminated data $Y_{c,t}$ as follows (cf. Fried et al. 2015; Kim and Lee 2020; Lee and Jo 2023b):

$$Y_{c,t} = Y_t + P_t Y_{o,t}, \tag{4.19}$$

where Y_t are generated from (2.1) with the Poisson-Lindley innovation Z_t, P_t are iid Bernoulli random variables with success probability $p = 0.05$, and $Y_{o,t}$ are iid Poisson Lindley random variables with $\psi = 0.5$. Tables 3 and 4 exhibit that MDPDE produces smaller MSEs than QMLE, which supports the superiority of MDPDE over QMLE when outliers contaminate the data.

Tables 5 and 6 exhibit the MSEs of MDPDE and QMLE when a significant outlier exists. Herein, the big outlier is generated from a Poisson-Lindley distribution

Table 4 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 2, 2)$, and outliers exist

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.38	0.18	0.29	0.48	0.11	0.09	0.20	0.25	1.50	1.50
	Variance $\times 10^2$	0.24	0.15	0.26	0.21	0.07	0.07	0.29	0.49	5.35	4.43
	MSE $\times 10^2$	0.48	0.27	0.52	0.38	0.15	0.15	0.56	0.89	34.85	36.12
0.1	Mean	0.42	0.21	0.34	0.52	0.11	0.09	0.20	0.28	1.42	1.41
	Variance $\times 10^2$	0.25	0.13	0.21	0.16	0.06	0.01	0.24	0.09	1.19	0.92
	MSE $\times 10^2$	0.45	**0.25**	0.46	**0.35**	0.14	**0.09**	**0.49**	**0.77**	34.59	35.59
0.2	Mean	0.42	0.21	0.31	0.54	0.11	0.07	0.19	0.22	1.45	1.42
	Variance $\times 10^2$	0.23	0.15	0.21	0.17	0.06	0.02	0.27	0.14	2.34	1.43
	MSE $\times 10^2$	**0.42**	0.28	**0.44**	0.43	0.13	0.10	0.51	0.78	33.57	35.05
0.3	Mean	0.42	0.21	0.33	0.54	0.11	0.07	0.20	0.23	1.50	1.46
	Variance $\times 10^2$	0.27	0.15	0.24	0.18	0.05	0.03	0.28	0.23	4.88	3.97
	MSE $\times 10^2$	0.52	0.27	0.53	0.47	**0.12**	0.11	0.54	0.83	33.30	**34.74**
0.5	Mean	0.42	0.21	0.33	0.55	0.11	0.08	0.21	0.25	1.60	1.62
	Variance $\times 10^2$	0.27	0.17	0.25	0.23	0.05	0.05	0.27	0.37	8.70	12.62
	MSE $\times 10^2$	0.52	0.33	0.52	0.59	0.13	0.13	0.54	0.91	**29.36**	36.42
1	Mean	0.43	0.20	0.33	0.55	0.11	0.08	0.20	0.27	1.86	2.06
	Variance $\times 10^2$	0.38	0.18	0.31	0.32	0.07	0.06	0.31	0.53	16.82	21.15
	MSE $\times 10^2$	0.73	0.33	0.64	0.82	0.15	0.14	0.61	1.04	31.50	38.61

with $\psi = 0.1$ and its location is generated from a uniform distribution over $[\frac{n}{3}, \frac{2n}{3}]$ with $n = 200$. The result shows that the MDPDE has smaller MSEs than QMLE for all cases, confirming the validity of MDPDE as well.

Tables 7, 8, 9 and Tables 10, 11, 12 respectively portray the results on the cases of $\rho = 0.5$ and -0.5, showing a pattern similar to Tables 1, 2, 3, 4, 5 and 6. Thus far, we have conducted experiments to assess the performance of MDPDE in the presence of additive outliers and a big single outlier. Overall, our findings consolidate the excellent performance of MDPDE against outliers.

Finally, Tables 13, 14, 15, 16 display the empirical sizes and powers of the change point tests in Remark 3. The empirical sizes and powers are calculated as the ratios of the rejection numbers of the null hypothesis out of 1000 repetitions. In particular, to assess the empirical power, we assume that the change point occurs at $[n/2]$. Tables 13 and 14 present that there is no significant distortion in the size of the CUSUM test, regardless of the presence of outliers generated from (4.19), but still $\gamma = 0.1, 0.2$ produce more stable results by a slight margin in the presence of outliers. Tables 15 and 16 present that all the tests produce very high powers.

Table 5 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1)$, and a big outlier exists

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.52	0.29	0.19	0.59	0.25	0.37	0.37	0.25	0.71	0.70
	Variance $\times 10^2$	0.21	0.21	0.13	0.12	0.11	0.61	0.18	0.06	0.95	1.42
	MSE $\times 10^2$	0.42	0.36	0.28	0.21	0.40	1.16	1.39	0.56	52.43	108.90
0.1	Mean	0.49	0.30	0.21	0.61	0.15	0.31	0.22	0.20	0.80	0.86
	Variance $\times 10^2$	0.20	0.19	0.14	0.11	0.09	0.59	0.14	0.29	0.93	1.38
	MSE $\times 10^2$	**0.33**	**0.31**	0.26	**0.20**	0.41	1.15	0.85	0.37	48.54	98.33
0.2	Mean	0.53	0.31	0.21	0.63	0.14	0.31	0.22	0.20	0.95	0.97
	Variance $\times 10^2$	0.22	0.21	0.13	0.11	0.05	0.59	0.13	0.29	1.84	1.76
	MSE $\times 10^2$	0.47	0.36	0.24	0.25	**0.39**	**0.86**	0.86	**0.36**	41.98	91.48
0.3	Mean	0.55	0.31	0.21	0.63	0.14	0.33	0.22	0.21	1.07	1.08
	Variance $\times 10^2$	0.23	0.24	0.14	0.13	0.07	0.56	0.14	0.28	1.96	1.90
	MSE $\times 10^2$	0.58	0.42	0.26	0.31	0.40	1.17	0.86	0.56	37.14	85.11
0.5	Mean	0.56	0.30	0.20	0.66	0.16	0.35	0.23	0.22	1.25	1.24
	Variance $\times 10^2$	0.26	0.28	0.13	0.17	0.18	0.37	0.34	0.23	0.78	0.83
	MSE $\times 10^2$	0.79	0.49	**0.22**	0.56	0.50	1.05	1.04	0.54	30.67	75.81
0.1	Mean	0.60	0.28	0.17	0.71	0.20	0.37	0.31	0.24	1.29	1.29
	Variance $\times 10^2$	0.28	0.32	0.16	0.25	0.33	0.16	0.70	0.10	0.52	0.64
	MSE $\times 10^2$	1.47	0.59	0.35	1.62	0.67	0.91	**0.76**	0.43	**29.91**	**72.96**

Our findings in this specific parameter setting demonstrate the effectiveness of the change point tests.

5 Real data analysis

In this section, we illustrate a real data example using the number of monthly earthquake cases (magnitude more than 3) in the United States from January 2000 to December 2019. Time series data can be obtained from the earthquake catalog in the USGS earthquake hazards program (http://usgs.gov). This earthquake data has been intensively studied by researchers on the topics including earthquake forecasting, hazard assessment to seismicity analysis, and identifying temporal and spatial patterns in earthquake datasets, refer to Thomas (1994), Gerstenberger et al. (2005), and Goebel et al. (2017) for a general background. Among the earthquake datasets, we select the bivariate time series of the monthly number of earthquake cases in Washington (WA) and Nevada (NE), with 240 observations. The sample mean and variance are 2.99 and 87.48 for WA and 2.15 and 11.75 for NE. Also, the zero and one portions are 0.33 and 0.26 for WA and 0.27

Table 6 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 2, 2)$, and a big outlier exists

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.37	0.18	0.28	0.48	0.12	0.11	0.24	0.34	1.44	1.40
	Variance $\times 10^2$	0.29	0.13	0.27	0.17	0.06	0.05	0.27	0.38	7.58	8.67
	MSE $\times 10^2$	0.53	0.24	0.49	0.30	0.15	0.13	0.550	0.84	31.68	36.63
0.1	Mean	0.41	0.19	0.32	0.51	0.11	0.08	0.19	0.27	1.51	1.52
	Variance $\times 10^2$	0.32	0.15	0.25	0.17	0.06	0.05	0.28	0.35	4.64	7.49
	MSE $\times 10^2$	0.55	0.27	**0.48**	**0.29**	0.14	0.13	**0.53**	**0.82**	30.79	36.07
0.2	Mean	0.41	0.19	0.32	0.51	0.11	0.08	0.20	0.26	1.58	1.55
	Variance $\times 10^2$	0.30	0.14	0.32	0.18	0.06	0.05	0.29	0.40	7.68	7.84
	MSE $\times 10^2$	**0.52**	**0.23**	0.60	0.33	**0.13**	**0.12**	0.57	0.87	29.01	**32.09**
0.3	Mean	0.41	0.19	0.32	0.52	0.11	0.08	0.21	0.26	1.64	1.60
	Variance $\times 10^2$	0.33	0.17	0.28	0.18	0.06	0.06	0.28	0.44	8.46	10.11
	MSE $\times 10^2$	0.58	0.30	0.54	0.32	0.14	0.13	0.54	0.92	**26.40**	32.24
0.5	Mean	0.41	0.20	0.32	0.52	0.11	0.09	0.21	0.27	1.75	1.77
	Variance $\times 10^2$	0.37	0.17	0.29	0.23	0.06	0.06	0.28	0.52	12.35	16.26
	MSE $\times 10^2$	0.65	0.30	0.55	0.44	0.15	0.14	0.55	1.02	27.13	33.46
1	Mean	0.41	0.20	0.32	0.53	0.10	0.09	0.19	0.29	1.93	2.12
	Variance $\times 10^2$	0.43	0.18	0.37	0.31	0.08	0.07	0.32	0.55	15.15	20.23
	MSE $\times 10^2$	0.75	0.33	0.70	0.61	0.16	0.15	0.63	1.02	26.62	37.56

and 0.27 for NE, indicating a high frequency of zero and one observations in both time series. Figures 1 and 2 plot the monthly time series of WA and NE, particularly showing a possibility of outliers in each series. The autocorrelation function (ACF) and partial autocorrelation function (PACF) of the time series, and the cross-correlation function (CCF) are displayed in Figs. 3 and 4, respectively.

To analyze this bivariate time series data, we fit a 0,1-inflated BINAR(1) model with Poisson-Lindley innovations and choose an optimal γ among $\{0.1, 0.2 \ldots, 1\}$ for MDPDE using the criterion of Hong and Kim (2001), who proposed to select γ that minimizes the trace of the estimated asymptotic variance of $\hat{\theta}_{\gamma,n}$(cf. Lee et al. (2023)). The criterion chooses $\gamma = 0.2$ and we obtain MDPDEs with this choice as:

$$\hat{\alpha}_{11,n} = 0.33, \hat{\alpha}_{12,n} = 0.09, \hat{\alpha}_{21,n} = 0.11, \hat{\alpha}_{22,n} = 0.18,$$
$$\hat{\rho}_{11,n} = 0.29, \hat{\rho}_{12,n} = 0.40, \hat{\rho}_{21,n} = 0.14, \hat{\rho}_{22,n} = 0.42, \hat{\psi}_{1,n} = 0.57, \hat{\psi}_{2,n} = 0.85.$$

The test $\hat{T}_n^{\gamma}$ with $\gamma = 0.2$ in (3.17) detects one change point at $t = 60$, indicated by the red vertical lines in Figs. 1 and 2. However, this result is unreliable as the outliers exist before and after the change point.

Table 7 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1)$, $\phi = 0.5$, and no outliers exist

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.49	0.30	0.20	0.59	0.19	0.30	0.28	0.20	1.00	1.03
	Variance $\times 10^2$	0.22	0.23	0.15	0.15	0.30	0.60	0.58	0.30	2.78	2.69
	MSE $\times 10^2$	**0.23**	**0.23**	**0.15**	**0.15**	**0.30**	**0.57**	**0.56**	**0.28**	**2.77**	**2.78**
0.1	Mean	0.51	0.30	0.20	0.61	0.23	0.28	0.32	0.18	0.78	0.76
	Variance $\times 10^2$	0.24	0.25	0.15	0.15	0.20	0.55	0.48	0.27	2.07	1.60
	MSE $\times 10^2$	0.25	0.25	0.15	0.17	0.30	0.58	0.56	0.28	6.87	7.01
0.2	Mean	0.51	0.30	0.20	0.61	0.23	0.28	0.32	0.19	0.80	0.78
	Variance $\times 10^2$	0.25	0.25	0.15	0.16	0.20	0.56	0.50	0.28	2.68	1.86
	MSE $\times 10^2$	0.26	0.26	0.15	0.17	0.30	0.59	0.58	0.29	6.37	6.45
0.3	Mean	0.51	0.30	0.20	0.61	0.23	0.28	0.32	0.19	0.83	0.81
	Variance $\times 10^2$	0.26	0.26	0.15	0.16	0.20	0.58	0.53	0.29	3.19	2.26
	MSE $\times 10^2$	0.28	0.27	0.15	0.18	0.30	0.60	0.59	0.29	5.85	5.83
0.5	Mean	0.51	0.30	0.20	0.61	0.22	0.28	0.31	0.19	0.90	0.87
	Variance $\times 10^2$	0.29	0.28	0.16	0.18	0.23	0.60	0.59	0.30	4.31	3.11
	MSE $\times 10^2$	0.31	0.29	0.16	0.20	0.30	0.61	0.62	0.30	5.16	4.67
1	Mean	0.51	0.30	0.20	0.61	0.21	0.29	0.30	0.19	1.03	0.99
	Variance $\times 10^2$	0.35	0.32	0.18	0.23	0.28	0.65	0.67	0.31	4.62	4.05
	MSE $\times 10^2$	0.39	0.33	0.18	0.25	0.31	0.65	0.67	0.31	4.73	4.05

To identify and correct the outliers, we use the method described in Remark 4. As a result, three AO's are identified at $t = 58, 100$ and 136, specified by the red points in Figs. 1 and 2. After the correction of these three outliers, MDPDEs with $\gamma = 0.2$ are obtained as

$$\hat{\alpha}_{11,n} = 0.33, \hat{\alpha}_{12,n} = 0.11, \hat{\alpha}_{21,n} = 0.12, \hat{\alpha}_{22,n} = 0.16,$$
$$\hat{\rho}_{11,n} = 0.39, \hat{\rho}_{12,n} = 0.13, \hat{\rho}_{21,n} = 0.23, \hat{\rho}_{22,n} = 0.12, \hat{\psi}_{1,n} = 0.33, \hat{\psi}_{2,n} = 0.48.$$

which significantly differ from what we have obtained previously. In particular, ρ and ψ's estimators appear to experience a dramatic change compared to the others. This result implies that the outliers significantly influences on the estimation of the zero–one inflation and Poisson-Lindley parameters. Unlike before, $\hat{\mathcal{T}}_n^{\gamma}$ with $\gamma = 0.2$ is revealed to detect a change point at $t = 91$ after the correction of outliers, as seen in Figs. 5 and 6, which looks more reasonable compared to the previously obtained change point of $t = 60$. This result reaffirms that outliers can severely impact both the parameter estimation and change point test and decrease their accuracies.

MDPDEs with $\gamma = 0.2$ obtained from the two subseries before and after the change point are as follows:

Table 8 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1)$, $\phi = 0.5$, and outliers exist

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.49	0.29	0.20	0.59	0.19	0.30	0.29	0.20	0.88	0.88
	Variance $\times 10^2$	0.19	0.22	0.13	0.14	0.30	0.61	0.60	0.28	2.56	2.38
	MSE $\times 10^2$	0.25	0.26	0.16	0.18	0.30	0.61	0.60	0.29	4.92	4.73
0.1	Mean	0.51	0.31	0.21	0.61	0.21	0.25	0.31	0.17	0.72	0.71
	Variance $\times 10^2$	0.20	0.24	0.14	0.13	0.25	0.36	0.49	0.20	0.66	0.30
	MSE $\times 10^2$	0.24	0.25	0.15	**0.15**	**0.28**	0.60	**0.51**	0.28	8.30	8.47
0.2	Mean	0.51	0.31	0.20	0.61	0.22	0.25	0.31	0.17	0.74	0.73
	Variance $\times 10^2$	0.21	0.23	0.15	0.15	0.24	0.37	0.49	0.22	1.19	0.74
	MSE $\times 10^2$	**0.24**	**0.25**	0.15	0.19	0.30	**0.60**	0.52	0.28	7.61	7.70
0.3	Mean	0.52	0.31	0.20	0.62	0.22	0.25	0.32	0.18	0.76	0.75
	Variance $\times 10^2$	0.22	0.24	0.15	0.16	0.24	0.44	0.51	0.24	1.75	1.10
	MSE $\times 10^2$	0.26	0.26	**0.15**	0.21	0.30	0.62	0.55	**0.28**	7.12	7.24
0.5	Mean	0.52	0.31	0.20	0.62	0.22	0.26	0.31	0.18	0.82	0.80
	Variance $\times 10^2$	0.25	0.26	0.16	0.19	0.24	0.54	0.55	0.27	3.09	2.11
	MSE $\times 10^2$	0.30	0.27	0.16	0.25	0.30	0.65	0.59	0.29	6.14	6.10
1	Mean	0.52	0.31	0.20	0.62	0.21	0.27	0.30	0.18	0.97	0.94
	Variance $\times 10^2$	0.33	0.32	0.17	0.25	0.29	0.64	0.65	0.30	4.75	3.61
	MSE $\times 10^2$	0.40	0.33	0.17	0.34	0.31	0.69	0.65	0.31	**4.79**	**3.94**

$$\hat{\alpha}_{11,n} = 0.39, \hat{\alpha}_{12,n} = 0.26, \hat{\alpha}_{21,n} = 0.14, \hat{\alpha}_{22,n} = 0.23,$$
$$\hat{\rho}_{11,n} = 0.14, \hat{\rho}_{12,n} = 0.14, \hat{\rho}_{21,n} = 0.16, \hat{\rho}_{22,n} = 0.15, \hat{\psi}_{1,n} = 0.30, \hat{\psi}_{2,n} = 0.42.$$

and

$$\hat{\alpha}_{11,n} = 0.21, \hat{\alpha}_{12,n} = 0.15, \hat{\alpha}_{21,n} = 0.16, \hat{\alpha}_{22,n} = 0.21,$$
$$\hat{\rho}_{11,n} = 0.24, \hat{\rho}_{12,n} = 0.22, \hat{\rho}_{21,n} = 0.13, \hat{\rho}_{22,n} = 0.28, \hat{\psi}_{1,n} = 0.23, \hat{\psi}_{2,n} = 0.18.$$

The MDPDEs before and after the change point exhibit a significant difference, particularly in $\hat{\alpha}_{11,n}$, $\hat{\alpha}_{12,n}$, $\hat{\rho}_{11,n}$, $\hat{\rho}_{22,n}$, and $\hat{\psi}_{2,n}$. Detecting the change point in the model is crucial due to the significant difference in MDPDEs before and after the change point, and this is also discussed in Lee and Jo (2023b). All our findings strongly confirm the functionality of our proposed methods in real applications.

Table 9 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1)$, $\phi = 0.5$, and a big outlier exists

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.49	0.29	0.19	0.59	0.23	0.36	0.36	0.24	0.76	0.70
	Variance $\times 10^2$	0.21	0.22	0.14	0.14	0.18	0.30	0.20	0.07	0.59	0.51
	MSE $\times 10^2$	0.26	0.23	0.16	0.14	0.31	0.69	0.66	0.32	6.21	8.95
0.1	Mean	0.51	0.30	0.20	0.60	0.22	0.29	0.34	0.19	0.71	0.72
	Variance $\times 10^2$	0.26	0.25	0.15	0.12	0.22	0.49	0.40	0.27	0.48	0.74
	MSE $\times 10^2$	0.27	0.25	0.16	**0.13**	0.29	**0.49**	0.57	**0.27**	8.63	8.20
0.2	Mean	0.51	0.30	0.20	0.61	0.22	0.29	0.33	0.19	0.75	0.77
	Variance $\times 10^2$	0.23	0.22	0.15	0.14	0.22	0.54	0.42	0.29	1.51	2.07
	MSE $\times 10^2$	**0.25**	**0.22**	0.15	0.16	**0.28**	0.55	**0.54**	0.29	7.54	6.92
0.3	Mean	0.51	0.30	0.20	0.61	0.22	0.29	0.33	0.19	0.80	0.82
	Variance $\times 10^2$	0.24	0.22	0.14	0.14	0.23	0.57	0.47	0.30	2.46	2.72
	MSE $\times 10^2$	0.26	0.23	**0.15**	0.17	0.29	0.58	0.57	0.30	6.19	5.93
0.5	Mean	0.51	0.30	0.20	0.62	0.22	0.28	0.32	0.19	0.89	0.89
	Variance $\times 10^2$	0.24	0.23	0.15	0.16	0.26	0.60	0.53	0.30	3.56	3.34
	MSE $\times 10^2$	0.27	0.23	0.15	0.21	0.31	0.61	0.59	0.30	4.57	4.34
1	Mean	0.52	0.30	0.20	0.62	0.20	0.29	0.30	0.19	1.05	1.03
	Variance $\times 10^2$	0.32	0.28	0.17	0.20	0.30	0.66	0.65	0.30	3.82	3.74
	MSE $\times 10^2$	0.37	0.28	0.17	0.26	0.31	0.66	0.65	0.31	**4.10**	**3.88**

6 Concluding remarks

In this study, we introduced an MV-inflated BINAR(1) model and substantiated the strong consistency and asymptotic normality of QMLE and MDPDE under regularity conditions. Moreover, the change point test and outlier detection method based on MDPDE were outlined for a practical application. To evaluate the performance of MDPDE in the presence of outliers, we conducted Monte Carlo simulations and real data analysis using the number of monthly earthquake cases in the United States. All the acquired results affirmed the validity of our proposed methods. While we focused on stationary time series models here, as mentioned by a referee, one can also consider an extension of our methods to time series of counts with more complicated characteristics, for example, time series with a (stochastic) trend or periodicity (seasonality). As this problem extends beyond the scope of our current study, it is deferred as a subject for future research projects.

Table 10 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 2, 2)$, $\phi = -0.5$, and no outliers exist

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.39	0.20	0.30	0.49	0.10	0.10	0.19	0.30	2.05	2.09
	Variance $\times 10^2$	0.28	0.16	0.25	0.21	0.08	0.07	0.32	0.51	9.94	14.73
	MSE $\times 10^2$	**0.28**	**0.16**	**0.25**	**0.21**	**0.08**	**0.07**	**0.31**	**0.51**	**10.28**	**14.56**
0.1	Mean	0.39	0.19	0.30	0.49	0.09	0.10	0.19	0.30	2.03	2.04
	Variance $\times 10^2$	0.29	0.16	0.26	0.21	0.08	0.07	0.32	0.52	10.51	15.24
	MSE $\times 10^2$	0.29	0.16	0.26	0.22	0.08	0.07	0.32	0.52	10.64	15.42
0.2	Mean	0.39	0.19	0.30	0.49	0.09	0.10	0.19	0.30	2.04	2.04
	Variance $\times 10^2$	0.30	0.16	0.26	0.22	0.08	0.07	0.31	0.53	10.88	14.94
	MSE $\times 10^2$	0.30	0.16	0.26	0.22	0.08	0.07	0.31	0.53	11.02	15.10
0.3	Mean	0.39	0.19	0.30	0.49	0.09	0.10	0.10	0.30	2.03	2.04
	Variance $\times 10^2$	0.31	0.16	0.26	0.22	0.08	0.07	0.31	0.54	11.31	15.17
	MSE $\times 10^2$	0.31	0.16	0.26	0.23	0.08	0.07	0.31	0.54	11.45	15.33
0.5	Mean	0.39	0.19	0.30	0.49	0.09	0.10	0.19	0.30	2.03	2.03
	Variance $\times 10^2$	0.33	0.17	0.27	0.24	0.08	0.07	0.32	0.56	12.39	16.29
	MSE $\times 10^2$	0.33	0.17	0.27	0.24	0.08	0.07	0.32	0.56	12.53	16.42
1	Mean	0.39	0.19	0.30	0.49	0.10	0.10	0.19	0.29	2.03	2.02
	Variance $\times 10^2$	0.40	0.18	0.31	0.29	0.08	0.07	0.32	0.60	15.19	20.44
	MSE $\times 10^2$	0.40	0.18	0.31	0.29	0.08	0.07	0.32	0.60	15.30	20.44

Proof

Proof of Theorem 1 From (3.8), one can see that

$$\begin{aligned} 0 \le -\ell_t(\theta) &\le -\sum_{i=1}^{2} J(Y_{t-1,1}, Y_{t-1,2}, \alpha_{i1}, \alpha_{i2}, 0) + \sum_{i=1}^{2} |\log f_{\varrho_i,\psi_i}(Y_{ti})| \\ &\le (Y_{t-1,1} + Y_{t-1,2}) \sum_{i=1}^{2} |\log(1-\alpha_{i1}) + \log(1-\alpha_{i2})| + \sum_{i=1}^{2} |\log f_{\varrho_i,\psi_i}(Y_{ti})|, \end{aligned}$$

due to the fact: $J(m, n, p_1, p_2, 0) = (1-p_1)^m(1-p_2)^n$ for any $m, n \ge 0$. Then, provided f_{ψ_i} is the one in (2.6), satisfying $E|\log h(Y_{ti})| < \infty,\ i = 1, 2$, which holds particularly true for the zero-inflated Poisson innovation case, we have

$$E\Big(\sup_{\theta\in\Theta} |\ell_{ti}(\theta)|\Big) < \infty.$$

Table 11 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 2, 2)$, $\phi = -0.5$, and outliers exist

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.38	0.19	0.29	0.49	0.11	0.10	0.20	0.28	1.58	1.50
	Variance $\times 10^2$	0.24	0.15	0.21	0.18	0.07	0.07	0.29	0.49	6.40	5.12
	MSE $\times 10^2$	0.26	0.16	0.24	0.19	0.08	0.07	0.31	0.55	23.68	29.79
0.1	Mean	0.38	0.19	0.29	0.49	0.11	0.10	0.20	0.28	1.62	1.49
	Variance $\times 10^2$	0.23	0.14	0.25	0.19	0.06	0.07	0.29	0.52	7.27	4.46
	MSE $\times 10^2$	**0.25**	**0.15**	0.25	**0.19**	0.08	0.07	0.30	0.54	21.38	29.99
0.2	Mean	0.38	0.19	0.29	0.49	0.10	0.09	0.20	0.28	1.67	1.53
	Variance $\times 10^2$	0.25	0.15	0.22	0.20	0.07	0.07	0.29	0.52	8.26	6.93
	MSE $\times 10^2$	0.26	0.15	**0.22**	0.20	**0.08**	**0.07**	0.30	**0.53**	19.04	28.30
0.3	Mean	0.39	0.19	0.29	0.49	0.10	0.09	0.20	0.29	1.71	1.57
	Variance $\times 10^2$	0.26	0.15	0.23	0.20	0.08	0.08	0.29	0.54	9.86	8.49
	MSE $\times 10^2$	0.26	0.16	0.23	0.20	0.08	0.08	**0.30**	0.54	17.69	26.56
0.5	Mean	0.39	0.19	0.29	0.49	0.10	0.10	0.20	0.29	1.78	1.65
	Variance $\times 10^2$	0.27	0.16	0.24	0.21	0.08	0.08	0.30	0.55	12.05	12.02
	MSE $\times 10^2$	0.28	0.16	0.24	0.22	0.08	0.08	0.30	0.55	**16.78**	**23.80**
1	Mean	0.39	0.19	0.29	0.49	0.10	0.10	0.19	0.29	1.84	1.76
	Variance $\times 10^2$	0.31	0.17	0.29	0.26	0.08	0.08	0.31	0.58	15.56	18.75
	MSE $\times 10^2$	0.32	0.17	0.29	0.26	0.08	0.08	0.32	0.58	18.00	24.02

This can be shown to hold for the Poisson-Lindley case as well. Then, using the continuity of $\ell_t(\theta)$ in θ and the uniform strong law of large numbers for stationary ergodic processes, we can get

$$\sup_{\theta\in\Theta}\left|\frac{1}{n}\sum_{t=1}^{n}\ell_{ti}(\theta) - E\ell_{ti}(\theta)\right| = o(1) \text{ a.s.},$$

which implies the strong convergence of $\hat{\theta}_n$ to θ_0, as $E\ell_t(\theta) < E\ell_t(\theta_0)$ for all $\theta \neq \theta_0$, due to the Kullback–Leibler divergence property and the model identifiability discussed earlier.

Moreover, using simple algebras, we can easily check that

$$E\Big(\sup_{\theta\in\Theta}\Big\|\frac{\partial \log f_{\theta,i}(Y_t|Y_{t-1})}{\partial\theta}\Big\|^2\Big) < \infty, \quad E\Big(\sup_{\theta\in\Theta}\Big\|\frac{\partial^2 \log f_{\theta,i}(Y_t|Y_{t-1})}{\partial\theta\partial\theta^T}\Big\|\Big) < \infty,$$

which results in

Table 12 Sample mean, variance, and MSE of estimators for the zero–one inflated BINAR(1) model when $(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2) = (0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 2, 2)$, $\phi = -0.5$, and a big outlier exists

γ	statistics	estimators									
		$\hat{a}_{11}$	$\hat{a}_{12}$	$\hat{a}_{21}$	$\hat{a}_{22}$	$\hat{\rho}_{11}$	$\hat{\rho}_{12}$	$\hat{\rho}_{21}$	$\hat{\rho}_{22}$	$\hat{\psi}_1$	$\hat{\psi}_2$
0(QMLE)	Mean	0.38	0.18	0.28	0.48	0.12	0.11	0.24	0.34	1.53	1.40
	Variance $\times 10^2$	0.27	0.13	0.25	0.17	0.01	0.06	0.14	0.29	10.89	10.78
	MSE $\times 10^2$	0.31	0.16	0.27	0.19	0.08	0.07	0.31	0.54	23.96	31.90
0.1	Mean	0.39	0.20	0.29	0.49	0.11	0.10	0.20	0.31	1.88	1.94
	Variance $\times 10^2$	0.27	0.14	0.27	0.18	0.06	0.07	0.30	0.55	9.29	17.50
	MSE $\times 10^2$	0.27	0.14	0.27	0.18	0.08	**0.07**	0.31	0.57	10.50	17.69
0.2	Mean	0.39	0.20	0.30	0.49	0.10	0.09	0.20	0.30	2.02	1.99
	Variance $\times 10^2$	0.25	0.13	0.26	0.18	0.08	0.07	0.32	0.53	9.94	16.24
	MSE $\times 10^2$	0.25	0.13	0.26	0.18	**0.07**	0.07	0.32	0.54	**9.95**	**16.18**
0.3	Mean	0.39	0.20	0.30	0.49	0.10	0.09	0.20	0.30	2.04	1.99
	Variance $\times 10^2$	0.25	0.13	0.25	0.18	0.08	0.08	0.31	0.54	9.92	16.61
	MSE $\times 10^2$	**0.25**	**0.13**	**0.25**	**0.18**	0.08	0.08	**0.31**	**0.54**	10.09	16.55
0.5	Mean	0.39	0.20	0.30	0.49	0.10	0.09	0.19	0.30	2.04	1.99
	Variance $\times 10^2$	0.27	0.14	0.27	0.21	0.08	0.08	0.31	0.55	10.64	18.17
	MSE $\times 10^2$	0.27	0.14	0.27	0.21	0.08	0.08	0.31	0.55	10.83	18.09
1	Mean	0.39	0.20	0.30	0.49	0.10	0.09	0.20	0.30	2.04	2.00
	Variance $\times 10^2$	0.33	0.16	0.31	0.29	0.08	0.08	0.31	0.59	14.16	22.23
	MSE $\times 10^2$	0.33	0.16	0.31	0.29	0.08	0.08	0.31	0.59	14.28	22.14

Table 13 Empirical sizes for the zero–one inflated BINAR(1) model when no ouliers exist

$(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2)$	size					
	$\gamma = 0$	0.1	0.2	0.3	0.5	1
(0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1)	0.52	0.51	0.54	0.48	0.50	0.51
(0.5, 0.3, 0.2, 0.6, 0.1, 0.2, 0.2, 0.2, 2, 2)	0.51	0.52	0.47	0.49	0.48	0.50
(0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 1, 1)	0.48	0.55	0.54	0.56	0.55	0.54
(0.4, 0.2, 0.3, 0.5, 0.2, 0.1, 0.2, 0.2, 2, 2)	0.52	0.45	0.48	0.47	0.48	0.56

Table 14 Empirical sizes for the zero–one inflated BINAR(1) model when ouliers exist

$(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2)$	size					
	$\gamma = 0$	0.1	0.2	0.3	0.5	1
(0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1)	0.51	0.48	0.48	0.53	0.55	0.54
(0.5, 0.3, 0.2, 0.6, 0.1, 0.2, 0.2, 0.2, 2, 2)	0.52	0.50	0.52	0.48	0.46	0.48
(0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 1, 1)	0.54	0.50	0.51	0.47	0.46	0.51
(0.4, 0.2, 0.3, 0.5, 0.2, 0.1, 0.2, 0.2, 2, 2)	0.53	0.51	0.50	0.53	0.52	0.50

Table 15 Empirical powers for the zero–one inflated BINAR(1) model when no ouliers exist

$(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2)$	power					
$\to (a'_{11}, a'_{12}, a'_{21}, a'_{22}, \rho'_{11}, \rho'_{12}, \rho'_{21}, \rho'_{22}, \psi'_1, \psi'_2)$	$\gamma = 0$	0.1	0.2	0.3	0.5	1
(0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1) $\to$ (0.4, 0.2, 0.2, 0.6, 0.2, 0.3, 0.2, 0.3, 1, 1)	0.89	0.88	0.89	0.85	0.84	0.91
(0.5, 0.3, 0.2, 0.6, 0.1, 0.2, 0.2, 0.2, 2, 2) $\to$ (0.5, 0.3, 0.2, 0.5, 0.1, 0.1, 0.1, 0.1, 2, 1)	0.93	0.92	0.94	0.95	0.95	0.93
(0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 1, 1) $\to$ (0.5, 0.3, 0.2, 0.6, 0.2, 0.2, 0.2, 0.2, 1, 1)	0.99	0.98	0.97	0.97	0.95	0.99
(0.4, 0.2, 0.3, 0.5, 0.2, 0.1, 0.2, 0.2, 2, 2) $\to$ (0.4, 0.2, 0.2, 0.5, 0.1, 0.2, 0.3, 0.1, 1, 1)	0.99	0.99	0.99	0.99	0.99	0.99

Table 16 Empirical powers for the zero–one inflated BINAR(1) model when ouliers exist

$(a_{11}, a_{12}, a_{21}, a_{22}, \rho_{11}, \rho_{12}, \rho_{21}, \rho_{22}, \psi_1, \psi_2)$	power					
$\to (a'_{11}, a'_{12}, a'_{21}, a'_{22}, \rho'_{11}, \rho'_{12}, \rho'_{21}, \rho'_{22}, \psi'_1, \psi'_2)$	$\gamma = 0$	0.1	0.2	0.3	0.5	1
(0.5, 0.3, 0.2, 0.6, 0.2, 0.3, 0.3, 0.2, 1, 1) $\to$ (0.4, 0.2, 0.2, 0.6, 0.2, 0.3, 0.2, 0.3, 1, 1)	0.97	0.99	0.99	0.98	0.98	0.99
(0.5, 0.3, 0.2, 0.6, 0.1, 0.2, 0.2, 0.2, 2, 2) $\to$ (0.5, 0.3, 0.2, 0.5, 0.1, 0.1, 0.1, 0.1, 2, 1)	0.98	0.99	0.98	0.98	0.99	0.99
(0.4, 0.2, 0.3, 0.5, 0.1, 0.1, 0.2, 0.3, 1, 1) $\to$ (0.5, 0.3, 0.2, 0.6, 0.2, 0.2, 0.2, 0.2, 1, 1)	0.99	0.98	0.98	0.97	0.98	0.99
(0.4, 0.2, 0.3, 0.5, 0.2, 0.1, 0.2, 0.2, 2, 2) $\to$ (0.4, 0.2, 0.2, 0.5, 0.1, 0.2, 0.3, 0.1, 1, 1)	0.99	0.99	0.99	0.99	0.99	0.99

$$E\left\|\frac{\partial \ell_t(\theta_0)}{\partial \theta}\right\|^2 < \infty, \quad E\left(\sup_{\theta\in\Theta}\left\|\frac{\partial^2 \ell_t(\theta)}{\partial\theta\partial\theta^T}\right\|\right) < \infty. \tag{7.20}$$

Then, applying the martingale central limit theorem to $\{\frac{\partial \ell_t(\theta_0)}{\partial \theta}\}$, we obtain

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial \ell_t(\theta_0)}{\partial \theta}\xrightarrow{d} N(0, \mathcal{K}_0).$$

Moreover, owing to (7.20), using the uniform law of strong large numbers combined with the ergodicity of $\{Y_t\}$ and the dominated convergence theorem, we can have

$$\left\|\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 \ell_t(\tilde\theta_n)}{\partial\theta\partial\theta^T} - E\left(\frac{\partial^2 \ell_t(\theta_0)}{\partial\theta\partial\theta^T}\right)\right\|_1 = o(1) \ \text{ a.s.},$$

where $\tilde\theta_n$ is an intermediate point between θ_0 and $\hat\theta_n$, and $||A||_1 = \sum_{i,j=1}^{d} |A_{ij}|$ for any $d \times d$ matrix $A = (A_{ij})$, so that

$$\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 \ell_t(\tilde{\theta}_n)}{\partial\theta\partial\theta^T} \to \mathcal{J}_0 \text{ a.s.},$$

which asserts the theorem. □

Acknowledgements The authors are grateful to the Editor, an AE, and two anonymous referees for their careful reading and valuable comments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (NRF-2021R1A2C1004009).

Data availability The data can be obtained from the earthquake catalog in the USGS earthquake hazards program (http://usgs.gov).

Declarations

Conflict of interest. The authors declare no Conflict of interest. Sangyeol Lee is an Associate Editor of Journal of the Korean Statistical Society. Associate Editor status has no bearing on editorial consideration.

References

Al-Osh, M. A., & Alzaid, A. A. (1987). First-order integer-valued autoregressive INAR(1) process. *Journal of Time Series Analysis, 8*, 261–275.

Altun, E. (2020). A new generalization of geometric distribution with properties and applications. *Communications in Statistics-Simulation and Computation, 49*, 793–807.

Berkes, I., Horvath, L., & Kokoszaka, P. (2004). Testing for parametric constancy in GARCH(p, q) models. *Statistics & Probability Letters, 70*, 263–273.

Chang, I., Tiao, G., & Chen, C. (1988). Estimation of time series parameters in the presence of outliers. *Technometrics, 30*, 193–204.

Chen, C., Khamthong, K., & Lee, S. (2019). Markov switching integer-valued generalized auto-regressive conditional heteroscedastic models for dengue counts. *Journal of the Royal Statistical Society: Series C: Applied Statistics, 68*, 963–983.

Csörgő, M. and Horváth, L. (1997). Limit Theorems in Change-Point Analysis. John Wiley & Sons Inc. New York., 18:ISBN:978–0–471–95522–1.

Darolles, S., Fol, G., Lu, Y., & Sun, R. (2019). Bivariate integer-autoregressive process with an application to mutual fund flows. *Journal of Multivariate Analysis, 173*, 181–203.

Durio, A., & Isaia, E. (2011). The minimum density power divergence approach in building robust regression models. *Informatica, 22*, 43–56.

Ferland, R., Latour, A., & Oraichi, D. (2006). Integer-valued GARCH process. *Journal of Time Series Analysis, 27*, 923–942.

Fokianos, K., Rahbek, A., & Tjøstheim, D. (2009). Poisson autoregression. *Journal of American Statistical Association, 104*, 1430–1439.

Fox, A. J. (1972). Outliers in time series. *Journal of the Royal Statistical Society, Series B: Statistical Methodology, 34*, 350–363.

Fried, R., Agueusop, I., Bornkamp, B., Fokianos, K., Fruth, J., & Ickstadt, K. (2015). Retro spective bayesian outlier detection in INGARCH series. *Statistics and Computing, 25*, 365–374.

Fujisawa, H., & Eguchi, S. (2006). Robust estimation in the normal mixture model. *Journal of Statistical Planning & Inferences, 136*, 3989–4011.

Gerstenberger, M., Wiemer, S., Jones, L., & Reasenberg, P. (2005). Real-time forecasts of tomorrow's earthquakes in California. *Nature, 435*, 328–331.

Goebel, T., Weingarten, M., Chen, X., Haffener, J., & Brodsky, E. (2017). The 2016 Mw5.1 Fairview, Oklahoma earthquakes: Evidence for long-range poroelastic triggering at > 40 km from fluid disposal wells. *Earth and Planetary Science Letters, 472*, 50–61.

Gombay, E. (2008). Change detection in autoregressive time series. *Journal of Multivariate Analysis, 99*, 451–464.

Hong, C., & Kim, Y. (2001). Automatic selection of the tuning parameter in the minimum density power divergemce estimation. *Journal of the Korean Statistical Society, 30*, 453–465.

Jazi, M., Jones, G., & Lai, C. (2012). Integer valued AR(1) with geometric innovations. *Journal of the Iranian Statistical Society, 11*, 173–190.

Jo, M., & Lee, S. (2023). Inference for integer valued AR model with zero inflated poisson innovations. *Journal of the Korean Data & Information Science Society, 34*, 1031–1039.

Jung, R., & Tremayne, A. (2006). Coherent forecasting in integer time series models. *International Journal of Forecasting, 22*, 223–238.

Kim, B., & Lee, S. (2020). Robust estimation for general integer-valued time series models. *Annals of the Institute of Statistical Mathematics, 72*, 1371–1396.

Kim, B., & Lee, S. (2024). Robust estimation for general integer-valued time series models based on the exponential-polynomial divergence. *Journal of Statistical Computation and Simulation, 94*, 1300–1316.

Lambert, D. (1992). Zero inflated poisson regression, with an application to defects in manufacturing. *Technometrics, 34*, 1–14.

Lee, S. (2020). Location and scale-based CUSUM test with application to autoregressive models. *Journal of Statistical Computation and Simulation, 90*, 2309–2328.

Lee, S., Ha, J., Na, O., & Na, S. (2003). The CUSUM test for parameter change in time series models. *Scandinavian Journal of Statistics, 30*, 781–796.

Lee, S., & Jo, M. (2023). Bivariate random coefficient integer-valued autoregressive models: Parameter estimation and change point test. *Journal of Time Series Analysis, 44*, 644–666.

Lee, S., & Jo, M. (2023). Robust estimation for bivariate integer valued autoregressive models based on minimum density power divergence. *Journal of Statistical Computation and Simulation, 93*, 3156–3184.

Lee, S., & Kim, D. (2023). Multiple values-inflated exponential family integer-valued GARCH models. *Journal of Statistical Computation and Simulation, 93*, 1297–1317.

Lee, S., Kim, D., & Kim, B. (2023). Modeling and inference for multivariate time series of counts based on the INGARCH scheme. *Computational Statistics & Data Analysis, 177*, 107579.

Lee, S., & Na, O. (2005). Test for parameter change based on the estimator minimizing density-based divergence measures. *Annals of the Institute of Statistical Mathematics, 57*, 553–573.

Lee, Y., & Lee, S. (2019). CUSUM test for general nonlinear integer-valued GARCH models. *Annals of the Institute of Statistical Mathematics, 71*, 1033–1057.

Livio, T., Khan, N., Bourguignon, M., and Bakouch, H. (2018). An INAR(1) model with poisson-lindley innovations. Economics Bulletin, 38.

Lu, F., & Wang, D. (2022). A new estimation for INAR (1) process with poisson distribution. *Computational Statistics, 37*, 1185–1201.

McKenzie, E. (1985). Some simple models for discrete variate time series. *Journal of American Water Resource Association, 21*, 645–650.

Mohammadi, Z., Sajjadnia, Z., Bakouch, H., & Sharafi, M. (2022). Zero-and-one inflated poisson-lindley INAR(1) process for modelling count time series with extra zeros and ones. *Journal of Statistical Computation and Simulation, 92*, 2018–2040.

Mohammadpour, M., Bakouch, H., & Shirozhan, M. (2018). Poisson-lindley INAR(1) model with applications. *Brazilian Journal of Probability and Statistics, 32*, 262–280.

Mukherjee, T., Mandal, A., & Basu, A. (2019). The B-exponential divergence and its generalizations with applications to parametric estimation. *Statistical Methods & Applications, 28*, 241–257.

Pedeli, X., & Karlis, D. (2011). A bivariate INAR(1) process with application. *Statistical Modelling, 11*, 325–349.

Pedeli, X., & Karlis, D. (2013). On estimation of the bivariate Poisson INAR process. *Communications in Statistics - Simulation and Computation, 42*, 514–533.

Singh, P., Mandal, A., & Basu, A. (2021). Robust inference using the exponential-polynomial divergence. *Journal of Statistical Theory and Practice, 15*, 1–22.

Thomas, H. (1994). The uniform california earthquake rupture forecast,. Version 1.

Toma, A., & Broniatowski, M. (2011). Dual divergence estimators and tests: Robustness results. *Journal of Multivariate Analysis, 102*, 20–36.

Tsay, R., Peña, D., & Pankratz, A. (2000). Outliers in multivariate time series. *Biometrika, 87*, 789–804.

Warwick, J. (2005). A data-based method for selecting tuning parameters in minimum distance estimators. *Computational Statistics & Data Analysis, 48*, 571–585.

Warwick, J., & Jones, M. C. (2005). Choosing a robustness tuning parameter. *Journal of Statistical Computation and Simulation, 75*, 581–588.

Weiß, C. H. (2015). A poisson INAR(1) model with serially dependent innovations. *Metrika, 78*, 829–851.

Weiß, C. H. (2018). An Intorduction to Discrete-Valued Time Series. Wiley, New York., ISBN-13:978-1119096962.