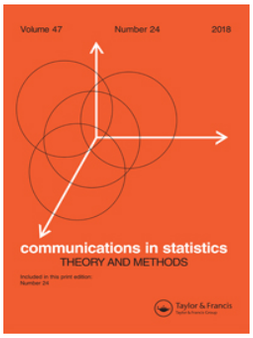

Communications in Statistics - Theory and Methods

ISSN: 0361-0926 (Print) 1532-415X (Online) Journal homepage: http://www.tandfonline.com/loi/lsta20

Phase I monitoring and change point estimation of autocorrelated poisson regression profiles

M. R. Maleki, A. Amiri, A. R. Taheriyoun & P. Castagliola

To cite this article: M. R. Maleki, A. Amiri, A. R. Taheriyoun & P. Castagliola (2018) Phase I monitoring and change point estimation of autocorrelated poisson regression profiles, Communications in Statistics - Theory and Methods, 47:24, 5885-5903, DOI: 10.1080/03610926.2017.1402052

To link to this article: https://doi.org/10.1080/03610926.2017.1402052

Published online: 05 Dec 2017.

COMMUNICATIONS IN STATISTICS—THEORY AND METHODS
2018, VOL. 47, NO. 24, 5885–5903
https://doi.org/10.1080/03610926.2017.1402052

Phase I monitoring and change point estimation of autocorrelated poisson regression profiles

M. R. Maleki[a], A. Amiri [a], A. R. Taheriyoun [b], and P. Castagliola [c]

[a]Department of Industrial Engineering, Faculty of Engineering, Shahed University, Tehran, Iran; [b]Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, G.C. Tehran, Iran; [c]Universit de Nantes & LS2N UMR CNRS, Nantes, France

ABSTRACT
The objective of this paper is to study the Phase I monitoring and change point estimation of autocorrelated Poisson profiles where the response values within each profile are autocorrelated. Two charts, the SLRT and the Hotelling's T^2, are proposed along with an algorithm for parameter estimation. The detecting power of the proposed charts is compared using simulations in terms of the signal probability criterion. The performance of the SLRT method in estimating the change point in the regression parameters is also evaluated. Moreover, a real data example is presented to illustrate the application of the methods.

ARTICLE HISTORY
Received 13 August 2017
Accepted 3 November 2017

KEYWORDS
Change point; likelihood ratio test; Phase I; signal probability; within-profile autocorrelation.

MATHEMATICS SUBJECT CLASSIFICATION

1. Introduction

In some practical systems, the quality of a process is well characterized by a functional relationship between a response variable and one or several explanatory variables. Analyzing the stability of this relationship over time through statistical techniques is referred to as "profile monitoring". For detailed information about profile monitoring approaches in Phases I and II refer to review papers by Woodall et al. (2004) and Woodall (2007) and the book by Noorossana, Saghaei, and Amiri (2011). In most researches in the literature of profile monitoring, the response values within each profile are assumed to be independent from each other. However, in many practical situations, the independency assumption of response values within each profile is violated. In recent years, taking into account the autocorrelation structure in profile monitoring approaches is discussed by Noorossana, Amiri, and Soleimani (2008), Jensen, Birch, and Woodall (2008), Soleimani, Noorossana, and Amiri (2009), Amiri, Jensen, and Kazemzadeh (2010), Noorossana, Saghaei, and Dorri (2010), Abdel-Salam, Birch, and Jensen (2013), Keramatpour et al. (2013), Narvand, Soleimani, and Raissi (2013), Soleimani, Noorossana, and Niaki (2013b), Soleimani, Narvand, and Raissi (2013a), Keramatpour, Niaki, and Amiri (2014), Soleimani and Hadizadeh (2014), Soleimani and Noorossana (2014), Zhang et al. (2014), and Khedmati and Niaki (2016).

In all of the above-mentioned papers, the response variable is assumed to follow a normal distribution. However, in practice, it is likely to face with conditions where the response values belong to a discrete distribution such as binary, binomial, Poisson, etc. To the best

CONTACT A. Amiri amiri@shahed.ac.ir Department of Industrial Engineering, Faculty of Engineering, Shahed University, Tehran, P.O. Box 18151-159, Iran.

of our knowledge, only few methods have been proposed to investigate autocorrelated profiles with discrete response variables. Koosha and Amiri (2013) explored the effect of within-profile autocorrelation on Phase I monitoring of logistic regression profiles. They used two approaches, namely the modification of the upper control limit and the generalized linear mixed model (GLMM), to account for the within-profile autocorrelation. Maleki, Amiri, and Taheriyoun (2017) proposed a state space model for Phase II monitoring of autocorrelated binary profiles when the successive response values within each profile are autocorrelated. The method they suggested is not only able to detect changes in the autocorrelated logistic regression parameters but, also in the autocorrelation coefficient. Then, Maleki, Amiri, and Taheriyoun (2017b) proposed two maximum likelihood estimators to identify the change point under step changes and linear trends in Phase II monitoring of binary regression profiles in the presence of within-profile autocorrelation. They evaluated the performance of the proposed estimators in terms of the accuracy and the precision criteria through simulation studies. They also presented a numerical example to illustrate the application of their proposed estimators under both step change and linear trend.

Although analyzing autocorrelated profiles under binary response values has been receiving little attention, monitoring autocorrelated profiles with other kind of discrete distributions has clearly been neglected in the literature. In many practical profile monitoring applications, as a response variable, it is common to deal with count data in the case of within-profile autocorrelation. As one of the most important count data-type distribution in profile monitoring applications, the quality practitioners would face with Poisson response variables where the observations within each profile are autocorrelated. To illustrate the motivation of our work, as an environmental application, consider the number of agglomerates which are ejected from a volcano in successive days. In this example, the count of agglomerates (the response variable) is a function of the agglomerates diameters (the explanatory variable). Clearly, the count of agglomerates in a given day will affect the future ones. Hence, due to the autocorrelation structure between the response values, an autocorrelated Poisson regression model should be used to express the relationship between the response and the explanatory variable. As another example, consider the count of stolen goods in successive months of a given city which is considered as the response variable. Here, the explanatory variable would be the count of theft in the adjacent city. In this example, the count of theft for each month not only is affected by the value of explanatory variable, but also by the count of theft during the preceding months.

As far as we know, analyzing such profiles with Poisson response variables are restricted to Amiri, Koosha, and Azhdari (2011) and Amiri, Koosha, Azhdari, and Wang (2015) where the response values are assumed to be independent. Hence, due to the potential application of profile monitoring with autocorrelated count data in real manufacturing and non-manufacturing situations as well as to fill the mentioned research gap, the Phase I monitoring as well as the change point estimation of autocorrelated Poisson regression profiles are both investigated in this paper. In the proposed model, the process parameters are linked linearly to their past values as well as to the observed values of the response variable via an INGARCH(1,1) model. This paper is organized as follows. The proposed regression model to account for the within-profile autocorrelation along with the parameter estimation method are discussed in Section 2. The proposed control charts to monitor the autocorrelated Poisson profile as well as to identify the time of the step change in the vector of model parameters are introduced in Section 3. Then, in Section 4, the performance of the proposed control charts, namely, the SLRT and the Hotelling's T^2 charts, to detect different out-of-control scenarios, is evaluated and compared in terms of the out-of-control signal probability criterion. In Section 5, the performance of the SLRT chart, a change point-based method, to estimate the time of the step change in the vector of regression model coefficients is assessed in terms of accuracy

and precision as well as the empirical distribution of the estimated change point parameter. In Section 6, a real data example is presented to illustrate the implementation of the proposed framework. Finally, concluding remarks along with some recommendations for future researches are given in Section 7.

2. The autocorrelated Poisson regression profile and the parameter estimation approach

First, in this section, the Poisson regression model considering the within-profile autocorrelation is discussed. Then, the particle swarm optimization (PSO) algorithm used in this paper to estimate the regression parameters will be described.

2.1. The proposed autocorrelated Poisson model

Recently, modeling time series of count data has received a growing attention by researchers and it has been applied in several fields. The wide application of integer-valued time series are highlighted by the quality practitioners. The integer-valued generalized autoregressive conditional heteroscedastic (INGARCH (p, q)) model is a popular tool to model autocorrelation structure of count data (Zhu 2012). This model, proposed by Ferland, Latour, and Oraichi (2006), is an integer-valued analogue of a GARCH(p, q) process. The INGARCH model can be used to model processes where the value of the response variable in a given treatment not only depends on the preceding response values but also on the value of the explanatory variable in the current treatment. For the explicit expression of the model, in this section, the autocorrelated Poisson regression model using the INGARCH(1,1) structure is introduced. Let $y_{i,j}$ be the response value at ith; $i = 1, \ldots, n$ experimental setting of profile $j; j = 1, \ldots, m$. We assume that the observations $\{y_{1,j}, \ldots, y_{n,j}\}$ within each profile j are autocorrelated and follow an INGARCH(1,1) model where the intensity parameter for observation i and profile j is equal to $\lambda_{i,j}(\boldsymbol{\theta})$. In this case, $\lambda_{i,j}(\boldsymbol{\theta})$ is expressed in terms of the vector $\boldsymbol{\theta} = (d, a, b, \boldsymbol{\beta})^{\intercal}$, $\lambda_{i-1,j}$ (the Poisson parameter corresponding to the previous observation), $y_{i-1,j}$ (the response value corresponding to the previous observation) as well as the vector of explanatory variables $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,p})^{\intercal}$ for observation i. Then, the Poisson parameter for observation i and profile j is given as

$$\lambda_{i,j}(\boldsymbol{\theta}) = d + a\lambda_{i-1,j}(\boldsymbol{\theta}) + by_{i-1,j} + \mathbf{x}_i^{\intercal}\boldsymbol{\beta} \tag{1}$$

where $d, a, b > 0$ are three parameters to be fixed, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^{\intercal}$ is the vector of coefficients associated with the explanatory variables and $\lambda_{0,j}$ and $y_{0,j}$ are two initial values which have to be fixed. We can rewrite (1) as

$$\begin{aligned}\lambda_{i,j} = {} & d(1 + a + \cdots + a^{i-1}) + a^i\lambda_{0,j} \\ & + b(a^{i-1}y_{0,j} + a^{i-2}y_{1,j} + \cdots + ay_{i-2,j} + y_{i-1,j}) \\ & + (a^{i-1}\mathbf{x}_1 + \cdots + a\mathbf{x}_{i-1} + \mathbf{x}_i)^{\intercal}\boldsymbol{\beta}\end{aligned} \tag{2}$$

For profile $j = 1, \ldots, m$, the joint probability function of the observations $\{y_{0,j}, y_{1,j}, \ldots, y_{n,j}\}$ is equal to

$$\begin{aligned}f(y_{0,j}, y_{1,j}, \ldots, y_{n,j}) = {} & f(y_{n,j}|y_{n-1,j}, \ldots, y_{0,j}) \\ & \times f(y_{n-1,j}|y_{n-2,j}, \ldots, y_{0,j}) \\ & \times \cdots \times f(y_{1,j}|y_{0,j}) \times f(y_{0,j})\end{aligned} \tag{3}$$

Using the Markov property, one can rewrite (3) as follows:

$$f(y_{0,j}, y_{1,j}, \ldots, y_{n,j}) = f(y_{n,j}|y_{n-1,j}) \times f(y_{n-1,j}|y_{n-2,j}) \times \cdots \times f(y_{2,j}|y_{1,j}) \times f(y_{1,j}|y_{0,j}) \times f(y_{0,j}) \tag{4}$$

Then, for profile $j = 1, \ldots, m$, the likelihood function of vector $\boldsymbol{\theta}$ given the initial value of $y_{0,j}$ and based on observations $y_{1,j}, \ldots, y_{n,j}$ is equal to

$$\begin{aligned} L_j(\boldsymbol{\theta}) &= \left(\prod_{i=1}^{n} L_{i,j}(\boldsymbol{\theta})\right) f(y_{0,j}) \\ &= \left(\prod_{i=1}^{n} \frac{e^{-\lambda_{i,j}(\boldsymbol{\theta})}(\lambda_{i,j}(\boldsymbol{\theta}))^{y_{i,j}}}{y_{i,j}!}\right)\left(\frac{e^{-\lambda_{0,j}}\lambda_{0,j}}{y_{0,j}!}\right) \end{aligned} \tag{5}$$

where $L_{i,j}(\boldsymbol{\theta}) = f(y_{i,j}|y_{i-1,j})$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$. Taking the logarithm of $L_j(\boldsymbol{\theta})$, leads to $\ln L_j(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ln L_{i,j}(\boldsymbol{\theta}) + \ln f(y_{0,j})$. Since the last term of $\ln L_j(\boldsymbol{\theta})$ does not depend on parameter $\boldsymbol{\theta}$, we can rewrite it as:

$$\ln L_j(\boldsymbol{\theta}) = \sum_{i=1}^{n} \left(y_{i,j} \ln(\lambda_{i,j}(\boldsymbol{\theta})) - \lambda_{i,j}(\boldsymbol{\theta})\right) + C \tag{6}$$

where C is a constant that does not depend on vector $\boldsymbol{\theta}$. Under some mild restrictions on the parameter space, one may assume that $\ln L_j(\boldsymbol{\theta})$ is a differentiable function, thus $\frac{\partial \ln L_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{0}$. Now assume that there exists a sub-space at the parameter space such that $\ln L_j(\boldsymbol{\theta})$ is twice differntiable. Then, use of the Taylor expansion allows to obtain:

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\partial \ln L_{i,j}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} + (\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^{\mathsf{T}} \left[\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln L_{i,j}(\boldsymbol{\theta})}{(\partial\boldsymbol{\theta})\,(\partial\boldsymbol{\theta})^{\mathsf{T}}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right] = \mathbf{0} \tag{7}$$

where $\boldsymbol{\theta}_0$ is the vector of the model parameters when the process is in-control. Therefore:

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^{\mathsf{T}} = \left[-\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln L_{i,j}(\boldsymbol{\theta})}{(\partial\boldsymbol{\theta})\,(\partial\boldsymbol{\theta})^{\mathsf{T}}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right]^{-1} \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \frac{\partial \ln L_{i,j}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \tag{8}$$

Using the law of large numbers and continuous mapping theorem, we have:

$$\left[\frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 \ln L_{i,j}(\boldsymbol{\theta})}{(\partial\boldsymbol{\theta})\,(\partial\boldsymbol{\theta})^{\mathsf{T}}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right]^{-1} \to E\left[-\frac{\partial^2 \ln L_j(\boldsymbol{\theta})}{(\partial\boldsymbol{\theta})\,(\partial\boldsymbol{\theta})^{\mathsf{T}}}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}\right]^{-1} \tag{9}$$

where the right hand is the information matrix. The normality of MLE estimators is related to the use of a central limit theorem for the convergence at $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} \frac{\partial \ln L_{i,j}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}$, where the rate of convergance is at least $O(\sqrt{n})$. As an example, for a sample size of $n = 100$, the normal approximation of $\widehat{\boldsymbol{\theta}}$, computes $P(\widehat{\boldsymbol{\theta}} \in A)$ for any Borel set A with an error of less than 0.1.

Here we define the ML equation system of $\ln L_j(\boldsymbol{\theta})$ as

$$\mathbf{g}(\boldsymbol{\theta}) = \frac{\partial \ln L_j(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{i=1}^{n} \left(\frac{y_{i,j}}{\lambda_{i,j}(\boldsymbol{\theta})} - 1\right) \frac{\partial \lambda_{i,j}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \tag{10}$$

The maximum point of $\ln L_j(\boldsymbol{\theta})$, called as $\widehat{\boldsymbol{\theta}}$, can be achieved by solving $\mathbf{g}(\boldsymbol{\theta}) = 0$. How to estimate the vector of regression parameters is discussed in Sub section 2.2.

2.2. Parameter estimation by PSO algorithm

As noted in the previous subsection, $\widehat{\boldsymbol{\theta}}$ is obtained by solving $\mathbf{g}(\boldsymbol{\theta}) = 0$. However, there is no way to find a closed-form solution for $\widehat{\boldsymbol{\theta}}$. In addition, the parameter estimation methods proposed in the literature of profile monitoring to estimate the regression parameters are based on the independency assumption of the response values within each profile which is violated in many practical applications. As an example, by assuming the case of independent response values, Amiri, Koosha, Azhdari, and Wang (2015) used an iterative reweighted least-square (IRLS) method to estimate the regression parameters for Poisson regression profiles. On the other hand, the run time by using exact algorithms to maximize the logarithm of the likelihood function significantly depends on the Hessian matrix of $\ln L_j(\boldsymbol{\theta})$. As the determinant value of the Hessian matrix increases, the run time of the algorithm will increase considerably. Hence, it is important here to use an algorithm to obtain $\widehat{\boldsymbol{\theta}}$ for which the run time is as short as possible. This issue is more crucial in processes where the rate of production per time unit is high. To deal with the mentioned issue and to take into account the within-profile autocorrelation structure, in this paper we develop a PSO type algorithm to obtain $\widehat{\boldsymbol{\theta}}$ that maximizes the logarithm of the likelihood function.

The PSO metaheuristic algorithm which has been extended by Eberhart and Kennedy (1995) is a global search algorithm based on the swarm behavior of particles. A point in the problem space is referred as a particle which is considered as a candidate solution in the optimization problem. In an iterative procedure, this computational algorithm moves toward an optimum solution by improving a candidate solution with regard to a given measure of quality. In PSO algorithm, the collective behavior of a population not only depends on the individual behavior of each particle, but it also depends on the interaction between them. As a consequence, there is a complex nonlinear relationship between the individual and the collective behaviors in this algorithm. Each particle is initialized with a random position (solution) and it seeks the optimal solution such that it can keep track of its position, velocity (change pattern of solution) and fitness. The position and the velocity of each particle is adjusted by considering its own experience and social cooperation by its fitness to the environment. As a result, three factors including (1) the current position of the particle, (2) the best position of the particle among its previous positions (personal best), and (3) the best position among all particles in the population (global best) can affect the behavior of a given particle. In other words, each particle improves its position based on the current velocity as well as the personal best and the global best. Therefore, the velocity and position of particle $k = 1, \ldots, K$, during iteration $(u+1)$ can be expressed by (11) and (12), defined as:

$$\mathbf{v}_k^{(u+1)} = w^{(u+1)}\mathbf{v}_k^{(u)} + c_1\mathbf{r}_1\left(\mathbf{p}_k^{(u)} - \mathbf{x}_k^{(u)}\right) + c_2\mathbf{r}_2\left(\mathbf{p}_g^{(u)} - \mathbf{x}_k^{(u)}\right) \tag{11}$$

$$\mathbf{x}_k^{(u+1)} = \mathbf{x}_k^{(u)} + \mathbf{v}_k^{(u+1)} \tag{12}$$

where c_1 and c_2 are the so-called *cognitive learning rate* and *social learning rate*, respectively, while $\mathbf{r}_1$ and $\mathbf{r}_2$ are random vectors with a uniform distribution within the range [0, 1]. In (11), $\mathbf{p}_k^{(u)}$ and $\mathbf{p}_g^{(u)}$ represent the personal best of particle k and the global best at iteration u, respectively. Note that, w at iteration $(u+1)$ is defined as the inertia weight which can be obtained as

$$w^{(u+1)} = w_{damp} \times w^{(u)} \tag{13}$$

where parameter w_{damp} is selected within the range [0, 1] which implies that the effect of inertia in consecutive iterations decreases in comparison with $\mathbf{p}_k^{(u)}$ and $\mathbf{p}_g^{(u)}$.

3. Proposed control charts

In this section, the proposed control charts for Phase I monitoring of Poisson regression profiles under within-profile autocorrelation are expressed. Note that the proposed SLRT chart not only detects different changes, but it also has capability to identify the change point in the vector of model parameters.

3.1. SLRT method

Suppose that a step change in vector $\boldsymbol{\theta}$ occurs for profile $\tau \in \{1, 2, \ldots, m-1\}$. Then, $L_{\text{all}}(\boldsymbol{\theta})$ (the likelihood function concerning the complete data), $L_{\text{bef}}(\boldsymbol{\theta})$ (the likelihood function concerning the first τ profiles), and $L_{\text{aft}}(\boldsymbol{\theta})$ (the likelihood function concerning the remaining $m-\tau$ profiles) are calculated according to (14), (15) and (16), respectively,

$$L_{\text{all}}(\boldsymbol{\theta}) = \prod_{j=1}^{m} f(y_{0,j}) \prod_{i=1}^{n} \frac{e^{-\lambda_{i,j}(\boldsymbol{\theta})}(\lambda_{i,j}(\boldsymbol{\theta}))^{y_{i,j}}}{y_{i,j}!} \tag{14}$$

$$L_{\text{bef}}(\boldsymbol{\theta}) = \prod_{j=1}^{\tau} f(y_{0,j}) \prod_{i=1}^{n} \frac{e^{-\lambda_{i,j}(\boldsymbol{\theta})}(\lambda_{i,j}(\boldsymbol{\theta}))^{y_{i,j}}}{y_{i,j}!} \tag{15}$$

$$L_{\text{aft}}(\boldsymbol{\theta}) = \prod_{j=\tau+1}^{m} f(y_{0,j}) \prod_{i=1}^{n} \frac{e^{-\lambda_{i,j}(\boldsymbol{\theta})}(\lambda_{i,j}(\boldsymbol{\theta}))^{y_{i,j}}}{y_{i,j}!} \tag{16}$$

Obviously the values of the log-likelihood $\ln(L_{\text{all}}(\boldsymbol{\theta}))$, $\ln(L_{\text{bef}}(\boldsymbol{\theta}))$, and $\ln(L_{\text{aft}}(\boldsymbol{\theta}))$ are calculated as follows:

$$\begin{aligned}\ln\left(L_{\text{all}}(\boldsymbol{\theta})\right) = &\sum_{j=1}^{m}\sum_{i=1}^{n} y_{i,j}\ln(\lambda_{i,j}(\boldsymbol{\theta})) - \lambda_{i,j}(\boldsymbol{\theta}) - \ln(y_{i,j}!) \\ &+ \sum_{j=1}^{m} y_{0,j}\ln(\lambda_{0,j}) - \lambda_{0,j} - \ln(y_{i,j}!)\end{aligned} \tag{17}$$

$$\begin{aligned}\ln\left(L_{\text{bef}}(\boldsymbol{\theta})\right) = &\sum_{j=1}^{\tau}\sum_{i=1}^{n} y_{i,j}\ln(\lambda_{i,j}(\boldsymbol{\theta})) - \lambda_{i,j}(\boldsymbol{\theta}) - \ln(y_{i,j}!) \\ &+ \sum_{j=1}^{\tau} y_{0,j}\ln(\lambda_{0,j}) - \lambda_{0,j} - \ln(y_{i,j}!)\end{aligned} \tag{18}$$

$$\begin{aligned}\ln\left(L_{\text{aft}}(\boldsymbol{\theta})\right) = &\sum_{j=\tau+1}^{m}\sum_{i=1}^{n} y_{i,j}\ln(\lambda_{i,j}(\boldsymbol{\theta})) - \lambda_{i,j}(\boldsymbol{\theta}) - \ln(y_{i,j}!) \\ &+ \sum_{j=\tau+1}^{m} y_{0,j}\ln(\lambda_{0,j}) - \lambda_{0,j} - \ln(y_{i,j}!)\end{aligned} \tag{19}$$

Here, through the extended PSO algorithm, the estimated parameter vectors $\widehat{\boldsymbol{\theta}}_{\text{all}}$, $\widehat{\boldsymbol{\theta}}_{\text{bef}}$, and $\widehat{\boldsymbol{\theta}}_{\text{aft}}$ are obtained by maximizing (17), (18), and (19), respectively. Let ℓ_{all}, ℓ_{bef}, and ℓ_{aft} be the values of $L_{\text{all}}(\boldsymbol{\theta})$, $L_{\text{bef}}(\boldsymbol{\theta})$, and $L_{\text{aft}}(\boldsymbol{\theta})$ obtained by replacing vector $\boldsymbol{\theta}$ by $\widehat{\boldsymbol{\theta}}_{\text{all}}$, $\widehat{\boldsymbol{\theta}}_{\text{bef}}$, and $\widehat{\boldsymbol{\theta}}_{\text{aft}}$ in (14), (15) and (16), respectively. The LRT statistic for Phase I monitoring of autocorrelated Poisson regression profiles considering the within-profile structure expressed in (1) is

$$LRT_{\tau} = -2\left(\ln \ell_{\text{all}} - (\ln \ell_{\text{bef}} + \ln \ell_{\text{aft}})\right) \tag{20}$$

Finally, the likelihood ratio statistic is standardized as follows:

$$SLRT_\tau = \frac{LRT_\tau - E(LRT_\tau)}{\sigma(LRT_\tau)} \tag{21}$$

where $E(LRT_\tau)$ and $\sigma(LRT_\tau)$ are the expected value and standard deviation of LRT_τ, respectively, which are computed by Monte Carlo simulation in practice. The control chart triggers an out-of-control signal when $SLRT_\tau > UCL$ where the upper control limit UCL is obtained by simulation in order to satisfy a fixed probability of Type-I error α. The guideline for designing and implementing the proposed SLRT chart is depicted in Figure 1.

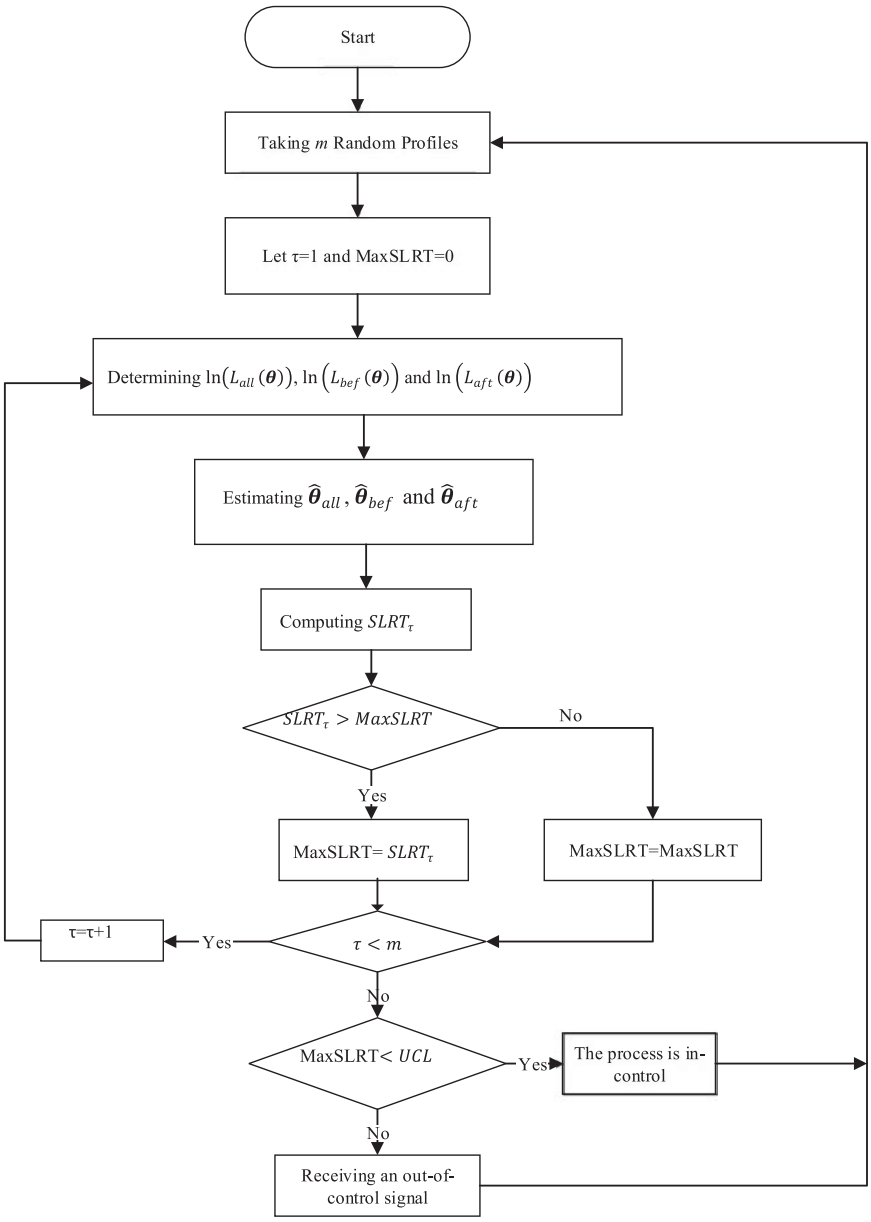

Figure 1. The flowchart of the proposed SLRT chart.

3.2. Hotelling's T^2 control chart

As the first study in the profile monitoring area concerning discrete distributions for the response variable, Yeh, Huwang, and LI (2009) proposed five Hotelling's T^2 based charts for Phase I monitoring of binary regression profiles under assumption of independence of the observations. This chart has been extended by Maleki, Amiri, and Taheriyoun (2017) to monitor binary regression profiles under within-profile autocorrelation structure in Phase II. Here, we develop this approach to a Phase I monitoring of Poisson regression profiles where the response values within each profile are autocorrelated based on the INGARCH(1,1) model. The chart statistic for profile $j \in \{1, \ldots, m\}$ is expressed as

$$T_j^2 = (\widehat{\boldsymbol{\theta}}_j - \bar{\widehat{\boldsymbol{\theta}}})^\intercal \mathbf{S}_{\widehat{\boldsymbol{\theta}}}^{-1} (\widehat{\boldsymbol{\theta}}_j - \bar{\widehat{\boldsymbol{\theta}}}) \tag{22}$$

where $\widehat{\boldsymbol{\theta}}_j$ is the estimated vector of model parameters which is obtained by maximizing $L_j(\boldsymbol{\theta})$ using the PSO algorithm, $\bar{\widehat{\boldsymbol{\theta}}} = \frac{1}{m}\sum_{j=1}^{m} \widehat{\boldsymbol{\theta}}_j$ and $\mathbf{S}_{\widehat{\boldsymbol{\theta}}}$ is the sample variance-covariance matrix of the estimated model parameters corresponding to profiles $j = 1, \ldots, m$ which is defined as

$$\mathbf{S}_{\widehat{\boldsymbol{\theta}}} = \frac{1}{2\,(m-1)} \sum_{j=1}^{m-1} (\widehat{\boldsymbol{\theta}}_{j+1} - \widehat{\boldsymbol{\theta}}_j)(\widehat{\boldsymbol{\theta}}_{j+1} - \widehat{\boldsymbol{\theta}}_j)^\intercal \tag{23}$$

The chart triggers an out-of-control signal when the extended Hotelling's T^2 statistic exceeds an upper control limit UCL obtained by simulation in order to satisfy a fixed probability of Type-I error α.

4. Performance evaluation of the proposed method to detect shifts in model parameters

In this section, the performance of the proposed monitoring schemes to detect different shifts in the vector of model parameters is evaluated in terms of the signal probability criterion. Without loss of generality, we assume that $m = 30$, $n = 20$, $p = 1$, and $\lambda_0 = 1$. The vector of explanatory variable, which is fixed from profile to profile is considered equal to $\mathbf{x} = (0, 01, 0, 02, \ldots, 0.20)^\intercal$ and when the process is in-control, the vector $\boldsymbol{\theta}$ is equal to $(d, a, b, \beta)^\intercal = (1, 0.2, 0.2, 0.25)^\intercal$. The UCL values for the proposed SLRT and Hotelling's T^2 charts are set equal to 6.7238 and 22.3083, respectively, in order to have $\alpha = 0.05$.

The out-of-control signal probabilities under different step shifts in the model parameters for different values of parameter τ, namely, $\tau \in \{5, 10, 15\}$ are given in Tables 1–4. The step changes in the model parameters are denoted by $q + \delta_q\sigma_q$, with $q \in \{d, a, b, \beta\}$, where δ_q and σ_q are the magnitude of shift for parameter q and the standard deviation of this parameter, respectively. Recall that τ is the time where the process changes to an out-of-control state. It is

Table 1. Signal probability values under shift from d to $d + \delta_d\sigma_d$.

τ	δ_d	0.5	0.75	1	1.5	2	2.5
5	T^2	0.061	0.068	0.077	0.085	0.087	0.097
	SLRT	0.066	0.074	0.098	0.133	0.202	0.417
10	T^2	0.052	0.066	0.067	0.074	0.083	0.094
	SLRT	0.082	0.091	0.101	0.128	0.155	0.301
15	T^2	0.053	0.061	0.066	0.084	0.087	0.099
	SLRT	0.083	0.085	0.086	0.117	0.138	0.171

Table 2. Signal probability values under shift from a to $a + \delta_a\sigma_a$.

τ	δ_a	0.5	0.75	1	1.5	2	2.5
5	T^2	0.065	0.069	0.071	0.088	0.094	0.109
	SLRT	0.095	0.108	0.111	0.726	1	1
10	T^2	0.067	0.068	0.079	0.103	0.108	0.118
	SLRT	0.075	0.083	0.129	0.517	1	1
15	T^2	0.062	0.074	0.075	0.086	0.090	0.104
	SLRT	0.072	0.084	0.112	0.234	0.998	1

also worth to mention that the signal probability criterion is defined as the number of out-of-control signals divided by the total number of simulation replicates. The performance of the proposed control charts to detect different shifts in parameter d in units of σ_d is summarized in Table 1. As we can see, for both methods, as δ_d increases, the out-of-control signal probability increases. Table 1 shows that for all values of parameter τ, the proposed SLRT chart outperforms the Hotelling's T^2 chart under all shifts induced in parameter d. The results of Table 1 also reveal that for a small shift $\delta_d = 0.5$, the proposed SLRT method has its best performance when τ is equal to 15. However, as the magnitude of the shift increases to 0.75 and 1, the best performance of this method is obtained when $\tau = 10$. Finally, for larger shifts of $\delta_d \in \{1.5, 2, 2.5\}$, the proposed SLRT control chart for $\tau = 5$ outperforms the other two ones. It can be generally concluded that, decreasing the value of parameter τ in the SLRT method leads to improving the capability of this method to detect large shifts in parameter d.

Table 2 shows the obtained out-of-control signal probabilities for different values of τ under shifts in parameter a in units of σ_a. As expected, it is observed from Table 2 that, as the magnitude of the shift in the parameter a increases, the performance of both proposed methods improves. However, for all out-of-control scenarios, the performance of the proposed SLRT chart to detect changes in the parameter a is better than the Hotelling's T^2 chart. As we can see, in all out-of-control shifts in parameter a, the performance of the SLRT method in the cases of $\tau = 5$ and $\tau = 10$ is better than $\tau = 15$.

Table 3. Signal probability values under shift from b to $b + \delta_b\sigma_b$.

τ	δ_b	0.5	0.75	1	1.5	2	2.5
5	T^2	0.053	0.058	0.060	0.069	0.079	0.090
	SLRT	0.057	0.058	0.059	0.169	0.861	1
10	T^2	0.061	0.067	0.075	0.083	0.098	0.107
	SLRT	0.053	0.073	0.074	0.11	0.643	1
15	T^2	0.051	0.058	0.074	0.088	0.090	0.101
	SLRT	0.054	0.062	0.067	0.071	0.375	0.965

Table 4. Signal probability values under shift from β to $\beta + \delta_\beta\sigma_\beta$.

τ	δ_β	0.5	0.75	1	1.5	2	2.5
5	T^2	0.058	0.063	0.068	0.071	0.076	0.079
	SLRT	0.060	0.069	0.071	0.072	0.08	0.082
10	T^2	0.057	0.068	0.070	0.073	0.075	0.078
	SLRT	0.070	0.073	0.076	0.076	0.077	0.081
15	T^2	0.052	0.067	0.068	0.070	0.075	0.076
	SLRT	0.056	0.068	0.072	0.073	0.079	0.080

The performance of the proposed control charts to detect different shifts in the parameter b in units of σ_b is compared in Table 3. As expected similar to the results of Tables 1 and 2, the power of the proposed SLRT and Hotelling's T^2 control charts improves as the magnitude of shift in the parameter b increases. It is seen that for all values of parameter τ and under various values of δ_b, the proposed SLRT chart considerably outperforms the Hotelling's T^2 chart. Similar to Table 2, in the cases of $\tau = 5$ and $\tau = 10$, the out-of-control signal probability values obtained from the proposed SLRT control chart are larger than the case when $\tau = 15$.

The signal probability values for both proposed charts under different shift magnitudes in parameter β in units of σ_β are given in Table 4. The results show that similar to the previous Tables, the performance of both methods improves as the magnitude of the shift increases. Table 4 represents that under all out-of-control scenarios, the SLRT chart outperforms the Hotelling's T^2 chart. However, the signal probability values of both methods to detect shifts in parameter β are almost close to each other. It can also be concluded that the SLRT chart under $\tau = 10$ has the best detecting performance when the magnitude of the shift is equal to $\delta_\beta \in \{0.5, 0.75, 1, 1.5\}$ while, in the case of $\delta_\beta \in \{2, 2.5\}$, for $\tau = 5$, this chart outperforms the corresponding results for $\tau = 10$ and $\tau = 15$. Generally, we can conclude that it is better to select $\tau = 5$ and $\tau = 10$ for detecting large and small shifts in parameter β, respectively. Table 4 also shows that the sensitivity of both SLRT and Hotelling's T^2 control charts to detect out-of-control situations in parameter β is smaller than those to detect shifts in the other parameters.

5. Performance evaluation of the proposed method to estimate change point in model parameters

In this section, the performance of the proposed SLRT method, a change point based control chart, to estimate the time of change in the vector of the model parameters is investigated using the same process data of Section 4. For this aim, three criteria namely the accuracy, precision and empirical distribution of estimated change point parameter are used for each regression coefficient and obtained by simulation experiments. The values of accuracy $\mathrm{acc}(\hat{\tau})$ and precision $\mathrm{prec}(\hat{\tau})$ for $\tau \in \{5, 10, 15\}$ are reported in Figures 2–9. Note that these two

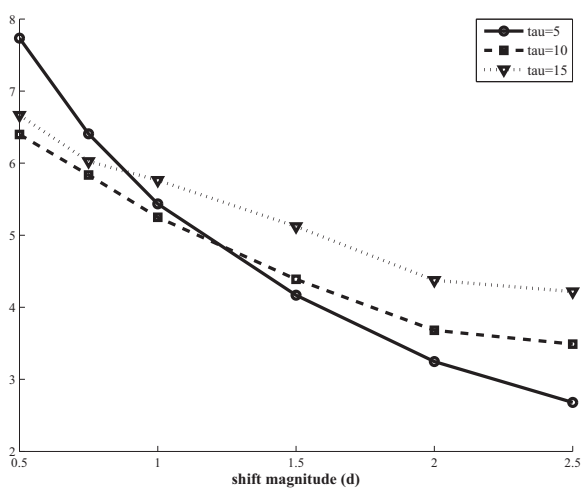

Figure 2. The values of $\mathrm{acc}(\hat{\tau})$ under shifts in the parameter d.

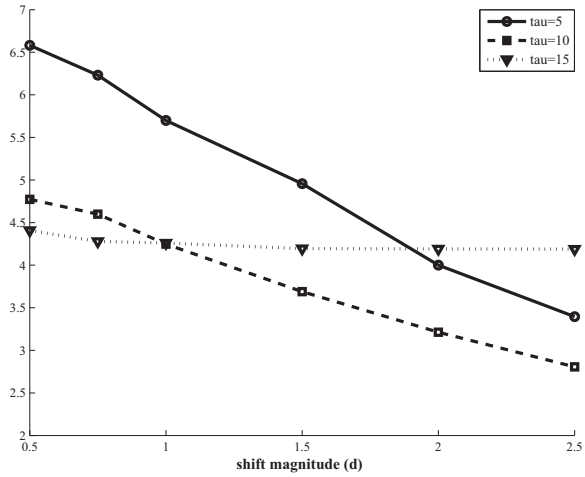

Figure 3. The values of prec($\hat{\tau}$) under shifts in the parameter d.

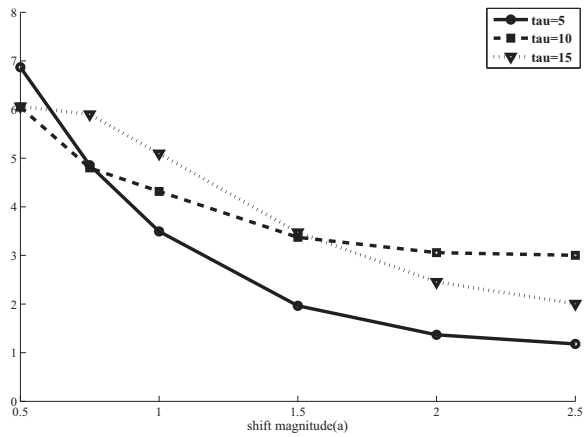

Figure 4. The values of acc($\hat{\tau}$) under shifts in the parameter a.

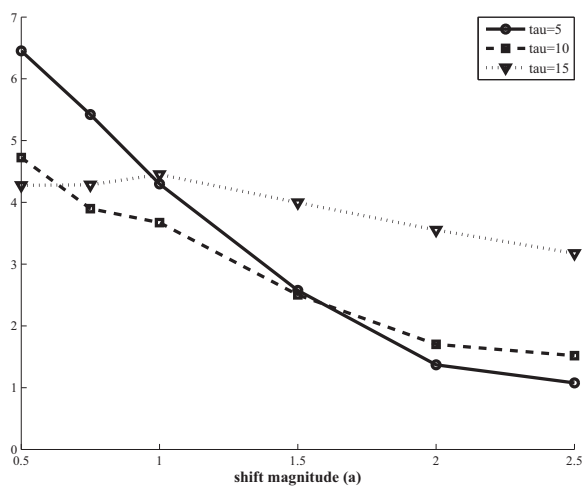

Figure 5. The values of prec($\hat{\tau}$) under shifts in the parameter a.

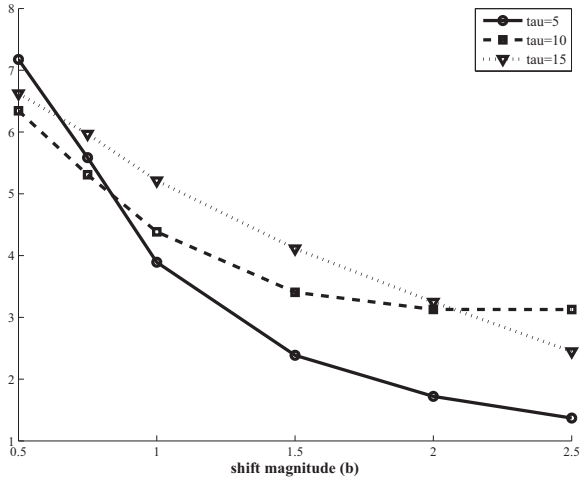

Figure 6. The values of acc($\hat{\tau}$) under shifts in the parameter *b*.

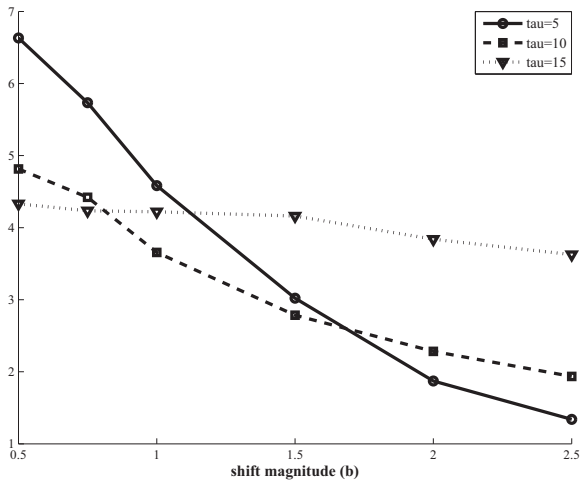

Figure 7. The values of prec($\hat{\tau}$) under shifts in the parameter *b*.

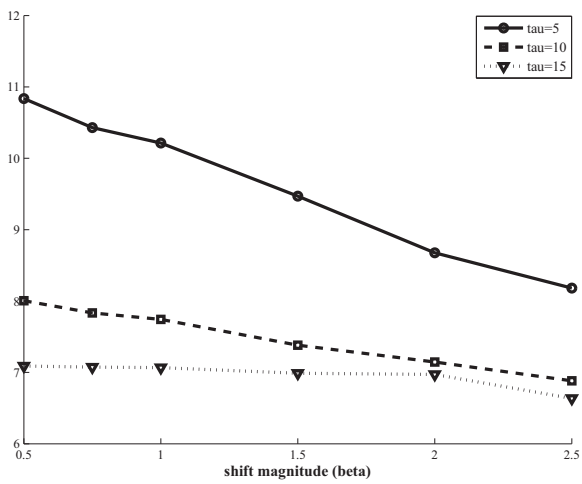

Figure 8. The values of acc($\hat{\tau}$) under shifts in the parameter β.

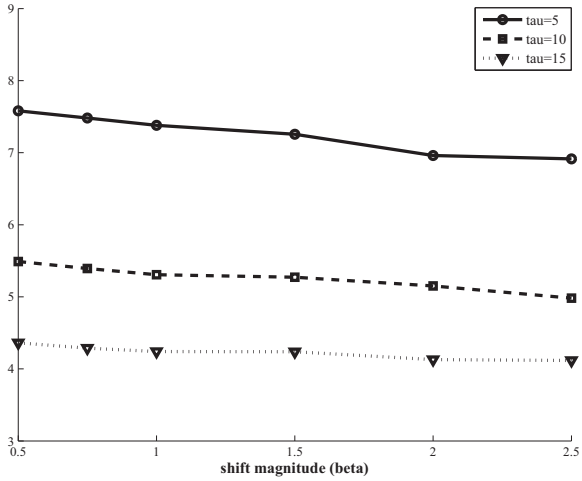

Figure 9. The values of prec$(\hat{\tau})$ under shifts in the parameter β.

criteria are defined as

$$\text{acc}(\hat{\tau}) = \hat{E}(|\hat{\tau} - \tau|) = \sum_{\upsilon=1}^{N} \frac{|\hat{\tau}_\upsilon - \tau|}{N}$$

$$\text{prec}(\hat{\tau}) = \hat{\sigma}(|\hat{\tau} - \tau|) = \sqrt{\sum_{\upsilon=1}^{N} \frac{(|\hat{\tau}_\upsilon - \tau| - \hat{E}(|\hat{\tau} - \tau|))^2}{N-1}}$$

respectively, where $\hat{\tau}_\upsilon$ is the estimated change point at υth replicate, τ is the actual change point and N is the total number of replicates. The cumulative distribution function of $|\hat{\tau} - \tau|$, denoted by $p(|\hat{\tau} - \tau| \leq u)$, is given in Tables 5–8, for $u = 0, 1, \ldots, 5$. The performance of the

Table 5. Empirical distribution of $|\hat{\tau} - \tau|$ under shifts in d.

τ	δ_d	0.5	0.75	1	1.5	2	2.5
5	$p(\lvert\hat{\tau} - \tau\rvert \leq 0)$	0.048	0.085	0.091	0.098	0.121	0.131
	$p(\lvert\hat{\tau} - \tau\rvert \leq 1)$	0.118	0.190	0.237	0.343	0.427	0.520
	$p(\lvert\hat{\tau} - \tau\rvert \leq 2)$	0.176	0.264	0.286	0.413	0.496	0.582
	$p(\lvert\hat{\tau} - \tau\rvert \leq 3)$	0.217	0.320	0.347	0.476	0.550	0.635
	$p(\lvert\hat{\tau} - \tau\rvert \leq 4)$	0.251	0.339	0.373	0.499	0.562	0.643
	$p(\lvert\hat{\tau} - \tau\rvert \leq 5)$	0.295	0.377	0.417	0.526	0.587	0.656
τ	δ_d	0.5	0.75	1	1.5	2	2.5
10	$p(\lvert\hat{\tau} - \tau\rvert \leq 0)$	0.051	0.056	0.057	0.074	0.080	0.086
	$p(\lvert\hat{\tau} - \tau\rvert \leq 1)$	0.101	0.128	0.143	0.198	0.234	0.237
	$p(\lvert\hat{\tau} - \tau\rvert \leq 2)$	0.146	0.171	0.187	0.238	0.268	0.270
	$p(\lvert\hat{\tau} - \tau\rvert \leq 3)$	0.180	0.211	0.216	0.268	0.284	0.289
	$p(\lvert\hat{\tau} - \tau\rvert \leq 4)$	0.210	0.247	0.249	0.297	0.300	0.309
	$p(\lvert\hat{\tau} - \tau\rvert \leq 5)$	0.255	0.285	0.292	0.331	0.338	0.312
τ	δ_d	0.5	0.75	1	1.5	2	2.5
15	$p(\lvert\hat{\tau} - \tau\rvert \leq 0)$	0.065	0.092	0.118	0.178	0.284	0.332
	$p(\lvert\hat{\tau} - \tau\rvert \leq 1)$	0.108	0.137	0.161	0.215	0.309	0.356
	$p(\lvert\hat{\tau} - \tau\rvert \leq 2)$	0.146	0.175	0.187	0.237	0.333	0.378
	$p(\lvert\hat{\tau} - \tau\rvert \leq 3)$	0.181	0.195	0.213	0.263	0.341	0.390
	$p(\lvert\hat{\tau} - \tau\rvert \leq 4)$	0.211	0.234	0.245	0.288	0.360	0.400
	$p(\lvert\hat{\tau} - \tau\rvert \leq 5)$	0.247	0.270	0.274	0.313	0.372	0.407

Table 6. Empirical distribution of $|\hat{\tau} - \tau|$ under shifts in a.

τ	δ_a	0.5	0.75	1	1.5	2	2.5
5	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.058	0.090	0.103	0.124	0.136	0.152
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.168	0.286	0.420	0.685	0.805	0.862
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.222	0.368	0.496	0.708	0.822	0.875
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.273	0.421	0.548	0.729	0.834	0.878
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.300	0.443	0.564	0.742	0.834	0.878
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.339	0.479	0.585	0.754	0.839	0.879
τ	δ_a	0.5	0.75	1	1.5	2	2.5
10	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.050	0.052	0.057	0.060	0.062	0.071
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.112	0.146	0.170	0.181	0.200	0.220
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.148	0.160	0.182	0.235	0.237	0.243
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.148	0.187	0.197	0.249	0.261	0.264
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.150	0.189	0.238	0.259	0.279	0.293
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.151	0.197	0.269	0.283	0.310	0.330
τ	δ_a	0.5	0.75	1	1.5	2	2.5
15	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.103	0.137	0.206	0.428	0.564	0.627
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.142	0.181	0.241	0.445	0.578	0.629
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.168	0.203	0.261	0.456	0.583	0.629
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.207	0.233	0.284	0.463	0.584	0.630
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.248	0.262	0.305	0.471	0.585	0.631
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.278	0.284	0.320	0.480	0.586	0.631

proposed SLRT estimator to identify the time of change in parameter d in terms of accuracy and precision criteria is illustrated in Figures 2 and 3 while the probability values are summarized in Table 5. Figure 2 shows that for the $E(|\hat{\tau} - \tau|)$ criterion and under shift magnitudes of $\delta_d \in \{0.5, 0.75, 1\}$, the SLRT method considering $\tau = 10$ outperforms the other scenarios, i.e., $\tau = 5$ and $\tau = 15$. However, the performance of the proposed estimator in terms of the

Table 7. Empirical distribution of $|\hat{\tau} - \tau|$ under shifts in b.

τ	δ_b	0.5	0.75	1	1.5	2	2.5
5	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.073	0.074	0.092	0.140	0.141	0.165
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.159	0.238	0.344	0.579	0.684	0.772
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.210	0.304	0.417	0.625	0.721	0.799
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.263	0.388	0.488	0.675	0.753	0.824
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.288	0.408	0.512	0.682	0.758	0.824
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.329	0.432	0.532	0.698	0.762	0.828
τ	δ_b	0.5	0.75	1	1.5	2	2.5
10	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.053	0.056	0.068	0.071	0.072	0.08
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.119	0.165	0.183	0.189	0.234	0.239
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.160	0.192	0.223	0.240	0.254	0.287
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.195	0.201	0.253	0.256	0.281	0.302
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.196	0.241	0.264	0.279	0.313	0.314
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.202	0.274	0.284	0.332	0.333	0.357
τ	δ_b	0.5	0.75	1	1.5	2	2.5
15	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.067	0.11	0.156	0.328	0.45	0.591
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.106	0.156	0.207	0.347	0.465	0.599
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.139	0.183	0.233	0.370	0.470	0.604
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.171	0.214	0.254	0.387	0.473	0.606
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.204	0.248	0.279	0.397	0.476	0.606
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.234	0.279	0.305	0.404	0.478	0.607

Table 8. Empirical distribution of $|\hat{\tau} - \tau|$ under shifts in β.

τ	δ_β	0.5	0.75	1	1.5	2	2.5
5	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.037	0.037	0.041	0.041	0.047	0.049
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.082	0.084	0.084	0.099	0.099	0.125
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.114	0.121	0.122	0.145	0.149	0.180
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.143	0.155	0.162	0.190	0.199	0.233
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.165	0.185	0.189	0.218	0.231	0.265
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.209	0.222	0.230	0.250	0.268	0.306
τ	δ_β	0.5	0.75	1	1.5	2	2.5
10	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.032	0.037	0.039	0.041	0.040	0.044
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.081	0.084	0.090	0.091	0.091	0.092
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.113	0.124	0.128	0.136	0.131	0.142
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.145	0.157	0.158	0.163	0.176	0.182
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.172	0.192	0.196	0.203	0.216	0.219
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.219	0.224	0.242	0.246	0.257	0.258
τ	δ_β	0.5	0.75	1	1.5	2	2.5
15	$p(\lvert\hat{\tau} - \tau\rvert \le 0)$	0.039	0.044	0.045	0.046	0.052	0.061
	$p(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.075	0.079	0.083	0.086	0.088	0.097
	$p(\lvert\hat{\tau} - \tau\rvert \le 2)$	0.104	0.120	0.121	0.122	0.132	0.136
	$p(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.137	0.145	0.153	0.158	0.168	0.173
	$p(\lvert\hat{\tau} - \tau\rvert \le 4)$	0.169	0.184	0.184	0.197	0.196	0.208
	$p(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.197	0.220	0.225	0.230	0.239	0.242

$\text{acc}(\hat{\tau})$ criterion for $\tau = 5$ is better than the other scenarios for parameter τ when the magnitude of the shift in parameter d is considered equal to 1.5, 2 and 2.5. In general, for large shifts in parameter d, the performance of SLRT estimator for identifying the change point in parameter d under $\tau = 5$ is better than the other scenarios for parameter τ. However, under small shifts in this parameter, the SLRT method has its best performance to estimate the change point in parameter d when the shift is occurred at $\tau = 10$. Figure 3 shows that the dispersion of $\hat{\tau}$ in the case of $\tau = 10$ is smaller than other ones when $\delta_d \in \{1, 1.5, 2, 2.5\}$. Clearly from Table 5, we can notice that for all values of parameter τ, as the magnitude of the step shift in the parameter d increases, the percentage of exact estimation increases. The probability values summarized in Table 5 shows that in general, the performance of SLRT method for $\tau = 10$ is better than the performance of this method under the other scenarios for parameter τ.

Figure 4 shows that under moderate ($\delta_a \in \{1, 1.5\}$) and large shifts ($\delta_a \in \{2, 2.5\}$) in parameter a, the proposed estimator considering $\tau = 5$ performs better than the other scenarios of parameter τ in terms of the $\text{acc}(\hat{\tau})$ criterion. From Figure 5, we can conclude that the SLRT method for $\tau = 5$ performs with more accuracy than the other cases of parameter τ when δ_a is considered equal to 2 and 2.5. However for shifts $\delta_a \in \{0.75, 1, 1.5\}$, the best performance of this method in terms of the $\text{prec}(\hat{\tau})$ criterion is obtained when the change is induced at profile $\tau = 10$. The probability values reported in Table 6 indicates that the performance of the proposed method to identify the change point in the parameter a is satisfactory.

As seen in Figure 6, the best performance of the SLRT method in terms of accuracy criterion for shift magnitudes $\delta_b \in \{1, 1.5, 2, 2.5\}$ is obtained when $\tau = 5$. Figure 7 shows that under small shifts ($\delta_b \in \{0.5, 0.75\}$), moderate shifts ($\delta_b \in \{1, 1.5\}$) and large shifts ($\delta_b \in \{2, 2.5\}$), the SLRT method has its best performance when $\tau = 15$, $\tau = 10$ and $\tau = 5$, respectively. As expected, Table 7 reveals that as the magnitude of the shift in parameter b increases,

the obtained probability values will also increase. Table 7 also shows that the obtained results in terms of the probability values under first scenario ($\tau = 5$) outperforms the others.

The expected value and the standard deviation for the difference of actual and estimated change point parameter are depicted in Figures 8 and 9, respectively. From these figures it can be concluded that the performance of the proposed estimator in terms of both accuracy and precision criteria is deteriorated as the value of parameter τ decreases. The obtained probability values given in Table 8 shows that the capability of the SLRT method to identify the time of change in the parameter β is not as satisfactory as those when the shift is induced in the other model parameters.

6. Real data example

In this section, the performance of the proposed framework is illustrated by a real data example concerning the number of thefts in two cities in Australia, namely Albury and Armidale Dumaresq. In this data set (considered as a Phase I analysis and available in data.gov.au), the monthly number of receiving or handing stolen goods are gathered during period 1995 to 2015. Hence, we have $m = 21$ profiles (each year is considered as a profile) such that each one contains $n = 12$ observations. The number of thefts in Albury and Armidale Dumaresq are considered as the response and explanatory variable, respectively. Here, due to the three following reasons, we fit an autocorrelated Poisson regression profile using an INGARCH(1,1) model given in Equation (1) to the data set:

1. The data set (the number of thefts in Albury) only contains positive integers which can be modeled using a Poisson distribution.
2. Based on experts feedback, the number of thefts for a given month is clearly affected by those of the preceding months (the data are autocorrelated).
3. Since Albury and Armidale Dumaresq are nearby, they have lots of similarities such as the same national cultures, etc. Hence, the number of thefts in Albury in a given month can be affected by those in Armidale Dumaresq.

In addition, the mean value of the studentized residuals of this nonlinear model (Figure 10) shows that the values of these residuals are not considerably large. This may confirm the accuracy of this model for the data.

It is also important to note that two kinds of causes, local and global ones, impact the count of crimes in these cities. Studying the counts of thefts in two different cities with different locations but the same calendar (having the same holidays, national days, etc.) and

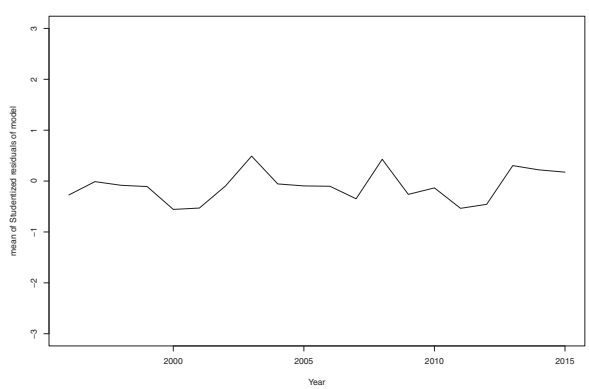

Figure 10. The mean of studentized residuals of the model over time.

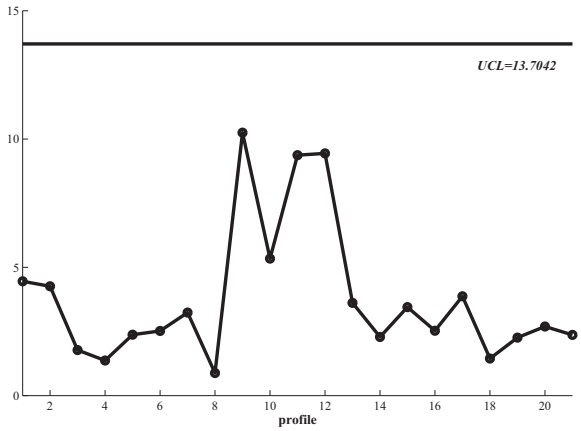

Figure 11. T^2 statistic values for data example.

national cultures helps us to monitor the global causes by considering the variation of the counts in other cities. It is worth mentioning that the global causes can be produced by many unknown and unmeasurable reasons such, for instance, the date of hot seasons, the rate of tourism in some days, the date of holidays and the governmental rules where all of them are almost the same in two seasons but it is not possible to quantify the effect of these sources. Therefore, we add the term $\mathbf{x}_i^\intercal \boldsymbol{\beta}$, the number of crimes in Armidale Dumaresq, to consider the total effect of these sources by considering the variation of the number of crimes in the city. First of all, from the data set, we obtain $\widehat{\boldsymbol{\theta}} = (1.0772, 0.4323, 0.3046, 0.2975)^\intercal$ using the PSO algorithm. Here, any changes in the vector of model parameters in Equation (1) indicates an out-of-control signal. After obtaining $\widehat{\boldsymbol{\theta}}$, using simulations, we have generated 10000 replicates of $m = 21$ profiles with $n = 12$ experimental settings in order to obtain the values of UCL for both the proposed SLRT and Hotelling's T^2 charts such that the probability of Type-I error, $\alpha = 0.05$. Afterwards, these control charts are applied on the real data and the profiles whose corresponding statistics exceeds UCL are eliminated. In this case, the value of vector $\widehat{\boldsymbol{\theta}}$ is recomputed based on the remaining profiles and then the UCL values of the proposed control charts are updated. This procedure is repeated until no more out-of-control signal occurs in the proposed control charts. The maximum value of $SLRT_\tau$; $\tau \in \{1, 2, \ldots, 20\}$ for this data set is obtained equal to 0.3710 while the $UCL = 5.3772$. The values of Hotelling's T^2 statistic for the data example is depicted in Figure 11. As seen, for the Hotelling's T^2 chart, the value of all T^2 statistics are smaller than the $UCL = 13.7042$. Since for both proposed control charts, the data set are in-control, the estimated parameters in Phase I analysis can be applied to monitor the upcoming profiles in Phase II.

7. Conclusion remarks and recommendations for future studies

In this paper, we studied the Phase I monitoring of autocorrelated Poisson regression profiles considering an INGARCH(1,1) model under within-profile autocorrelation. Two control charts, namely the SLRT and the Hotelling's T^2 charts, were proposed to detect different step shifts in the vector of model parameters. We investigated and compared the efficiency of the proposed charts in terms of the out-of-control signal probability criterion. Moreover, we investigated the capability of the proposed SLRT chart in estimating the time of change in the vector of regression model parameters. Based on the obtained results, we found that the detecting capability of the proposed SLRT chart is superior than the one of the Hotelling's

T^2 chart. The results also indicated that the performance of the SLRT method to identify the change point in the vector of model parameters is satisfactory, except for parameter β. Finally, we illustrated the application of the proposed methods in a Phase I analysis of autocorrelated Poisson regression profile by a real case study. The future study would be directed to considering between-profile autocorrelation for monitoring Poisson regression profiles. Besides, the effect of parameter estimation on the performance of control charts for monitoring autocorrelated Poisson regression profiles is also recommended as a future research.

ORCID

A. Amiri http://orcid.org/0000-0002-2385-8910
A. R. Taheriyoun http://orcid.org/0000-0002-2282-7741
P. Castagliola http://orcid.org/0000-0002-9532-4029

References

Abdel-Salam, A. S. G., J. B. Birch, and W. A. Jensen. 2013. A semiparametric mixed model approach to phase i profile monitoring. *Quality and Reliability Engineering International* 29 (4):555–69.

Amiri, A., W. A. Jensen, and R. B. Kazemzadeh. 2010. A case study on monitoring polynomial profiles in the automotive industry. *Quality and Reliability Engineering International* 26 (5):509–20.

Amiri, A., M. Koosha, and A. Azhdari. 2011. Profile monitoring for poisson responses. In *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Singapore, Singapore, 1481–84.

Amiri, A., M. Koosha, A. Azhdari, and G. Wang. 2015. Phase I monitoring of generalized linear model-based regression profiles. *Journal of Statistical Computation and Simulation* 85 (14):2839–59.

Eberhart, R., and J. Kennedy. 1995. A new optimizer using particle swarm theory. In *IEEE Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 39–43.

Ferland, R., A. Latour, and D. Oraichi. 2006. Integer-valued GARCH process. *Journal of Time Series Analysis* 27 (6):923–42.

Jensen, W. A., J. B. Birch, and W. H. Woodall. 2008. Monitoring correlation within linear profiles using mixed models. *Journal of Quality Technology* 40 (2):167–83.

Keramatpour, M., S. T. A. Niaki, M. Khedmati, and M. Soleymanian. 2013. Monitoring and change point estimation of AR(1) autocorrelated polynomial profiles. *International Journal of Engineering-Transactions C: Aspects* 26 (9):933–42.

Keramatpour, M., S. T. A. Niaki, and A. Amiri. 2014. Phase-II monitoring of AR(1) autocorrelated polynomial profiles. *Journal of Optimization in Industrial Engineering* 7 (14):53–9.

Khedmati, M., and S. T. A. Niaki. 2016. Phase II monitoring of general linear profiles in the presence of between-profile autocorrelation. *Quality and Reliability Engineering International* 32 (2):443–52.

Koosha, M., and A. Amiri. 2013. Generalized linear mixed model for monitoring autocorrelated logistic regression profiles. *The International Journal of Advanced Manufacturing Technology* 64 (1–4): 487–95.

Maleki, M. R., A. Amiri, and A. R. Taheriyoun. 2017a. Phase II monitoring of binary profiles in the presence of within-profile aautocorrelation based on markov model. *To appear in Communications in Statistics-Simulation and Computation*.

Maleki, M. R., A. Amiri, and A. R. Taheriyoun. 2017b. Identifying the time of step change and drift in phase II monitoring of autocorrelated logistic regression profiles. Submitted.

Narvand, A., P. Soleimani, and S. Raissi. 2013. Phase II monitoring of auto-correlated linear profiles using linear mixed model. *Journal of Industrial Engineering International* 9 (1):1–9.

Noorossana, R., A. Amiri, and P. Soleimani. 2008. On the monitoring of autocorrelated linear profiles. *Communications in Statistics—Theory and Methods* 37 (3):425–42.

Noorossana, R., A. Saghaei, and M. Dorri. 2010. Linear profile monitoring in the presence of non-normality and autocorrelation. *International Journal of Industrial Engineering and Production Research* 21 (4):221–30.

Noorossana, R., A. Saghaei, and A. Amiri. 2011. *Statistical analysis of profile monitoring*, Vol. 865. Hoboken, New Jersey: John Wiley & Sons.

Soleimani, P., and R. Hadizadeh. 2014. Monitoring simple linear profiles in the presence of garch and non-normality effects. In *International Conference on Control, Decision and Information Technologies (CoDIT)*, 393–99. Metz, France: IEEE.

Soleimani, P., and R. Noorossana. 2014. Monitoring multivariate simple linear profiles in the presence of between profile autocorrelation. *Communications in Statistics—Theory and Methods* 43 (3):530–46.

Soleimani, P., R. Noorossana, and A. Amiri. 2009. Simple linear profiles monitoring in the presence of within profile autocorrelation. *Computers & Industrial Engineering* 57 (3):1015–21.

Soleimani, P., A. Narvand, and S. Raissi. 2013a. Online monitoring of auto correlated linear profiles via mixed model. *The International Journal of Manufacturing Technology and Management* 27 (4–6):238–50.

Soleimani, P., R. Noorossana, and S. T. A. Niaki. 2013b. Monitoring autocorrelated multivariate simple linear profiles. *The International Journal of Advanced Manufacturing Technology* 67 (5–8):1857–65.

Woodall, W. H. 2007. Current research on profile monitoring. *Production* 17 (3):420–25.

Woodall, W. H., D. J. Spitzner, D. C. Montgomery, and S. Gupta. 2004. Using control charts to monitor process and product quality profiles. *Journal of Quality Technology* 36 (3):309–20.

Yeh, A. B., L. Huwang, and Y. M. Li. 2009. Profile monitoring for a Binary response. *IIE Transactions* 41 (11):931–41.

Zhang, Y., Z. He, C. Zhang, and W. H. Woodall. 2014. Control charts for monitoring linear profiles with within-profile correlation using gaussian process models. *Quality and Reliability Engineering International* 30 (4):487–501.

Zhu, F. 2012. Modeling time series of counts with COM-poisson INGARCH models. *Mathematical and Computer Modelling* 56 (9):191–203.