Article

Monitoring the Zero-Inflated Time Series Model of Counts with Random Coefficient

Cong Li [1], Shuai Cui [1] and Dehui Wang [1,2,*]

1 School of Mathematics, Jilin University, Changchun 130012, China; li_cong@jlu.edu.cn (C.L.); cuishuai@jlu.edu.cn (S.C.)
2 School of Economics, Liaoning University, Shenyang 110036, China
* Correspondence: wangdh@jlu.edu.cn

Abstract: In this research, we consider monitoring mean and correlation changes from zero-inflated autocorrelated count data based on the integer-valued time series model with random survival rate. A cumulative sum control chart is constructed due to its efficiency, the corresponding calculation methods of average run length and the standard deviation of the run length are given. Practical guidelines concerning the chart design are investigated. Extensive computations based on designs of experiments are conducted to illustrate the validity of the proposed method. Comparisons with the conventional control charting procedure are also provided. The analysis of the monthly number of drug crimes in the city of Pittsburgh is displayed to illustrate our current method of process monitoring.

Keywords: CUSUM control chart; INAR-type time series; statistical process monitoring; random survival rate; zero-inflation

1. Introduction

This work is motivated by an empirical analysis and process control of a monthly drug crime series, which contains excess zeros (over 40%) and shows clear serial dependence (see Section 5 for more details). To solve this problem, an appropriate integer-valued model is selected, further, control charts based on this model are developed. Among kinds of integer-valued models, a specific kind featured with first-order integer-valued autoregressive (INAR(1)) models plays an very important role and has been widely studied in the literature. In reality, serial dependence among the count data have been demonstrated to arise extensively in practice, typical examples are infectious disease counts, defect counts and unemployment counts, etc. These data are important indicators of the epidemic study, quality control and economics analysis, and the process monitoring is essential to detect the shifts in them.

The first INAR(1) model proposed by Al-Osh and Alzaid [1] is in the following form

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

where the binomial thinning operator "$\circ$" is defined by Steutel and Van Harn [2], $\alpha \circ X_{t-1} = \sum_{i=1}^{X_{t-1}} Y_i$, $\alpha \in (0,1)$, $\{Y_i\}_N$ is a sequence of independent and identically distributed (i.i.d.) random variables with Bernoulli(α) distribution; and $\{\varepsilon_t\}_N$ is a sequence of i.i.d. random variables, independent of all $\{Y_i\}$. The INAR(1) model is currently applied in various kinds of real-world problems because of its good interpretability. As one example, we let X_t represent the number of patients of an infectious disease in a community at time t, ε_t the number of new patients entering the community at time t, and suppose each patient at time $t-1$ survives at time t with survival probability α. As for the crime data, $\alpha \circ X_{t-1}$ can be considered as the number of re-offendings provoked by X_{t-1} with probability α. Depending on the nature of this kind of observed data, the INAR(1) models have been modified and

generalized with respect to their orders (Ristić and Nastić [3], Nastić, Laketa and Ristić [4]), dimensions (Pedeli and Karlis [5], Khan, Cekim and Ozel [6]), marginal distributions (Alzaid and Al-Osh [7], Alzaid and Al-Osh [8], Jazi, Jones and Lai [9], Ristić, Nastić and Bakouch [10], Barreto-Souza [11]), thinning operators (Ristić, Bakouch and Nastić[12], Liu and Zhu [13]), and mixed models (Ristić and Nastić [3], Li, Wang and Zhang [14], Orozco, Sales, Fernández and Pinho [15]). For more literature, we refer to review papers (Weiß [16], Scotto, Weiß and Gouveia [17]). Differing with the models that based on a fixed survival rate α, Zheng, Basawa and Datta [18] proposed the random coefficient INAR(1) model, supposing that the parameter α may be affected by various environmental factors, and could vary randomly over time. Some of the generalizations of random coefficient INAR models can also be found in Kang and Lee [19] and Zhang, Wang and Zhu [20]. In particular, considering both random survival probability and zero-inflation phenomenon, Bakouch, Mohammadpour and Shirozhan [21] purposed a zero-inflated geometric INAR(1) time series with random coefficient (short for the ZIGINAR$_{RC}$(1) process). The ZIGINAR$_{RC}$(1) model has simple structure and good properties, which turns out to be the best fit for the real data studied by us.

As the serial dependence shows big influence on the performance of the control chart, the traditional control charts under the assumption of independent observations are not appropriate in many cases. Therefore, the monitoring of INAR(1) models has received much attention. The related research includes but not limited to the control charts for the generally developed Poisson INAR(1) models (Weiß [22], Weiß and Testik [23], Weiß and Testik [24], Yontay, Weiß, Testik and Bayindir [25]), for zero-inflated or zero-deflated INAR(1) models (Rakitzis, Weiß and Castagliola [26], Li, Wang and Sun [27], Fernandes, Bourguignon and Ho [28]), for the mixed INAR(1) model (Sales, Pinho, Vivacqua and Ho [29]), etc. While, to the best of our knowledge, methods for monitoring the zero-inflated INAR(1) model with random coefficient have not been studied in the literature so far, which is exactly what we are going to explore. As cumulative sum (CUSUM) control charts are known to be sensitive in detecting small shifts, we study the performance of the CUSUM chart for monitoring ZIGINAR$_{RC}$(1) process. We investigate the practical guidelines for the statistical design and the methods for evaluating the chart performance. Besides monitoring mean shifts of the ZIGINAR$_{RC}$(1) model, our scope is also to monitor correlation shifts in the model. Meanwhile, we compare the performance of the CUSUM chart with the conventional Shewhart chart.

The rest of the article is outlined as follows. The ZIGINAR$_{RC}$(1) process and some properties of this process are introduced in Section 2. In Section 3, we present the monitoring procedure to detect the mean and correlation shifts of the process. Extensive computation results are discussed in Section 4. In Section 5, the applicability of the process monitoring is investigated using the monthly number of drug crimes in Pittsburgh. Finally conclusions and possible future lines of research are shown in Section 6.

2. The ZIGINAR$_{RC}$(1) Process

A randomized binomial thinning operation in Bakouch, Mohammadpour and Shirozhan [21] is defined by

$$
\alpha_t \circ X = \begin{cases} \alpha \circ X, & \text{with probability } 1 - \beta, \\ 0, & \text{with probability } \beta, \end{cases}
$$

where $\alpha, \beta \in (0,1)$, α_t is a binary random variable independent of discrete random variable X, $P(\alpha_t = 0) = \beta = 1 - P(\alpha_t = \alpha)$.

Based on the definition of the randomized binomial thinning operation, the ZIGINAR$_{RC}$(1) model $\{X_t\}$ presented by Bakouch, Mohammadpour and Shirozhan [21], is given by

$$
X_t = \alpha_t \circ X_{t-1} + \varepsilon_t, \quad t = 1, 2, \cdots,
$$

where the marginal distribution is a zero-inflated Geometric distribution (denoted as ZIG(p, θ)), $P(X_t = 0) = p + (1-p)/(1+\theta)$, $P(X_t = j) = (1-p)\theta^j/(1+\theta)^{j+1}$, $j = 1, 2, \cdots$, and $p \in (0,1)$, $\theta > 0$. $\{\varepsilon_t\}_N$ is independent of the past of the solution $\{X_s; s < t\}$ and the binary sequence $\{\alpha_t\}$, parameters are also constrained by the condition $p/(\beta + p(1-\beta)) < \alpha < 1$.

The ZIGINAR$_{RC}$(1) process is quite suitable for modelling some real-life phenomena in which counted events may survive or vanish with the random survival probability α_t. Such series are studied in Section 5 with an example of the counts of the drug crimes, where the re-offending rate may be affected by public security situation and financial situation. The mean, variance, and first-order autocorrelation function of the process are, respectively,

$$\mu_X \triangleq E(X_t) = (1-p)\theta, \quad \sigma_X^2 \triangleq \text{Var}(X_t) = (1-p)\theta[(1+p)\theta + 1],$$

$$\rho_X \triangleq Corr(X_t, X_{t+1}) = \alpha(1-\beta).$$

Obviously the process is characterized by the property of overdispersion, i.e., the variance greater than the expectation. Figure 1 shows some sample paths of simulated ZIGINAR$_{RC}$(1) processes for $\theta = 1, 3, 5$; $p = 0.2, 0.5$; $\alpha = 0.5, 0.8$ and $\beta = 0.3, 0.8$. As we can see, the model has larger process mean with larger θ, and larger percentage of zeros with larger p.
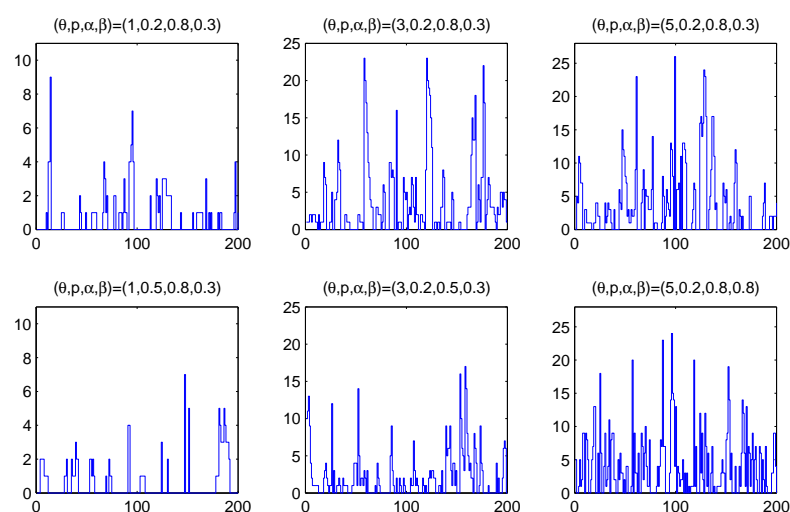
Figure 1. Sample path of ZIGINAR$_{RC}$(1) processes for $\theta = 1, 3, 5$, $p = 0.2, 0.5$, $\alpha = 0.5, 0.8$ and $\beta = 0.3, 0.8$.

Following Theorem 2.1 in Bakouch, Mohammadpour and Shirozhan [21], the ZIGINAR$_{RC}$(1) model has a unique, strictly stationary solution given by

$$\sum_{i=1}^{\infty} \left(\prod_{l=0}^{i-1} \alpha_{t-l} \right) \circ \varepsilon_{t-i} + \varepsilon_t.$$

Furthermore, the probability mass function of $\{\varepsilon_t\}_N$ is

$$P(\varepsilon_t = j) = \frac{p}{\beta + p(1-\beta)} I_{(j)} + \frac{(1-p)(1-\alpha)}{1 - \alpha[\beta + p(1-\beta)]} \frac{\theta^j}{(1+\theta)^{j+1}}$$
$$+ \frac{(1-p)(1-\beta)[\alpha(\beta + p(1-\beta)) - p]}{(1 - \alpha[\beta + p(1-\beta)])(\beta + p(1-\beta))}$$
$$\times \frac{[\alpha\theta(\beta + p(1-\beta))]^j}{[1 + \alpha\theta(\beta + p(1-\beta))]^{j+1}}, \qquad j = 0, 1, 2, \cdots.$$

where $I_{(j)} = 1$ for $j = 0$ and 0 else. It can be deduced that the innovation series $\{\varepsilon_t\}_N$ is a mixture of three random variables, a degenerate distribution at 0, Geometric$(\theta/(1+\theta))$ and Geometric$(\alpha\theta(\beta + p(1-\beta))/(1 + \alpha\theta(\beta + p(1-\beta))))$ distributions with three different mixing portions. The following form is the transition probability of the process $\{X_t\}_{N_0}$

$$P(X_t = j | X_{t-1} = i) = \beta P(\varepsilon_t = j) + (1-\beta) \sum_{l=0}^{min(i,j)} \binom{i}{l} \alpha^l (1-\alpha)^{i-l} P(\varepsilon_t = j - l),$$

$$i, j = 0, 1, \cdots.$$

Some other important probabilistic properties of the process, like spectral density, multi-step conditional mean and variance, extreme order statistics, distributional properties of length of run of zeros, have also been discussed in Bakouch, Mohammadpour and Shirozhan [21]. Furthermore, the unknown parameters of the model could be estimated by conditional least squares or maximum likelihood methods.

3. Monitoring Procedure

In this section, we present a CUSUM chart for monitoring the ZIGINAR$_{RC}$(1) process. As this process is used to fit the number of crimes, an increase in the process mean usually means a deteriorating public security environment, and an increase in the process correlation usually means more re-offendings. Thus, our purpose is to detect the increasing of both mean shifts and correlation shifts in the ZIGINAR$_{RC}$(1) process. According to the model properties, the process mean is affected by the parameters θ and p, the correlation is affected by the parameters α and β. Let θ_0, p_0, α_0 and β_0 (θ_1, p_1, α_1 and β_1) denote the in-control (out-of-control) parameters of the processes, and μ_0, σ_0, ρ_0 (μ_1, σ_1, ρ_1) be the corresponding in-control (out-of-control) process mean, standard deviation and first-order correlation.

The CUSUM charts are commonly used charts in statistical process control, which were first proposed by Page [30]. The essential assumption underlying the design of CUSUM charts is that the process observations are independent (Montgomery [31], Alencar, Ho and Albarracin [32], Bourguignon, Medeiros, Fernandes and Ho [33]). While the violation of this major assumption seriously affects the monitoring performance of the charts (Harris and Ross [34], Triantafyllopoulos and Bersimis [35], Albarracin, Alencar and Ho [36]). Some authors have studied the performance of CUSUM charts for some integer-valued models (Weiß and Testik [23], Weiß and Testik [24], Yontay, Weiß, Testik and Bayindir [25], Rakitzis, Weiß and Castagliola [26], Li, Wang and Sun [27], Lee and Kim [37], Lee, Kim and Kim [38]).

Scheme (The ZIGINAR$_{RC}$(1) CUSUM chart). *Let $\{X_t\}_{N_0}$ be a stationary ZIGINAR$_{RC}$(1) process, the CUSUM statistics C_t is defined as:*

$$C_t = \max(0, \; X_t - k + C_{t-1}), \; t = 1, 2, \cdots,$$

where k is a positive integer constant representing the reference ($k \geqslant \mu_0$). This chart is said to be out-of-control when C_t falls outside the control limit h ($h \in N$), that is, $C_t > h$.

The initial value of the CUSUM statistics is set equal to the integer constant c_0, i.e., $C_0 = c_0$ with $c_0 < h$. The performance evaluation of this chart is accomplished based on the average run length (ARL) measures, which is defined as the average number of points to be plotted on the chart until the first out-of-control signal triggers. As $\{X_t, C_t\}_{t \in N_0}$ of the ZIGINAR$_{RC}$(1) process is a bivariate Markov chain, the Markov chain approach proposed by Brook and Evans [39] is adapted to evaluate the exact ARLs. Though this method has been described in detail in the relevant literature by Weiß [22], Weiß and Testik [23] and Weiß and Testik [24], we briefly introduce this method here for completeness. The reachable control region ($\mathcal{CR}$) of $\{X_t, C_t\}_{N_0}$ is given by

$$\mathcal{CR} \triangleq \{(n,i) \in N_0 \times \{0, \cdots, h\} \,|\, \max(0, n - k + i) \in \{0, \cdots, h\}\}$$
$$= \{(n,i) | i \in \{0, \cdots, h\}, \; n \in \{\max(0, i + k - h), \cdots, i + k\}\}.$$

Obviously $\mathcal{CR}$ has a finite number of elements and could be ordered in a certain manner. The transition probability matrix of $\{X_t, C_t\}_{N_0}$ is $\boldsymbol{Q}^\top \triangleq (p(n,j|m,i))_{(n,j),(m,i) \in \mathcal{CR}}$,

$$p(n,j|m,i) \triangleq P(X_t = n, C_t = j | X_{t-1} = m, C_{t-1} = i) = I_{(j - \max(0, n - k + i))} P(X_t = n | X_{t-1} = m).$$

The initial probabilities are

$$p(n,j|c_0) \triangleq P(X_1 = n, C_1 = j | C_0 = c_0) = I_{(j - \max(0, n - k + c_0))} P(X_1 = n).$$

The conditional probability that the run length of $\{X_t, C_t\}_{N_0}$ equals r is defined by

$$p_{m,i}(r) \triangleq P((X_{r+1}, C_{r+1}) \notin \mathcal{CR}, (X_r, C_r), \cdots, (X_2, C_2) \in \mathcal{CR} | (X_1, C_1) = (m, i)),$$

where $(m,i) \in \mathcal{CR}$. Let the vector $\boldsymbol{\mu}_{(k)}$ denote the k-th factorial moments that $(\boldsymbol{u}_{(k)})_{m,i} \triangleq \sum_{r=1}^{\infty} r_{(k)} p_{m,i}(r)$ where $k \geqslant 1$ and $r_{(k)} \triangleq r \times (r-1) \times \cdots \times (r - k + 1)$. Then

$$p_{m,i}(r) = \sum_{(n,j) \in \mathcal{CR}} p_{n,j}(r-1) \times p(n,j|m,i),$$

$$(\boldsymbol{u}_{(1)})_{m,i} = 1 + \sum_{(n,j) \in \mathcal{CR}} p(n,j|m,i) \times (\boldsymbol{u}_{(1)})_{n,j}, \; i.e., \; (\boldsymbol{I} - \boldsymbol{Q})\boldsymbol{u}_{(1)} = \boldsymbol{1}.$$

The ARL is obtained as

$$\text{ARL} = \sum_{(m,i) \in \mathcal{CR}} (\boldsymbol{u}_{(1)})_{m,i} \times p(m,i|c_0).$$

For simplicity we do not repeat the proof methods, see Weiß [22], Weiß and Testik [23] and Weiß and Testik [24] for more details. It is expected that an efficient chart possesses a large in-control ARL (denoted as ARL_0) and a small out-of-control ARL. Along with the ARL, we also assess the performance of the charts through the standard deviation of run length (SDRL) suggested by Weiß [22]. The SDRL of the $\text{ZIGINAR}_{\text{RC}}(1)$ CUSUM chart could again be computed efficiently by applying the Markov chain method. The second order factorial moments $\boldsymbol{u}_{(2)}$ can be determined recursively from the relation $(\boldsymbol{I} - \boldsymbol{Q})\boldsymbol{u}_{(2)} = 2\boldsymbol{Q}\boldsymbol{u}_{(1)}$. Then the SDRL is

$$\text{SDRL} = \sqrt{\sum_{(m,i) \in \mathcal{CR}} ((\boldsymbol{u}_{(2)})_{m,i} + (\boldsymbol{u}_{(1)})_{m,i}) \times p(m,i|c_0) - \text{ARL}^2}.$$

To implement the proposed monitoring scheme, the chart design pairs (h, k) need to be designed in advance. Generally, a fixed ARL_0 value is set to be the target value, and (h, k) is set accordingly. Some guidelines for the choices of them will be given in the next section.

4. Computation Results

In this section, we evaluate the $\text{ZIGINAR}_{\text{RC}}(1)$ CUSUM chart performance basing on extensive numerical experiments and presume that the parameters in this model have already been known. In practice, the in-control parameters need to be estimated from the data, as shown in the next section. We search for possible chart designs (integer (h, k) pairs) in order to adjust the ARL_0 close to the target value. Here the target ARL_0 value is set to be 370, which is commonly used in the statistical process monitoring domain. Meanwhile, the values of ARL and SDRL are calculated accurately by the Markov chain method, and we only show the results with two decimal places for simplicity. We first compute ARL_0

and SDRL$_0$ of the CUSUM chart for different in-control process parameters and initial values in Table 1. The process parameters are: $\theta_0 = \{1, 5\}$; $p_0 = \{0.1, 0.3\}$; $\alpha_0 = \{0.5, 0.8\}$; $\beta_0 = \{0.5, 0.8\}$. Furthermore, initial values are $c_0 = \{0, 3, 6\}$. These chosen parameters could cover a broad range of different scenarios. Based on the results in Table 1, three important conclusions can be derived. First, it can be observed that when c_0 takes smaller value, the deviation of ARL$_0$ and SDRL$_0$ is small. When c_0 takes larger value, there might be a situation where the value of SDRL$_0$ is significantly greater than the value of ARL$_0$ (for example, ARL$_0 = 409.42$, SDRL$_0 = 442.13$ under $(\theta_0, p_0, \alpha_0, \beta_0, c_0) = (1, 0.3, 0.5, 0.8, 6)$). Thus, we assume that $c_0 = 0$ in the following studies to get better robust. Second, as the differences of the values between ARL and SDRL are small when $c_0 = 0$, we only use ARL as the measure in the following computations to save space. Last, the parameter θ_0 shows a great influence on the selection of control designs (h, k), with a larger θ_0 comes a larger pair of (h, k).

Table 1. ARL$_0$ and SDRL$_0$ of the CUSUM chart for various θ_0, p_0, α_0, β_0 and c_0.

(θ_0, p_0, α_0, β_0)	(h, k)	(c_0, ARL$_0$, SDRL$_0$)	(c_0, ARL$_0$, SDRL$_0$)	(c_0, ARL$_0$, SDRL$_0$)
(1, 0.1, 0.5, 0.5)	(9, 2)	(0, 340.55, 339)	(3, 336.84, 338.98)	(6, 322.88, 338.52)
(1, 0.1, 0.5, 0.8)	(8, 2)	(0, 428.55, 427.38)	(3, 423.5, 427.34)	(6, 398.83, 426.33)
(1, 0.1, 0.8, 0.5)	(12, 2)	(0, 368.36, 366.45)	(3, 365.76, 366.43)	(6, 358.76, 366.3)
(1, 0.1, 0.8, 0.8)	(9, 2)	(0, 428.69, 427.42)	(3, 424.79, 427.4)	(6, 408.35, 426.91)
(1, 0.3, 0.5, 0.5)	(8, 2)	(0, 385.69, 384.65)	(3, 381.9, 384.62)	(6, 365.66, 384.11)
(1, 0.3, 0.5, 0.8)	(7, 2)	(0, 444.16, 443.51)	(3, 438.89, 443.47)	(6, 409.42, 442.13)
(1, 0.3, 0.8, 0.5)	(10, 2)	(0, 359.91, 358.61)	(3, 357.26, 358.6)	(6, 349.32, 358.44)
(1, 0.3, 0.8, 0.8)	(8, 2)	(0, 469.37, 468.53)	(3, 465.3, 468.51)	(6, 446.23, 467.94)
(5, 0.1, 0.5, 0.5)	(60, 6)	(0, 379.61, 371.51)	(3, 379.07, 371.51)	(6, 378.25, 371.51)
(5, 0.1, 0.5, 0.8)	(49, 6)	(0, 376.02, 369.17)	(3, 375.4, 369.17)	(6, 374.38, 369.16)
(5, 0.1, 0.8, 0.5)	(75, 6)	(0, 371.37, 363.76)	(3, 370.91, 363.76)	(6, 370.28, 363.75)
(5, 0.1, 0.8, 0.8)	(54, 6)	(0, 378.57, 372.12)	(3, 378.01, 372.12)	(6, 377.14, 372.11)
(5, 0.3, 0.5, 0.5)	(46, 6)	(0, 383.15, 379.48)	(3, 382.64, 379.47)	(6, 381.86, 379.47)
(5, 0.3, 0.5, 0.8)	(38, 6)	(0, 386.29, 383.42)	(3, 385.68, 383.42)	(6, 384.7 , 383.42)
(5, 0.3, 0.8, 0.5)	(59, 6)	(0, 378.46, 374.67)	(3, 378.04, 374.67)	(6, 377.45, 374.67)
(5, 0.3, 0.8, 0.8)	(42, 6)	(0, 379.79, 377.04)	(3, 379.26, 377.04)	(6, 378.45, 377.03)

Due to its simplicity, the conventional Shewhart chart is very popular in monitoring the process shifts. The upper limit for the Shewhart chart is denoted as *UCL*. For observations, when the value of the process $\{X_t\}_{N_0}$ exceeds the threshold value *UCL* ($X_t > UCL$), a fault is declared. Figures 2 and 3 investigate the CUSUM method preliminarily by comparing it with the Shewhart method. In both of these figures, we assume that the in-control parameters are $\theta_0 = 2$, $p_0 = 0.2$, $\alpha_0 = 0.5$ and $\beta_0 = 0.5$, which are selected based on the real drug crime data in Section 5. According to these parameters, the CUSUM chart designs can be determined, respectively, as $h = 31, k = 2$ (corresponding ARL$_0 = 383.74$); $h = 19, k = 3$ (ARL$_0 = 396.12$); $h = 14, k = 4$ (ARL$_0 = 373.27$); $h = 11, k = 5$ (ARL$_0 = 370.77$); $h = 9, k = 6$ (ARL$_0 = 394.03$). Furthermore, the Shewhart chart limit $UCL = 13$ (ARL$_0 = 381.31$) can be used. It should be noted that two types of changes are considered in Figure 2, which both lead to the upward mean shifts. The first type of changes occurs only in the parameter θ, with other parameters invariant, the results are listed in Figure 2a. Similarly, the second type of changes occurs only in the parameter p, with other parameters invariant, the results are in Figure 2b. From Figure 2, we can conclude that the CUSUM chart with the design $h = 31, k = 2$ outperforms the other

CUSUM charts under most shifts, while the Shewhart chart performs worst among them. For the upward correlation shift scenarios, ARL values under two types of the parameter changes are displayed in Figure 3. The first one considers changes only in the parameter α, and the second one considers changes only in the parameter β. For each scenario, the Shewhart chart performs increase ratio of ARL with the increase of first-order correlation ρ_X, and the CUSUM chart has the better behaviour in the figure. In a comprehensive view, the conventional Shewhart chart is insensitive for upward mean shifts caused by changes in parameter p, and fails to detect shifts in the correlation. While the proposed CUSUM chart could overcome these limitations and has superiorities in various coefficient shifts compared with the Shewhart chart. From the figures, we can also conclude that the smaller the value of k, the more sensitive the CUSUM chart is. As the constraint $k \geqslant \mu_0$ is required to make the chart reasonable, it is natural to recommend $k = \lceil \mu_0 \rceil$ (the smallest integer no less than μ_0), then we aim to select the value of h such that ARL_0 is close to 370. Now the computations of the CUSUM chart are extended to general cases with designs of experiments as follow.
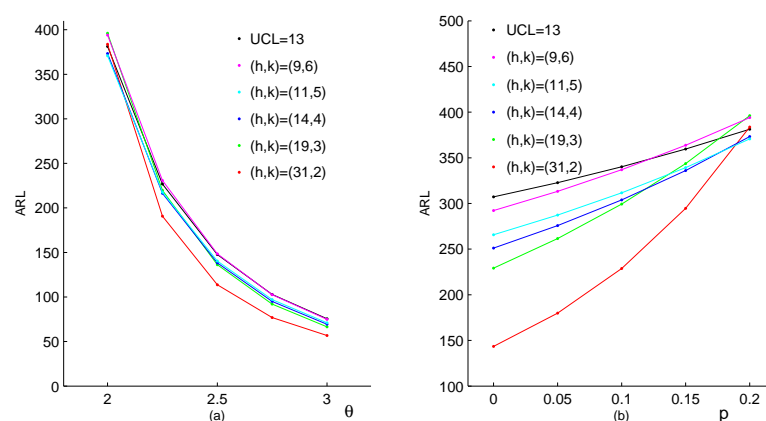
Figure 2. The performance of the Shewhart chart and the CUSUM chart to detect an increase in process mean, (**a**) changes only in θ, (**b**) changes only in p.
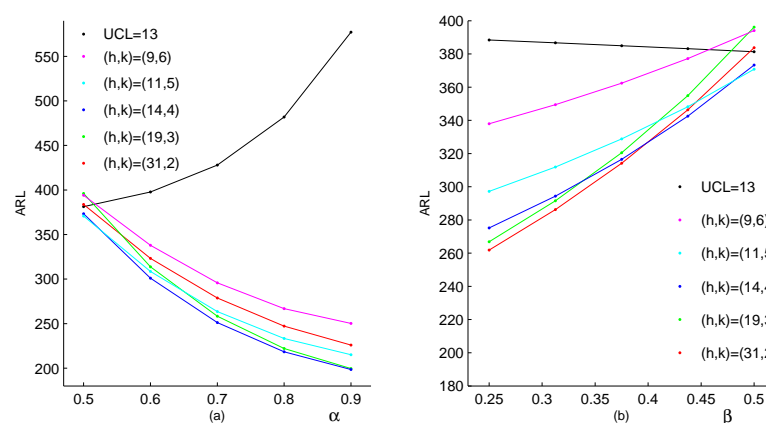
Figure 3. The performance of the Shewhart chart and the CUSUM chart to detect an increase in process first-order correlation, (**a**) changes only in α, (**b**) changes only in β.

In Tables 2–4, we focus on situations that there are increasing shifts in process mean, and the correlation remains the same. Each in-control parameter has three levels: $\theta_0 = \{1, 3, 5\}$; $p_0 = \{0.1, 0.2, 0.3\}$; $\alpha_0 = \{0.5, 0.6, 0.7\}$; $\beta_0 = \{0.5, 0.6, 0.7\}$. We consider the case when the changes only occur in θ, this is the most common case. The out-of-control process mean is $\mu_1 = \mu_0 + \delta\sigma_0$, the shift size δ considers potential values in set $\{0.5, 1, 1.5, 6\}$. The usual relative deviation (in %) in ARL is defined as $\text{dev}^{(\%)} = 100\% \times$

(ARL-ARL$_0$)/ARL$_0$ (Weiß and Testik [24]). From Tables 2–4, we can conclude that the ZIGINAR$_{RC}$(1) CUSUM chart performs quite well in detecting upward mean shifts for all scenarios. For the small shift size of $\delta = 0.5$, the CUSUM chart is efficient with the minimum 86.38% drop of ARL and the maximum 91.23% drop of ARL. Take larger δ for another illustration ($\delta = 1$), the drop of ARL is at least 92.81%, and up to 96.03% at most. It can also be obtained that δ has to be at least 6 to get an immediate signal with the out-of-control ARL closer to 1. In addition, extensive computation results show that the in-control parameters ($\theta_0, p_0, \alpha_0, \beta_0$) have little effect on the better performance of the CUSUM chart to detect the mean shifts.

Table 2. ARL profiles of the CUSUM chart versus mean shifts under $\theta_0 = 1$.

Process Parameters						ARL (dev$^{(\%)}$)				
p_0	α_0	β_0	μ_0	h	k	$\delta = 0$	$\delta = 0.5$	$\delta = 1$	$\delta = 1.5$	$\delta = 6$
0.1	0.5	0.5	0.9	22	1	348.22	38.62 (−88.91%)	19.31 (−94.45%)	12.94 (−96.28%)	3.44 (−99.01%)
0.1	0.5	0.6	0.9	21	1	357.74	36.77 (−89.72%)	18.28 (−94.89%)	12.21 (−96.59%)	3.21 (−99.1%)
0.1	0.5	0.7	0.9	20	1	365.71	34.91 (−90.45%)	17.27 (−95.28%)	11.51 (−96.85%)	3 (−99.18%)
0.1	0.6	0.5	0.9	24	1	370.79	42.46 (−88.55%)	21.23 (−94.27%)	14.23 (−96.16%)	3.78 (−98.98%)
0.1	0.6	0.6	0.9	22	1	358.58	38.82 (−89.17%)	19.32 (−94.61%)	12.91 (−96.4%)	3.39 (−99.05%)
0.1	0.6	0.7	0.9	21	1	378.05	36.86 (−90.25%)	18.23 (−95.18%)	12.15 (−96.79%)	3.16 (−99.16%)
0.1	0.7	0.5	0.9	25	1	358.79	44.75 (−87.53%)	22.43 (−93.75%)	15.04 (−95.81%)	3.99 (−98.89%)
0.1	0.7	0.6	0.9	23	1	359.01	40.95 (−88.59%)	20.39 (−94.32%)	13.63 (−96.2%)	3.57 (−99.01%)
0.1	0.7	0.7	0.9	21	1	351.73	37.15 (−89.44%)	18.41 (−94.77%)	12.28 (−96.51%)	3.19 (−99.09%)
0.2	0.5	0.5	0.8	18	1	371.21	37.8 (−89.82%)	18.14 (−95.11%)	12.01 (−96.76%)	3.26 (−99.12%)
0.2	0.5	0.6	0.8	17	1	378.49	35.72 (−90.56%)	17 (−95.51%)	11.22 (−97.04%)	3.01 (−99.2%)
0.2	0.5	0.7	0.8	16	1	382.96	33.57 (−91.23%)	15.88 (−95.85%)	10.45 (−97.27%)	2.77 (−99.28%)
0.2	0.6	0.5	0.8	19	1	360.14	40.29 (−88.81%)	19.42 (−94.61%)	12.87 (−96.43%)	3.49 (−99.03%)
0.2	0.6	0.6	0.8	18	1	382.06	38.15 (−90.01%)	18.17 (−95.24%)	11.99 (−96.86%)	3.2 (−99.16%)
0.2	0.6	0.7	0.8	16	1	346.1	33.8 (−90.23%)	16.06 (−95.36%)	10.58 (−96.94%)	2.81 (−99.19%)
0.2	0.7	0.5	0.8	20	1	349.23	42.95 (−87.7%)	20.77 (−94.05%)	13.77 (−96.06%)	3.71 (−98.94%)
0.2	0.7	0.6	0.8	18	1	338.16	38.56 (−88.6%)	18.49 (−94.53%)	12.21 (−96.39%)	3.25 (−99.04%)
0.2	0.7	0.7	0.8	17	1	365.85	36.24 (−90.09%)	17.17 (−95.31%)	11.29 (−96.91%)	2.98 (−99.19%)
0.3	0.5	0.5	0.7	15	1	376.37	37.77 (−89.96%)	17.57 (−95.33%)	11.54 (−96.93%)	3.25 (−99.14%)
0.3	0.5	0.6	0.7	14	1	374.56	35.37 (−90.56%)	16.31 (−95.65%)	10.67 (−97.15%)	2.97 (−99.21%)
0.3	0.5	0.7	0.7	13	1	368.03	32.86 (−91.07%)	15.04 (−95.91%)	9.81 (−97.33%)	2.71 (−99.26%)
0.3	0.6	0.5	0.7	16	1	369.03	40.69 (−88.97%)	19.04 (−94.84%)	12.51 (−96.61%)	3.5 (−99.05%)
0.3	0.6	0.6	0.7	14	1	323.29	35.58 (−88.99%)	16.61 (−94.86%)	10.89 (−96.63%)	3.03 (−99.06%)
0.3	0.6	0.7	0.7	13	1	328.19	33.03 (−89.94%)	15.25 (−95.35%)	9.96 (−96.97%)	2.75 (−99.16%)
0.3	0.7	0.5	0.7	17	1	360.19	43.81 (−87.84%)	20.59 (−94.28%)	13.53 (−96.24%)	3.75 (−98.96%)
0.3	0.7	0.6	0.7	15	1	335.26	38.66 (−88.47%)	18 (−94.63%)	11.79 (−96.48%)	3.24 (−99.03%)
0.3	0.7	0.7	0.7	14	1	357.18	36.06 (−89.9%)	16.52 (−95.37%)	10.76 (−96.99%)	2.93 (−99.18%)

Table 3. ARL profiles of the CUSUM chart versus mean shifts under $\theta_0 = 3$.

Process Parameters						ARL (dev$^{(\%)}$)				
p_0	α_0	β_0	μ_0	h	k	$\delta = 0$	$\delta = 0.5$	$\delta = 1$	$\delta = 1.5$	$\delta = 6$
0.1	0.5	0.5	2.7	54	3	364.48	38.52 (−89.43%)	19.06 (−94.77%)	12.72 (−96.51%)	3.37 (−99.08%)
0.1	0.5	0.6	2.7	51	3	366.54	36.32 (−90.09%)	17.87 (−95.12%)	11.9 (−96.75%)	3.11 (−99.15%)
0.1	0.5	0.7	2.7	48	3	365.43	34.07 (−90.68%)	16.69 (−95.43%)	11.09 (−96.97%)	2.87 (−99.21%)
0.1	0.6	0.5	2.7	58	3	373.32	41.8 (−88.8%)	20.71 (−94.45%)	13.83 (−96.3%)	3.66 (−99.02%)
0.1	0.6	0.6	2.7	54	3	374.45	38.8 (−89.64%)	19.09 (−94.9%)	12.71 (−96.61%)	3.32 (−99.11%)
0.1	0.6	0.7	2.7	50	3	369.98	35.75 (−90.34%)	17.51 (−95.27%)	11.63 (−96.86%)	3.01 (−99.19%)
0.1	0.7	0.5	2.7	61	3	365.9	44.55 (−87.82%)	22.1 (−93.96%)	14.76 (−95.97%)	3.89 (−98.94%)
0.1	0.7	0.6	2.7	56	3	366.07	40.68 (−88.89%)	20.03 (−94.53%)	13.33 (−96.36%)	3.47 (−99.05%)
0.1	0.7	0.7	2.7	52	3	374.26	37.49 (−89.98%)	18.34 (−95.1%)	12.17 (−96.75%)	3.14 (−99.16%)
0.2	0.5	0.5	2.4	43	3	371.94	36.87 (−90.09%)	17.43 (−95.31%)	11.48 (−96.91%)	3.1 (−99.17%)
0.2	0.5	0.6	2.4	40	3	363.7	34.31 (−90.57%)	16.11 (−95.57%)	10.58 (−97.09%)	2.82 (−99.22%)
0.2	0.5	0.7	2.4	38	3	377.02	32.59 (−91.36%)	15.16 (−95.98%)	9.92 (−97.37%)	2.62 (−99.31%)
0.2	0.6	0.5	2.4	46	3	368.96	39.91 (−89.18%)	18.93 (−94.87%)	12.47 (−96.62%)	3.35 (−99.09%)
0.2	0.6	0.6	2.4	43	3	379.62	37.34 (−90.16%)	17.5 (−95.39%)	11.48 (−96.98%)	3.04 (−99.2%)
0.2	0.6	0.7	2.4	39	3	360.73	33.72 (−90.65%)	15.72 (−95.64%)	10.29 (−97.15%)	2.7 (−99.25%)
0.2	0.7	0.5	2.4	49	3	363.38	43.17 (−88.12%)	20.52 (−94.35%)	13.51 (−96.28%)	3.6 (−99.01%)
0.2	0.7	0.6	2.4	45	3	370.51	39.59 (−89.31%)	18.57 (−94.99%)	12.17 (−96.72%)	3.2 (−99.14%)
0.2	0.7	0.7	2.4	41	3	369.61	35.87 (−90.3%)	16.68 (−95.49%)	10.89 (−97.05%)	2.84 (−99.23%)
0.3	0.5	0.5	2.1	36	3	380.79	36.83 (−90.33%)	16.87 (−95.57%)	11.01 (−97.11%)	3.08 (−99.19%)
0.3	0.5	0.6	2.1	33	3	357.51	33.81 (−90.54%)	15.39 (−95.7%)	10.01 (−97.2%)	2.78 (−99.22%)
0.3	0.5	0.7	2.1	31	3	360.6	31.82 (−91.18%)	14.33 (−96.03%)	9.28 (−97.43%)	2.55 (−99.29%)
0.3	0.6	0.5	2.1	38	3	358.5	39.26 (−89.05%)	18.15 (−94.94%)	11.86 (−96.69%)	3.3 (−99.08%)
0.3	0.6	0.6	2.1	35	3	355.89	36.32 (−89.79%)	16.56 (−95.35%)	10.76 (−96.98%)	2.96 (−99.17%)
0.3	0.6	0.7	2.1	33	3	378.11	34.37 (−90.91%)	15.41 (−95.92%)	9.96 (−97.37%)	2.71 (−99.28%)
0.3	0.7	0.5	2.1	41	3	359.33	43.13 (−88%)	19.97 (−94.44%)	13.03 (−96.37%)	3.58 (−99%)
0.3	0.7	0.6	2.1	38	3	377.93	40.22 (−89.36%)	18.23 (−95.18%)	11.81 (−96.88%)	3.2 (−99.15%)
0.3	0.7	0.7	2.1	34	3	362.28	35.83 (−90.11%)	16.08 (−95.56%)	10.38 (−97.13%)	2.8 (−99.23%)

Table 4. ARL profiles of the CUSUM chart versus mean shifts under $\theta_0 = 5$.

Process Parameters						ARL (dev$^{(\%)}$)				
p_0	α_0	β_0	μ_0	h	k	$\delta = 0$	$\delta = 0.5$	$\delta = 1$	$\delta = 1.5$	$\delta = 6$
0.1	0.5	0.5	4.5	85	5	365.12	38.31 (−89.51%)	18.91 (−94.82%)	12.61 (−96.55%)	3.33 (−99.09%)
0.1	0.5	0.6	4.5	80	5	364.27	36 (−90.12%)	17.67 (−95.15%)	11.76 (−96.77%)	3.07 (−99.16%)
0.1	0.5	0.7	4.5	76	5	371.67	34.09 (−90.83%)	16.64 (−95.52%)	11.04 (−97.03%)	2.86 (−99.23%)
0.1	0.6	0.5	4.5	91	5	370.55	41.46 (−88.81%)	20.5 (−94.47%)	13.68 (−96.31%)	3.61 (−99.03%)
0.1	0.6	0.6	4.5	85	5	374.72	38.61 (−89.7%)	18.95 (−94.94%)	12.6 (−96.64%)	3.29 (−99.12%)
0.1	0.6	0.7	4.5	79	5	373.87	35.71 (−90.45%)	17.43 (−95.34%)	11.56 (−96.91%)	2.99 (−99.2%)
0.1	0.7	0.5	4.5	97	5	373.34	44.81 (−88%)	22.15 (−94.07%)	14.77 (−96.04%)	3.89 (−98.96%)
0.1	0.7	0.6	4.5	89	5	373.74	40.89 (−89.06%)	20.06 (−94.63%)	13.34 (−96.43%)	3.47 (−99.07%)
0.1	0.7	0.7	4.5	81	5	365.37	36.94 (−89.89%)	18.03 (−95.07%)	11.96 (−96.73%)	3.08 (−99.16%)
0.2	0.5	0.5	4	118	4	373.44	45.69 (−87.77%)	24.01 (−93.57%)	16.42 (−95.6%)	4.62 (−98.76%)
0.2	0.5	0.6	4	112	4	369.47	43.25 (−88.29%)	22.65 (−93.87%)	15.45 (−95.82%)	4.3 (−98.84%)
0.2	0.5	0.7	4	107	4	369.02	41.18 (−88.84%)	21.48 (−94.18%)	14.63 (−96.04%)	4.03 (−98.91%)
0.2	0.6	0.5	4	122	4	366.95	47.66 (−87.01%)	25.12 (−93.15%)	17.21 (−95.31%)	4.86 (−98.68%)
0.2	0.6	0.6	4	115	4	365.11	44.74 (−87.75%)	23.47 (−93.57%)	16.03 (−95.61%)	4.47 (−98.78%)
0.2	0.6	0.7	4	110	4	371.48	42.56 (−88.54%)	22.22 (−94.02%)	15.14 (−95.92%)	4.17 (−98.88%)
0.2	0.7	0.5	4	128	4	371.06	50.55 (−86.38%)	26.69 (−92.81%)	18.29 (−95.07%)	5.17 (−98.61%)
0.2	0.7	0.6	4	120	4	372.35	47.07 (−87.36%)	24.71 (−93.36%)	16.88 (−95.47%)	4.71 (−98.74%)
0.2	0.7	0.7	4	112	4	368.48	43.62 (−88.16%)	22.79 (−93.82%)	15.53 (−95.79%)	4.28 (−98.84%)
0.3	0.5	0.5	3.5	80	4	367.77	38.09 (−89.64%)	19.11 (−94.8%)	12.93 (−96.48%)	3.77 (−98.97%)
0.3	0.5	0.6	3.5	75	4	363.52	35.68 (−90.18%)	17.79 (−95.11%)	12 (−96.7%)	3.45 (−99.05%)
0.3	0.5	0.7	3.5	71	4	367.7	33.69 (−90.84%)	16.7 (−95.46%)	11.23 (−96.95%)	3.2 (−99.13%)
0.3	0.6	0.5	3.5	85	4	366.71	40.93 (−88.84%)	20.6 (−94.38%)	13.95 (−96.2%)	4.06 (−98.89%)
0.3	0.6	0.6	3.5	79	4	366.41	37.96 (−89.64%)	18.95 (−94.83%)	12.78 (−96.51%)	3.67 (−99%)
0.3	0.6	0.7	3.5	74	4	372.31	35.4 (−90.49%)	17.55 (−95.29%)	11.8 (−96.83%)	3.35 (−99.1%)
0.3	0.7	0.5	3.5	90	4	364.09	43.96 (−87.93%)	22.17 (−93.91%)	15.02 (−95.87%)	4.35 (−98.81%)
0.3	0.7	0.6	3.5	83	4	368.41	40.36 (−89.04%)	20.16 (−94.53%)	13.59 (−96.31%)	3.88 (−98.95%)
0.3	0.7	0.7	3.5	76	4	365.16	36.7 (−89.95%)	18.2 (−95.02%)	12.23 (−96.65%)	3.46 (−99.05%)

The computation study in Tables 5 and 6 concerns the upward shifts in the process correlation. Two levels are accepted for each in-control parameter: $\theta_0 = \{1, 5\}$; $p_0 = \{0.1, 0.2\}$; $\alpha_0 = \{0.5, 0.6\}$; $\beta_0 = \{0.7, 0.8\}$. Two types of out-of-control pattern are considered here for comprehensive investigation, the first type is that the upward changes only exist in α (shown in Table 5), and the second type is that the downward changes only exist in β (shown Table 6). In Table 5, the shifts in the magnitude $\delta_\alpha = \alpha_1 - \alpha_0$ are from the set $\{0.1, 0.2, 0.3\}$. The results imply that the performance of CUSUM chart fluctuates greatly in detecting correlation shifts caused only by α. To be specific, when δ_α is 0.3, $\mathrm{dev}^{(\%)}$ ranges from -6.81% to -23.68%. Another finding based on the design of experiments in Table 5 is that both a smaller θ_0 and a smaller β_0 could slightly improve detection efficiency, while p_0 and α_0 could not. In Table 6, the shifts magnitude $\delta_\beta = \beta_1 - \beta_0$ are from the set $\{-0.1, -0.2, -0.3\}$. From Table 6, we can see the CUSUM chart is more efficient in detecting the correlation shifts caused by β. As the absolute value of δ_β gets bigger, the decreasing proportion of ARL gradually increased. When $\delta_\beta = -0.3$, $\mathrm{dev}^{(\%)}$ ranges from -21.3% to -40.39%. Meanwhile, we can further conclude that a smaller θ_0 and a larger α_0 often lead to better chart performance, and p_0, β_0 have little influence. Based on all the analysis above in this paragraph, we can further conclude that a smaller θ_0 and a larger value of initial correlation ρ_0 are helpful to detect the correlation shifts. Furthermore, that we cannot get an immediate signal when only correlation shifts occur.

**Table 5.** ARL profiles of the CUSUM chart versus correlation shifts caused by α.

Process Parameters							ARL $(\mathbf{dev}^{(\%)})$			
θ_0	p_0	α_0	β_0	ρ_0	h	k	$\delta_\alpha = 0$	$\delta_\alpha = 0.1$	$\delta_\alpha = 0.2$	$\delta_\alpha = 0.3$
1	0.1	0.5	0.7	0.15	20	1	365.71	339.16 (-7.26%)	316.72 (-13.4%)	298.07 (-18.5%)
1	0.1	0.5	0.8	0.1	19	1	371.95	353.06 (-5.08%)	336.45 (-9.54%)	322.04 (-13.42%)
1	0.1	0.6	0.7	0.18	21	1	378.05	351.73 (-6.96%)	329.83 (-12.75%)	311.98 (-17.48%)
1	0.1	0.6	0.8	0.12	19	1	353.06	336.45 (-4.7%)	322.04 (-8.79%)	309.72 (-12.28%)
1	0.2	0.5	0.7	0.15	16	1	382.96	346.1 (-9.63%)	316.19 (-17.44%)	292.28 (-23.68%)
1	0.2	0.5	0.8	0.1	15	1	384.25	357.92 (-6.85%)	335.41 (-12.71%)	316.44 (-17.65%)
1	0.2	0.6	0.7	0.18	16	1	346.1	316.19 (-8.64%)	292.28 (-15.55%)	273.5 (-20.98%)
1	0.2	0.6	0.8	0.12	15	1	357.92	335.41 (-6.29%)	316.44 (-11.59%)	300.62 (-16.01%)
5	0.1	0.5	0.7	0.15	76	5	371.67	340.88 (-8.28%)	315.5 (-15.11%)	295.25 (-20.56%)
5	0.1	0.5	0.8	0.1	72	5	377.08	355.15 (-5.82%)	336.23 (-10.83%)	320.32 (-15.05%)
5	0.1	0.6	0.7	0.18	79	5	373.87	344.75 (-7.79%)	321.44 (-14.02%)	303.45 (-18.84%)
5	0.1	0.6	0.8	0.12	73	5	367.42	347.51 (-5.42%)	330.75 (-9.98%)	316.97 (-13.73%)
5	0.2	0.5	0.7	0.15	107	4	369.02	353.52 (-4.2%)	339.82 (-7.91%)	328.1 (-11.09%)
5	0.2	0.5	0.8	0.1	103	4	372.99	362.65 (-2.77%)	353.34 (-5.27%)	345.14 (-7.47%)
5	0.2	0.6	0.7	0.18	110	4	371.48	356.88 (-3.93%)	344.32 (-7.31%)	333.97 (-10.1%)
5	0.2	0.6	0.8	0.12	104	4	369.05	359.52 (-2.58%)	351.13 (-4.86%)	343.92 (-6.81%)

**Table 6.** ARL profiles of the CUSUM chart versus correlation shifts caused by β.

Process Parameters							ARL (dev$^{(\%)}$)			
θ_0	p_0	α_0	β_0	ρ_0	h	k	$\delta_\alpha = 0$	$\delta_\alpha = 0.1$	$\delta_\alpha = 0.2$	$\delta_\alpha = 0.3$
1	0.1	0.5	0.7	0.15	20	1	365.71	321.34 (-12.13%)	284.33 (-22.25%)	252.99 (-30.82%)
1	0.1	0.5	0.8	0.1	19	1	371.95	325.82 (-12.4%)	287.74 (-22.64%)	255.77 (-31.24%)
1	0.1	0.6	0.7	0.18	21	1	378.05	324.46 (-14.18%)	281.01 (-25.67%)	245.09 (-35.17%)
1	0.1	0.6	0.8	0.12	19	1	353.06	303.31 (-14.09%)	263.34 (-25.41%)	230.49 (-34.72%)
1	0.2	0.5	0.7	0.15	16	1	382.96	325.3 (-15.06%)	280.11 (-26.86%)	243.87 (-36.32%)
1	0.2	0.5	0.8	0.1	15	1	384.25	324.54 (-15.54%)	278.48 (-27.53%)	241.96 (-37.03%)
1	0.2	0.6	0.7	0.18	16	1	346.1	288.47 (-16.65%)	244.5 (-29.36%)	209.96 (-39.34%)
1	0.2	0.6	0.8	0.12	15	1	357.92	295.66 (-17.39%)	249.23 (-30.37%)	213.36 (-40.39%)
5	0.1	0.5	0.7	0.15	76	5	371.67	322.61 (-13.2%)	282.59 (-23.97%)	249.42 (-32.89%)
5	0.1	0.5	0.8	0.1	72	5	377.08	325.96 (-13.56%)	284.72 (-24.49%)	250.82 (-33.48%)
5	0.1	0.6	0.7	0.18	79	5	373.87	316.52 (-15.34%)	271.1 (-27.49%)	234.34 (-37.32%)
5	0.1	0.6	0.8	0.12	73	5	367.42	310.19 (-15.58%)	265.46 (-27.75%)	229.57 (-37.52%)
5	0.2	0.5	0.7	0.15	107	4	369.02	340.45 (-7.74%)	313.89 (-14.94%)	289.18 (-21.64%)
5	0.2	0.5	0.8	0.1	103	4	372.99	344.5 (-7.64%)	318.09 (-14.72%)	293.53 (-21.3%)
5	0.2	0.6	0.7	0.18	110	4	371.48	337.45 (-9.16%)	306.12 (-17.59%)	277.2 (-25.38%)
5	0.2	0.6	0.8	0.12	104	4	369.05	336.01 (-8.95%)	305.7 (-17.17%)	277.78 (-24.73%)

5. Analyses of Drug Crime Count Time Series

In this section, we present a case study of crime count data in Pittsburgh. The data set contains multiple crime types, such as arson, drink-driving, robbery and so on. Monitoring of crime data is needed not only for early warnings of the organised crime, but also for assessments of the social security environment. For the crime data, the readers can download it from the Forecasting Principles site (http://www.forecastingprinciples.com, accessed on 20 March 2021), or email to the corresponding author to access. The subset we analyse is a monthly drug use count data collected from the 56th police car beat, which contains 144 observations from January 1990 to December 2001. There are 67 zeros in this drug use data (the proportion up to 46.53%), which have the greatest proportion among the other values for the data series. The sample mean, variance and first-order autocorrelation of the data are 1.7153, 6.4289 and 0.3886, respectively, which show strong overdispersion and autocorrelation. The sample path and the histogram of the series are in Figure 4. The histograms of estimated ZIG distribution, estimated Geometric distribution and estimated Poisson distribution are also given in Figure 4b, which indicate that the ZIG marginal is the most appropriate to describe the data. The sample autocorrelation function (ACF) and the sample partial autocorrelation function (PACF) in Figure 5 reveal that the series most likely comes from an AR-type process of order 3. While our intention is to illustrate the implementation of the proposed control chart, we will employ the first-order INAR models that are widely studied and applied in the literature. The consideration of more complex models will be left for future study.
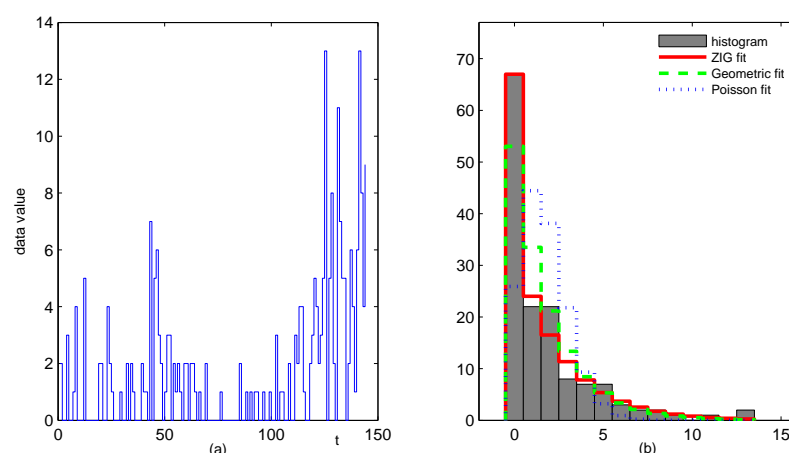
Figure 4. The plots about the zero-inflated drug crime series, (**a**) the sample path, (**b**) the histogram with ZIG fit, Geometric fit and Poisson fit.
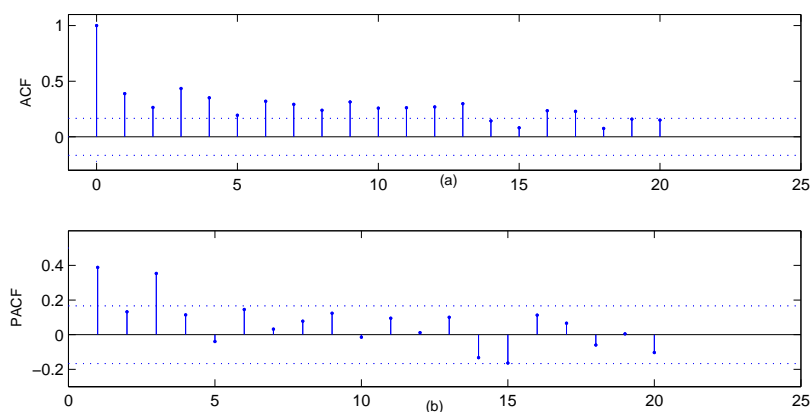

Figure 5. The plots about the zero-inflated drug crime series, (**a**) the ACF plot, (**b**) the PACF plot.

Except for the ZIGINAR$_{\text{RC}}$(1) model, some competitive models are also applied to the time series, such as Poisson INAR(1) (Al-Osh and Alzaid [1]), GINAR(1) (Alzaid and Al-Osh [7]), ZINAR(1) (Jazi, Jones and Lai [9]), ZMGINAR(1) (Barreto-Souza [11]), NGINAR(1) (Ristić, Bakouch and Nastić [12]), ZIMINAR(1) (Li, Wang and Zhang [14]). The Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are suggested to evaluate these models. Numerical results in Table 7 show that the ZIGINAR$_{\text{RC}}$(1) process has the best overall performance, compared with its competitors. Therefore, we assume that the drug crime data is from the ZIGINAR$_{\text{RC}}$(1) model and the estimated parameters in Table 7 are used in the process control procedures. Based on the computation results in Section 4, the CUSUM chart with designs $h = 34$, $k = 2$ (corresponding ARL$_0$ = 364.44) is the best choice, which is shown in Figure 6a. For comparison, we also present CUSUM control charts with designs h=15, $k = 4$ (ARL$_0$ = 358.40) in Figure 6b, designs $h = 12$, $k = 5$ (ARL$_0$ = 372.28) in Figure 6c, and the Shewhart chart with control limit $UCL = 13$ (ARL$_0$ = 340.25) in Figure 6d. We observe that all the CUSUM charts give out-of-control signals, while there is no outliers in the Shewhart chart. Because the Shewhart chart has been proved to be less effective than the CUSUM chart, the drug crime data set seems to be out-of-control with increasing mean shifts or increasing correlation shifts, and some investigation should be done for further explanation. The CUSUM control charts with three designs also display different detection efficiencies. The CUSUM chart with $k = 2$ first signals at $t = 131$ following with continuous alarms as t increases. The signals of the CUSUM chart with $k = 4$ are first given at $t = 133$, then go back below the control limit over a period of time, and come again at $t = 141$. While the outlier of the CUSUM

chart with $k = 5$ occurs at $t = 144$. The analysis above proves again that the CUSUM chart design $k = \lceil \mu_0 \rceil$ is the most effective in practice.

Table 7. The estimated parameters, AIC and BIC of candidate models.

Model	Estimated Parameters		AIC	BIC
Poisson INAR(1)	$\widehat{\lambda} = 1.2806$	$\widehat{\alpha} = 0.2614$	625.787	631.7266
GINAR(1)	$\widehat{p} = 0.3581$	$\widehat{\alpha} = 0.2005$	507.0551	512.9947
NGINAR(1)	$\widehat{\mu} = 1.7925$	$\widehat{\alpha} = 0.3028$	502.6322	508.5718
ZINAR(1)	$\widehat{\alpha} = 0.2157$ $\widehat{p} = 0.5351$	$\widehat{\lambda} = 2.9131$	527.1981	536.1075
ZMGINAR(1)	$\widehat{\mu} = 1.9712$ $\widehat{\pi} = 0.1197$	$\widehat{\alpha} = 0.29$	501.7477	510.6571
ZIMINAR(1)	$\widehat{\alpha} = 0.001$ $\widehat{p} = 0.4101$ $\widehat{\rho} = 0.548$	$\widehat{\beta} = 0.6731$ $\widehat{\lambda} = 2.2564$	503.457	518.3061
ZIGINAR$_{RC}$(1)	$\widehat{\alpha} = 0.547$ $\widehat{p} = 0.185$	$\widehat{\theta} = 2.0495$ $\widehat{\beta} = 0.5188$	494.906	506.7852
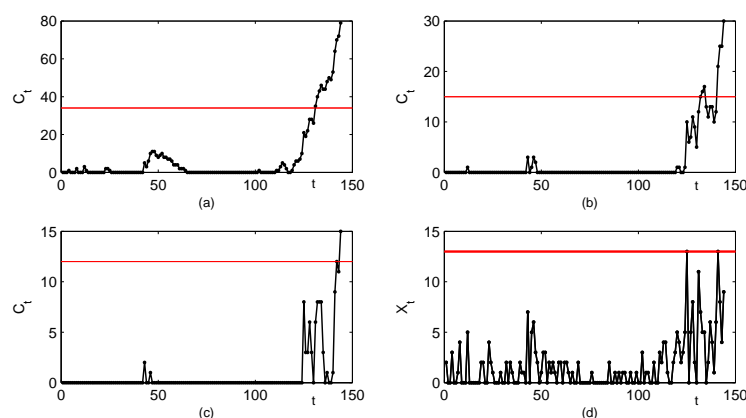
Figure 6. The control charts for the zero-inflated drug crime series, the CUSUM charts with designs (**a**) $h = 34$, $k = 2$, (**b**) $h = 15$, $k = 4$, (**c**) $h = 12$, $k = 5$, and (**d**) the Shewhart chart with $UCL = 13$.

6. Conclusions

In this paper, we have made contributions on monitoring the zero-inflated autocorrelated count data which can be described by an INAR(1) process with random coefficient. ARL and SDRL are adopted to be the measures and calculated by the Markov chain approach. The design parameter k in the CUSUM chart is proved to have great influence on monitoring efficiency, and the smallest integer no less than μ_0 is the recommended value of k. The proposed ZIGINAR$_{RC}$(1)CUSUM chart is proved to have superiorities in various coefficient shifts compared with the conventional Shewhart chart. Computation results also show that the CUSUM chart performs quite well in detecting upward mean shifts, and shows fluctuation in detecting upward correlation shifts. Based on the design of experiment, we also find that a larger value of initial correlation ρ_0 is helpful to detect the correlation shifts. An immediate signal occurs after large upward mean shifts, while does not occur when only correlation shifts exist.

There are some possible topics for our future research. First, we can consider a different monitoring scheme for the ZIGINAR$_{RC}$(1) process and conduct a comparison study. Second, we can explore the monitoring of p-th random coefficient INAR model, which is suitable for the count data with high order dependence. Third, we can study a

multivariate INAR(1) process with random coefficient to continuously monitor the serial correlated counts that we are interested in.

References

1. Al-Osh, M.A.; Alzaid, A.A. First-order integer-valued autoregressive (INAR(1)) process. *J. Time. Ser. Anal.* **1987**, *8*, 261–275. [CrossRef]
2. Steutel, F.W.; Van Harn, K. Discrete analogues of self-decomposability and stability. *Ann. Probab.* **1979**, *7*, 893–899. [CrossRef]
3. Ristić, M.M.; Nastić, A.S. A mixed INAR(p) model. *J. Time. Ser. Anal.* **2012**, *33*, 903–915. [CrossRef]
4. Nastić, A.S.; Laketa, P.N.; Ristić, M.M. Random environment INAR models of higher order. *Revstat-Stat. J.* **2019**, *17*, 35–65.
5. Pedeli, X.; Karlis, D. A bivariate INAR(1) process with application. *Stat. Model.* **2011**, *11*, 325–349. [CrossRef]
6. Khan, N.M.; Cekim, H.O.; Ozel, G. The family of the bivariate integer-valued autoregressive process (BINAR(1)) with Poisson-Lindley (PL) innovations. *J. Stat. Comput. Sim.* **2020**, *90*, 624–637. [CrossRef]
7. Alzaid, A.A.; Al-Osh, M.A. First-order integer-valued autoregressive (INAR(1)) process: Distributional and regression properties. *Stat. Neerl.* **1988**, *42*, 53–61. [CrossRef]
8. Alzaid, A.A.; Al-Osh, M.A. Some autoregressive moving average processes with generalized Poisson marginal distributions. *Ann. I. Stat. Math.* **1993**, *45*, 223–232. [CrossRef]
9. Jazi, M.A.; Jones, G.; Lai, C.D. First-order integer valued AR processes with zero inflated Poisson innovations. *J. Time. Ser. Anal.* **2012**, *33*, 954–963. [CrossRef]
10. Ristić, M.M.; Nastić, A.S.; Bakouch, H.S. Estimation in an integer-valued autoregressive process with negative binomial marginals (NBINAR(1)). *Commun. Stat-Theor. M.* **2012**, *41*, 606–618. [CrossRef]
11. Barreto-Souza, W. Zero-modified Geometric INAR(1) process for modelling count time series with deflation or inflation of zeros. *J. Time. Ser. Anal.* **2015**, *36*, 839–852. [CrossRef]
12. Ristić, M.M.; Bakouch, H.S.; Nastić, A.S. A new geometric first-order integer-valued autoregressive (NGINAR(1)) process. *J. Stat. Plan. Infer.* **2009**, *139*, 2218–2226. [CrossRef]
13. Liu, Z.; Zhu, F. A new extension of thinning-based integer-valued autoregressive models for count data. *Entropy* **2021**, *23*, 62. [CrossRef] [PubMed]
14. Li, C.; Wang, D.; Zhang, H. First-order mixed integer-valued autoregressive processes with zero-inflated generalized power series innovations. *J. Korean Stat. Soc.* **2015**, *44*, 232–246. [CrossRef]
15. Orozco, D.L.; Sales, L.O.; Fernández, L.M.; Pinho, A.L. A new mixed first-order integer-valued autoregressive process with Poisson innovations. *ASTA-Adv. Stat. Anal.* **2020**, 1–22. [CrossRef]
16. Weiß, C.H. Thinning operations for modeling time series of counts-a survey. *Asta-Adv. Stat. Anal.* **2008**, *92*, 319–341. [CrossRef]
17. Scotto, M.; Weiß, C.H.; Gouveia, S. Thinning-based models in the analysis of integer-valued time series: A review. *Stat. Model.* **2015**, *15*, 590–618. [CrossRef]
18. Zheng, H.; Basawa, I.V.; Datta, S. Inference for pth-order random coefficient integer-valued autoregressive processes. *J. Time. Ser. Anal.* **2006**, *27*, 411–440. [CrossRef]
19. Kang, J.; Lee, S. Parameter change test for random coefficient integer-valued autoregressive processes with application to polio data analysis. *J. Time. Ser. Anal.* **2009**, *30*, 239–258. [CrossRef]
20. Zhang, H.; Wang, D.; Zhu, F. Empirical likelihood inference for random coefficient INAR(p) process. *J. Time. Ser. Anal.* **2011**, *32*, 195–203. [CrossRef]
21. Bakouch, H.S.; Mohammadpour, M.; Shirozhan, M. A zero-inflated geometric INAR(1) process with random coefficient. *Appl. Math.* **2018**, *63*, 79–105. [CrossRef]
22. Weiß, C.H. EWMA monitoring of correlated processes of Poisson counts. *Qual. Technol. Quant. M.* **2009**, *6*, 137–153. [CrossRef]
23. Weiß, C.H.; Testik, M.C. CUSUM monitoring of first-order integer-valued autoregressive processes of Poisson counts. *J. Qual. Technol.* **2009**, *41*, 389–400. [CrossRef]
24. Weiß, C.H.; Testik, M.C. The Poisson INAR(1) CUSUM chart under overdispersion and estimation error. *IIE Trans.* **2011**, *43*, 805–818. [CrossRef]
25. Yontay, P.; Weiß, C.H.; Testik, M.C.; Bayindir, Z.P. A two-sided cumulative sum chart for first-order integer-valued autoregressive processes of poisson counts. *Qual. Reliab. Eng. Int.* **2013**, *29*, 33–42. [CrossRef]

26. Rakitzis, A.C.; Weiß, C.H.; Castagliola, P. Control charts for monitoring correlated Poisson counts with an excessive number of zeros. *Qual. Reliab. Eng. Int.* **2017**, *33*, 413–430. [CrossRef]

27. Li, C.; Wang, D.; Sun, J. Control charts based on dependent count data with deflation or inflation of zeros. *J. Stat. Comput. Sim.* **2019**, *89*, 3273–3289. [CrossRef]

28. Fernandes, F.H.; Bourguignon, M.; Ho, L.L. Control charts to monitor integer valued autoregressive process with inflation or deflation of zeros. *J. Stat. Manag. Syst.* **2020**, *23*, 1463–1484. [CrossRef]

29. Sales, L.O.; Pinho, A.L.; Vivacqua, C.A.; Ho, L.L. Shewhart control chart for monitoring the mean of Poisson mixed integer autoregressive processes via Monte Carlo simulation. *Comput. Ind. Eng.* **2020**, *140*, 106245. [CrossRef]

30. Page, E.S. Cumulative sum charts. *Technometrics* **1961**, *3*, 1–9. [CrossRef]

31. Montgomery, D.C. *Introduction to Statistical Quality Control*, 6th ed.; Wiley: New York, NY, USA, 2009.

32. Alencar, A.P.; Ho, L.L.; Albarracin, O.Y.E. CUSUM control charts to monitor series of negative binomial count data. *Stat. Methods. Med. Res.* **2017**, *26*, 1925–1935. [CrossRef]

33. Bourguignon, M.; Medeiros, R.M.; Fernandes, F.H.; Ho, L.L. Simple and useful statistical control charts for monitoring count data. *Qual. Reliab. Eng. Int.* **2021**, *37*, 541–566. [CrossRef]

34. Harris, T.J.; Ross, W.H. Statistical process control procedures for correlated observations. *Can. J. Chem. Eng.* **1991**, *69*, 48–57. [CrossRef]

35. Triantafyllopoulos, K.; Bersimis, S. Phase II control charts for autocorrelated processes. *Qual. Technol. Quant. M.* **2016**, *13*, 88–108. [CrossRef]

36. Albarracin, O.Y.E.; Alencar, A.P.; Ho, L.L. Effect of neglecting autocorrelation in regression EWMA charts for monitoring count time series. *Qual. Reliab. Eng. Int.* **2018**, *34*, 1752–1762. [CrossRef]

37. Lee, S.; Kim, D. Monitoring parameter change for time series models of counts based on minimum density power divergence estimator. *Entropy* **2020**, *22*, 1304. [CrossRef] [PubMed]

38. Lee, S.; Kim, C.K.; Kim, D. Monitoring volatility change for time series based on support vector regression. *Entropy* **2020**, *22*, 1312. [CrossRef]

39. Brook, D.A.; Evans, D. An approach to the probability distribution of CUSUM run length. *Biometrika* **1972**, *59*, 539–549. [CrossRef]