Journal de la Société Française de Statistique
Vol. 156 No. 4 (2015)

Sequential change-point detection in Poisson autoregressive models

Titre: Détection séquentielle de ruptures dans les models autorégressifs de Poisson

William Kengne[1] *

Abstract: We consider the sequential change-point detection in a general class of Poisson autoregressive models. The conditional mean of the process depends on a parameter $\theta_0^* \in \Theta \subset \mathbb{R}^d$ which may change over time as and when data are observed. We propose a closed and open-end procedure based on the maximum likelihood estimator of the parameter. Under the null hypothesis of no change, it is shown that the detector converges to a well know distribution. The (empirical) power and the efficiency in terms of the detection delay are assessed through a simulation study and a real data example is provided.

Résumé : Nous considérons la détection séquentielle de ruptures dans une classe assez générale de modèles de Poisson autorégressifs de séries temporelles à valeurs entières. La moyenne conditionnelle du processus dépend d'un paramètre $\theta_0^* \in \Theta \subset \mathbb{R}^d$ susceptible de changer dans le temps au fur et à mesure que les données sont observées. Nous proposons une procédure séquentielle dont le temps de suivi peut être fini ou infini basée sur l'estimateur du maximum de vraisemblance du paramètre. Sous l'hypothèse nulle selon laquelle aucun changement n'intervient dans le paramètre, la statistique de test converge vers une distribution connue. Des résultats de simulations nous permettent d'évaluer la puissance (empirique) ainsi que l'efficacité en terme du délai de détection et un exemple d'application aux données réelles est fourni.

Keywords: Sequential detection, change-point, time series of counts, Poisson autoregression, likelihood estimation
Mots-clés : Detection séquentielle, rupture, séries temporelles à valeurs entières, autoregression de Poisson, estimation par vraisemblance
AMS 2000 subject classifications: 62M10, 62L12, 62F05

1. Introduction

Detecting change in time series has become an important research topic in statistics in the last three decades, since it has been known that many data often suffer from structural change. Before any statistical inference, one must test if a change has not occurred in the model during the data generating period. Two approaches are generally considered for solving this problem : the off-line (or retrospective) detection and the on-line (or sequential) detection. The off-line detection is an approach used when all data are available ; see the book of Csörgö and Horváth (1997) for a large overview and Aue et al. (2009a); Bardet et al. (2012); Kengne (2012); Fryzlewicz and Subba Rao (2013); Fokianos et al. (2014) among others, for some recent works that have been done in this setting. The on-line (sequential) detection that we will focus here, refers to the change-point detection as and when new data are observed.

* This work has been developed within the Labex MME-DII (ANR-11-LBX-0023-01) ; `http://labex-mme-dii.u-cergy.fr/`.
[1] THEMA, Université de Cergy-Pontoise, 33 Boulevard du Port, 95011 Cergy-Pontoise Cedex, France.
E-mail: `william.kengne@u-cergy.fr`

Numerous works have been devoted to the sequential change-point detection, we refer to the book of Basseville et al. (1993) for a large surveys. But the important turning on this topic has been done with the paper of Chu et al. (1996). They have pointed out that repeating the retrospective test as new data are observed will increase the probability of type I error of the procedure. For the regression model, they addressed the sequential change detection as a classical hypothesis testing with a fixed probability of type I error and proposed two monitoring procedures based on cumulative sum (CUSUM) of residuals and recursive parameter fluctuations. Their approach has been generalized and extended in several directions. See Leisch et al. (2000); Horváth et al. (2004); Zeileis et al. (2005); Aue et al. (2006); Horvath et al. (2007); Aue et al. (2009b) for some various procedures for sequential change in linear models with independent and dependent innovations. Berkes et al. (2004) and Gombay and Serban (2009) focussed respectively on the sequential change-point detection in the parameters of GARCH and linear autoregressive processes ; whereas Na et al. (2011) proposed fluctuation-type test procedure for sequential change detection in a general class of time series models. Bardet and Kengne (2014) considered a large class of causal models (that including AR(∞), ARCH(∞), TARCH(∞),... processes) and developed a sequential procedure where the updated estimator is computed without the historical observations ; the consistency of the procedure has been established.

In this paper, we focus on sequential change detection in a large class of Poisson autoregressive models. More precisely, we consider a process $Y = (Y_t)_{t\in\mathbb{Z}}$ satisfying :

$$Y_t/\mathscr{F}_{t-1} \sim \text{ Poisson}(\lambda_t) \text{ with } \lambda_t = f_{\theta_0^*}(Y_{t-1},\dots) \tag{1}$$

where $\mathscr{F}_t = \sigma(Y_s,\ s \leq t)$ is the σ-field generated by the whole past of the process and f is a measurable non-negative function assumed to be know up to some parameter θ_0^* belonging to a compact set $\Theta \subset \mathbb{R}^d$.
Let $\{N_t(\cdot)\ ;\ t = 1,2,\cdots\}$ be a sequence of independent Poisson processes of unit intensity. Y_t can be seen as the number of events of $N_t(\cdot)$ in the time interval $[0,\lambda_t]$. The feedback mechanism in the conditional mean λ_t provides a large tool for modeling dependence structure in count events (number of new infections, number of transactions per minute, number of defect products,$\cdots$). The model with one order autoregression has been addressed by Fokianos et al. (2009); Fokianos and Tjøstheim (2012). They investigated the asymptotic properties of the maximum likelihood estimator by using a perturbation approach. Whereas Doukhan et al. (2012) use a weak dependence approach to prove the existence of a stationary solution of a class of model larger than (1). Neumann et al. (2011) focused on the absolute regularity and ergodicity of the model with one order autoregression. For surveys on the change-point and outliers detection in integer-valued autoregressive models and relate, we refer to Kang and Lee (2009, 2014); Fokianos and Fried (2010, 2012); Doukhan and Kengne (2015).
The previous papers have worked out the retrospective change-point detection on variants of the model (1) and related models. Sequential detection seems to be very important in finance, economics or epidemiology. For example, in the real data application in finance (see below), it is crucial to know early if the high variations in the number of transactions are noise effects or are due to a structural change in the dynamic of the transaction process, before taking an appropriate decision.

Now, assume that a trajectory $(Y_1,\cdots,Y_n)$ of the process Y generated as in (1) with a parameter θ_0^* has been observed. In the setting of the sequential detection, these available observations are called the **historical data**. New data $X_{n+1},X_{n+2},\cdots,X_k,\cdots$ will be observed and monitoring scheme starts at time $n+1$. For each new observation, one would like to know if it has been generated by the model depending on θ_0^* or some other parameter θ_1^* (with $\theta_1^* \neq \theta_0^*$). More precisely, we consider the following test problem :

$\mathbf{H_0}$: θ_0^* is constant over the observations $Y_1,\cdots,Y_n,Y_{n+1},\cdots$ *i.e.* $(Y_n)_{n\in\mathbb{N}}$ satisfying (1) with the parameter θ_0^* ;

$\mathbf{H_1}$: there exist $k^*>n$, $(\theta_0^*,\theta_1^*)\in\Theta^2$, with $\theta_0^*\neq\theta_1^*$, such that $(Y_1,\cdots,Y_{k^*})$ satisfying (1) with the parameter θ_0^* and and $(Y_{k^*+n})_{n\in\mathbb{N}}$ satisfying (1) with θ_1^*.

It is known that under H_0, the maximum likelihood estimator (MLE in the sequel) of the parameter θ_0^* is consistent (see Doukhan and Kengne, 2015). Therefore, as the fluctuation procedure proposed by Chu et al. (1996), we define a detector based on the difference between the historical and the updated parameter estimator. More precisely, at the time $k>n$, we compare the estimator computed on the historical data to that computed with all the data up to the time k (i.e. $X_1,\cdots,X_n,X_{n+1},\cdots,X_k$). Thus, under H_0 these two quantities are consistent estimators of θ_0^*, so they will be close and the detector will not be large enough. We will show through simulation study that the detector seems to be large enough under H_1. Hence, if the detector is larger than a suitable critical value, then H_0 is rejected and a model with a new parameter is considered; otherwise, the sequential monitoring scheme continues.

2. Assumptions and examples

2.1. Assumptions

We will use the following classical notations:

1. $\|y\|:=\sum_{j=1}^{p}|y_j|$ for any $y\in\mathbb{R}^p$;
2. for any compact set $\mathscr{K}\subseteq\mathbb{R}^d$ and for any function $g:\mathscr{K}\longrightarrow\mathbb{R}^{d'}$, $\|g\|_{\mathscr{K}}=\sup_{\theta\in\mathscr{K}}(\|g(\theta)\|)$;
3. for any set $\mathscr{K}\subseteq\mathbb{R}^d$, $\overset{\circ}{\mathscr{K}}$ denotes the interior of $\mathscr{K}$.

Throughout the sequel, we will assume that the function $\theta\mapsto f_\theta$ is twice continuously differentiable on Θ. The following Lipschitz-type condition $\mathrm{A}_0(\Theta)$ is classical to ensure the existence of solution of such model (see for instance Doukhan and Wintenberger, 2008) and the assumptions $\mathrm{A}_1(\Theta)$, $\mathrm{A}_2(\Theta)$ as well as $\mathrm{D}(\Theta)$, $\mathrm{Id}(\Theta)$ and $\mathrm{Var}(\Theta)$ are needed for inference on the model see Doukhan and Kengne (2015).
For $i=0,\,1,\,2$ and for any compact set $\mathscr{K}\subseteq\Theta$, define
Assumption $\mathbf{A}_i(\mathscr{K})$: $\|\partial^i f_\theta(0)/\partial\theta^i\|_\Theta<\infty$ and there exists a sequence of non-negative real numbers $(\alpha_k^{(i)}(\mathscr{K}))_{k\geq1}$ satisfying $\sum_{j=1}^{\infty}\alpha_k^{(0)}(\mathscr{K})<1$ (when $i=0$) and $\sum_{j=1}^{\infty}\alpha_k^{(i)}(\mathscr{K})<\infty$ (when

$i = 1,2$) such that

$$\left\| \frac{\partial^i f_\theta(y)}{\partial \theta^i} - \frac{\partial^i f_\theta(y')}{\partial \theta^i} \right\|_{\mathscr{K}} \leq \sum_{k=1}^{\infty} \alpha_k^{(i)}(\mathscr{K})|y_k - y'_k| \quad \text{for all } y, y' \in (\mathbb{R}^+)^{\mathbb{N}}.$$

Assumption D(Θ): $\exists \underline{c} > 0$ such that $\inf_{\theta \in \Theta}(f_\theta(y)) \geq \underline{c}$ for all $y \in (\mathbb{R}^+)^{\mathbb{N}}$.

Assumption Id(Θ): For all $(\theta, \theta') \in \Theta^2$, $\Big(f_\theta(Y_{t-1}, \dots) = f_{\theta'}(Y_{t-1}, \dots)$ a.s. for some $t \in \mathbf{Z}\Big) \Rightarrow \theta = \theta'$.

Assumption Var(Θ): For all $\theta \in \Theta$ and $t \in \mathbf{Z}$, the components of the vector $\dfrac{\partial f_\theta}{\partial \theta^i}(Y_{t-1,\dots})$ are a.s. linearly independent.

2.2. *Examples*

2.2.1. *Linear Poisson autoregression*

We consider an integer-valued time series $(Y_t)_{t \in \mathbf{Z}}$ satisfying for any $t \in \mathbf{Z}$

$$Y_t / \mathscr{F}_{t-1} \sim \text{Poisson}(\lambda_t) \text{ with } \lambda_t = \phi_0(\theta_0^*) + \sum_{k \geq 1} \phi_k(\theta_0^*) Y_{t-k} \tag{2}$$

where $\theta_0^* \in \Theta \subset \mathbb{R}^d$, the functions $\theta \mapsto \phi_k(\theta)$ are positive and satisfying $\sum_{k \geq 1} \|\phi_k(\theta)\|_\Theta < 1$. This model is also called an INARCH(∞), due to its similarity with the classical ARCH(∞) model. Assumptions $A_0(\Theta)$ holds automatically. If the function ϕ_k are twice continuous differentiable such that $\sum_{k \geq 1} \|\phi'_k(\theta)\|_\Theta < \infty$ and $\sum_{k \geq 1} \|\phi''_k(\theta)\|_\Theta < \infty$, then $A_1(\Theta)$ and $A_2(\Theta)$ hold. If $\inf_{\theta \in \Theta} \phi_0(\theta) > 0$ then D(Θ) holds. Moreover, if there exists a finite subset $I \subset \mathbb{N} - \{0\}$ such that the function $\theta \mapsto (\phi_k(\theta))_{k \in I}$ is injective, then assumption Id(Θ) holds and the model (2) is identifiable.
Note that the classical Poisson INGARCH(p,q) (see Ferland et al., 2006 or Weiß, 2009) obtained with

$$\lambda_t = \alpha_0(\theta_0^*) + \sum_{k=1}^{p} \alpha_k(\theta_0)\lambda_{t-k} + \sum_{k=1}^{q} \beta_k(\theta_0^*) Y_{t-k}$$

is a special case of the model (2) if the condition $\sum_{k=1}^{p} \|\alpha_k(\theta)\|_\Theta + \sum_{k=1}^{q} \|\beta_k(\theta)\|_\Theta < 1$ is satisfied. Fokianos and Fried (2010) focussed on the intervention effects (that can generated sudden mean shift or outliers see Fokianos and Fried, 2010) in this model whereas the diagnostic checking when $p = 0$ has been addressed by Zhu and Wang (2010). The INGARCH(1,1) is known to describe adequately the number of transactions in the stock Ericsson B (see Fokianos et al., 2009, 2013; Doukhan and Kengne, 2015).

2.2.2. *Poisson exponential autoregressive model*

This model, proposed by Fokianos et al. (2009) is defined by

$$Y_t / \mathscr{F}_{t-1} \sim \text{Poisson}(\lambda_t) \text{ with } \lambda_t = \big(\alpha_0 + \alpha_1 \exp(-\gamma \lambda_{t-1}^2)\big)\lambda_{t-1} + \beta Y_{t-1} \tag{3}$$

where $\alpha_0, \alpha_1, \beta, \gamma > 0$ are the parameters of the model. If $\alpha_0 + \alpha_1 + \beta < 1$, then the Lipschitz-type conditions above hold. Moreover, the assumptions D(Θ), Id(Θ) and Var(Θ) hold. See Fokianos et al. (2009) for an application of this model to transactions data.

3. The sequential procedure and the main results

3.1. Likelihood estimation

For any integer $\ell \geq 1$, denote
$$T_\ell = \{1, 2, \cdots, \ell\}.$$

Let $k \geq n \geq 1$. If $(X_1, \cdots, X_k)$ is generated according to (1) with the parameter θ, then for any $\ell \leq k$, the conditional (log)-likelihood (up to a constant) computed on T_ℓ, is given by

$$L(T_\ell, \theta) = \sum_{t=1}^{\ell} \big(Y_t \log \lambda_t(\theta) - \lambda_t(\theta)\big) = \sum_{t=1}^{\ell} \ell_t(\theta) \text{ with } \ell_t(\theta) = Y_t \log \lambda_t(\theta) - \lambda_t(\theta)$$

where $\lambda_t(\theta) = f_\theta(Y_{t-1}, \ldots)$. In the sequel, we use the notation $f_\theta^t := f_\theta(Y_{t-1}, \ldots)$. This (log)-likelihood can be approximated by (see also Doukhan and Kengne, 2015)

$$\widehat{L}(T_\ell, \theta) = \sum_{t=1}^{\ell} \big(Y_t \log \widehat{\lambda}_t(\theta) - \widehat{\lambda}_t(\theta)\big) \tag{4}$$

where $\widehat{\lambda}_t(\theta) := \widehat{f}_\theta^t := f_\theta(Y_{t-1}, \ldots, Y_1, 0, \ldots)$ and $\widehat{\lambda}_1(\theta) = f_\theta(0, \ldots)$. The maximum likelihood estimator of the parameter computed on T_ℓ is defined by

$$\widehat{\theta}(T_\ell) = \text{argmax}_{\theta \in \Theta}(\widehat{L}(T_\ell, \theta)). \tag{5}$$

The two following results have been established by Doukhan and Kengne (2015). Under H_0, if $\theta_0^* \in \overset{\circ}{\Theta}$ and $D(\Theta)$, $\text{Id}(\Theta)$, $\text{Var}(\Theta)$ and $A_i(\Theta)$ $i = 0, 1, 2$ hold with

$$\sum_{j \geq 1} \sqrt{j} \times \alpha_j^{(i)}(\Theta) < \infty \tag{6}$$

then

$$\widehat{\theta}(T_n) \xrightarrow[n \to +\infty]{a.s.} \theta_0^*$$

and

$$\sqrt{n}(\widehat{\theta}(T_n) - \theta_0^*) \xrightarrow[n \to +\infty]{\mathcal{D}} \mathcal{N}(0, \Sigma^{-1}) \tag{7}$$

where $\Sigma = E\Big(\frac{1}{f^0_{\theta_0^*}}\big(\frac{\partial}{\partial \theta} f^0_{\theta_0^*}\big)\big(\frac{\partial}{\partial \theta} f^0_{\theta_0^*}\big)'\Big)$ and where $'$ denotes the transpose. Under the above conditions, the matrix

$$\widehat{\Sigma}_n = \Big(\frac{1}{n} \sum_{t=1}^{n} \frac{1}{\widehat{f}_\theta^t} \big(\frac{\partial}{\partial \theta} \widehat{f}_\theta^t\big)\big(\frac{\partial}{\partial \theta} \widehat{f}_\theta^t\big)'\Big)\Big|_{\theta = \widehat{\theta}(T_n)}$$

is a consistent estimator of Σ. These results will be the main tools for the following sequential procedure.

3.2. *The sequential procedure*

In the sequel, $(Y_1,\cdots,Y_n)$ is supposed to be the historical available observations generated according to (1) with the parameter θ_0^*. At a monitoring instant k, we will assess the difference between the two estimators $\widehat{\theta}(T_k)$ and $\widehat{\theta}(T_n)$. More precisely, according to (7), we define for any $k > n$ the statistic $\widehat{D}_k$, called the **detector** by

$$\widehat{D}_k := \sqrt{n}\left\|\widehat{\Sigma}_n^{1/2}\left(\widehat{\theta}(T_k)-\widehat{\theta}(T_n)\right)\right\|.$$

This detector is well defined since the matrix $\widehat{\Sigma}_n$ is asymptotically symmetric and positive definite (see Doukhan and Kengne, 2015). Note that, if change does not occur at time $k > n$, both the estimators $\widehat{\theta}(T_k)$ and $\widehat{\theta}(T_n)$ are close and the detector $\widehat{D}_k$ is not too large.
Let $T > 1$ (T can be equal to infinity). The sequential monitoring scheme rejects H_0 at the first time k satisfying $n < k \leq [Tn]+1$ such that $\widehat{D}_k > c$ for a suitably chosen constant $c > 0$, where $[x]$ denote the integer part of x. The procedure is called **closed-end method** when $T < \infty$ and **open-end method** when $T = \infty$. The set $\{n+1,n+2,\cdots,[Tn]\}$ is called the monitoring period. The choice of T depends on the time that we when to monitor the procedure. Note that, this choice will not affect the efficiency of the procedure ; as we will see below, the critical value of the test takes into account the choice of T. But in practice, we cannot monitor until infinity thus, we have to choose $T < \infty$.
To build a detector that is sensitive to detect changes that occur at the beginning of the monitoring and those occur a long time after the beginning of the monitoring, we use the so-called boundary function $b : [1,\infty) \mapsto (0,\infty)$, assumed to be continuous and satisfying :

$$\underset{1\leq t<\infty}{\mathrm{Inf}}\ b(t) > 0.$$

Unlike Bardet and Kengne (2014), the decreasing assumption is not imposed to b.

The monitoring scheme rejects H_0 at the first time k (with $n < k \leq [Tn]+1$) such as $\widehat{D}_k > b(k/n)$. Hence define the stopping time:

$$\tau(n) := \mathrm{Inf}\left\{n < k \leq [Tn]+1 \ \big/ \ \widehat{D}_k > b(k/n)\right\}$$

with the convention that $\mathrm{Inf}\{\emptyset\} = \infty$. Therefore, we have

$$\begin{aligned} P\{\tau(n) < \infty\} &= P\Big\{ \frac{\widehat{D}_{k,\ell}}{b(k/n)} > 1 \ \text{ for some } k \ \text{ between } n \text{ and } [Tn]+1\Big\} \\ &= P\Big\{ \sup_{n<k\leq[Tn]+1} \frac{\widehat{D}_k}{b(k/n)} > 1\Big\}. \end{aligned} \tag{8}$$

The main aim is to choose a suitable boundary function $b(\cdot)$ such as for some given $\alpha \in (0,1)$,

$$\lim_{n\to\infty} P_{H_0}\{\tau(n) < \infty\} = \alpha \tag{9}$$

and $P_{H_1}\{\tau(n) < \infty\}$ is asymptotically close to one ; where the hypothesis H_0 and H_1 are specified in Section 1.

If the break is suspected to occur soon after the beginning of the monitoring, a boundary function b with the smallest values in the neighborhood of 1 is appropriate. Otherwise, choose a function b that takes its largest values in the neighborhood of 1. But, in practice, we do not often know the instant of the break. So, one will choose a constant boundary function $b \equiv c$ with $c > 0$, this leads to compute a threshold $c = c_\alpha$ that satisfying (9).
Moreover, if a change-point is detected under H_1 *i.e.* $\tau(n) < \infty$ and $\tau(n) > k^*$, then the detection delay is defined by

$$\widehat{d_n} = \tau(n) - k^*. \tag{10}$$

$\widehat{d_n}$ is used to assess the efficiency of the procedure to early detect changes in the model. The smaller is the detection delay, the better is the efficiency under the alternative.

3.3. Main results

Under H_0, the parameter θ_0^* does not change over the new observations. The following theorem displays the asymptotic behavior under the null hypothesis of the detector $\widehat{D}_k$ for the open and closed-end procedure.

Theorem 3.1. *Assume* $D(\Theta)$, $\mathrm{Id}(\Theta)$, *Var*(Θ) *and* $A_i(\Theta)$ $i = 0,1,2$ *hold with*

$$\sum_{j\geq 1} \sqrt{j} \times \alpha_j^{(i)}(\Theta) < \infty.$$

Under H_0 *with* $\theta_0^* \in \overset{\circ}{\Theta}$, *for the open-end* ($T = \infty$) *and closed-end* ($T < \infty$) *procedure it holds that*

$$\lim_{n\to\infty} P\{\tau(n) < \infty\} = P\Big\{ \sup_{1\leq s\leq T} \frac{\|W_d(s) - sW_d(1))\|}{s\, b(s)} > 1 \Big\},$$

where W_d *is a* d*-dimensional standard Brownian motion.*

In the case of using the most "natural" boundary function $b(\cdot) = c$ with $c > 0$, the following corollary is a direct application of Theorem 3.1.

Corollary 3.1. *Assume* $b(t) = c > 0$ *for all* $t \geq 0$. *Under the assumptions of Theorem 3.1, and with* $T \in (1,\infty)$,

$$\lim_{n\to\infty} P\{\tau(n) < \infty\} = P\{U_{d,T} > c\}$$

where

$$U_{d,T} = \sqrt{(T-1)/T} \sup_{0\leq s\leq 1} \|W_d(s)\| \ \ \textit{when} \ \ 1 < T < \infty \tag{11}$$

and

$$U_{d,\infty} = \sup_{0\leq s\leq 1} \|W_d(s)\|. \tag{12}$$

So that, at a nominal level $\alpha \in (0,1)$, the critical value of interest is $c = c_\alpha$ the $(1-\alpha)$-quantile of the distribution of $U_{d,T}$. Note that, the distribution of $\sup_{0<s\leq 1} \|W_d(s)\|$ is known see for instance Berkes et al. (2004). The quantiles of this distribution can also be computed through a

Monte Carlo simulation ; see the values display in Table 1 of Na et al. (2011) for $d = 1, 2, \cdots, 5$.

Under the alternative, one would like to show that

$$\sup_{n<k\leq[Tn]+1} \frac{\widehat{D}_k}{b(k/n)} \xrightarrow[n\to+\infty]{P} \infty \qquad (13)$$

which is enough for the consistency under H_1. But this seems to be not easy, since the convergence of $\widehat{\theta}(T_k)$ is not ensured under H_1. But we believe that (13) can be obtained with an adapted version of the procedure proposed by Bardet and Kengne (2014). This problem is the topic of a different research project. Nevertheless, the following simulation study shows that the procedure based on $\widehat{D}_k$ still works well under the alternative.

4. Some numerical results

We present some numerical results for sequential change-point detection in Poisson autoregressive models. These results are based on the application of the INGARCH$(1,1)$ model that satisfying

$$Y_t/\mathscr{F}_{t-1} \sim \text{Poisson}(\lambda_t) \text{ with } \lambda_t = \alpha_0 + \alpha_1\lambda_{t-1} + \beta_1 Y_{t-1} \qquad (14)$$

where $\alpha_0 > 0$ and $\alpha_1, \beta_1 \geq 0$. Denote $\theta^* = (\alpha_0, \alpha_1, \beta_1)$ the parameter of the model. As we mentioned above, this model is known to describe adequately the number of transactions in the stock Ericsson B on that we will focus in the real data example. In the sequel, we take $b(s) \equiv c$ is constant and deal with the closed-end procedure with $T = 2$; hence the procedure is monitored from $k = n+1$ to $k = 2n$. The nominal level used is $\alpha = 0.05$. According to Corollary 3.1, the critical value of the test satisfies $c_\alpha = \sqrt{(T-1)/T}\, c'_\alpha$ where c'_α is the $(1-\alpha)$-quantile of the distribution of $\sup_{0<s\leq 1}\|W_d(s)\|$ ($d = 3$ in this section). From Table 1 of Na et al. (2011), we get $c_\alpha = 2.130$.

In the sequel, the historical data $X_1, \cdots, X_n$ are generated by the model (14) depending on the parameter θ_0^*. Under the null hypothesis, the new observations $X_{n+1}, \cdots, X_{[Tn]}$ are also generated according to θ_0^*. Under the alternative, the new observations $X_{n+1}, \cdots, X_{k^*}$ are derived from the model that depends on θ_0^* while $X_{k^*+1}, \cdots, X_{[Tn]}$ depend on θ_1^*.

4.1. An illustration

We consider the above INGARCH$(1,1)$ model and generate the historical data of length $n = 500$ according to the parameter $\theta_0^* = (0.5, 0.7, 0.15)$. Figure 1(a) and (b) displays the statistics $(\widehat{D}_k)_{501\leq k\leq 1000}$ for a scenario without change and a scenario with break at $k^* = 1.25n = 625$. Figure 1 a-) shows that the detector $\widehat{D}_k$ is under the horizontal line which represents the critical value of the test. In Figure 1 b-), before change occurs, $\widehat{D}_k$ is under the horizontal line and increases with a high speed after change. Such growth over a long period indicates that something happening in the model.

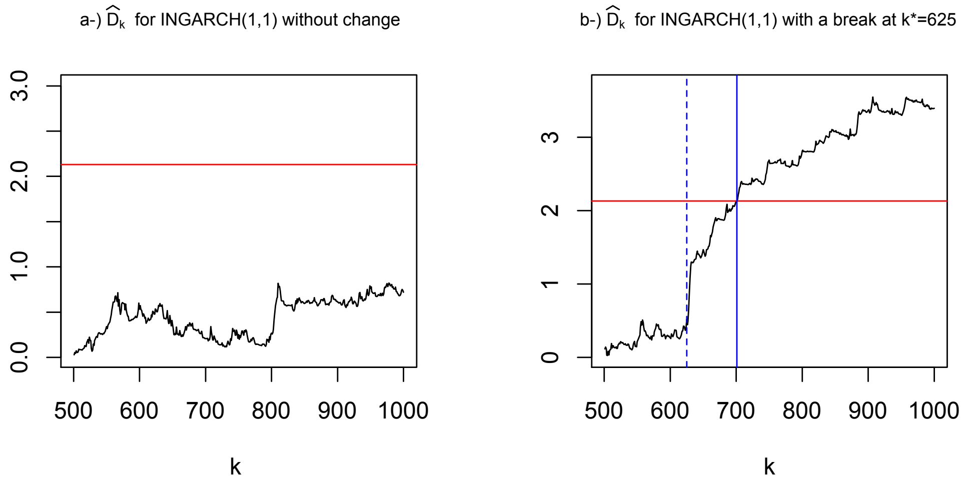

FIGURE 1. *Typical realization of the statistics* $(\widehat{D}_k)_{501\leq k\leq 1000}$ *for INGARCH*$(1,1)$ *with* $n=500$. *a-) The parameter* $\theta_0^*=(0.5,0.7,0.15)$ *is constant; b-) the parameter* $\theta_0^*=(0.5,0.7,0.15)$ *changes to* $\theta_1^*=(0.5,0.4,0.3)$ *at* $k^*=625$. *The horizontal solid line represents the limit of the critical region, the vertical dotted line indicates where the change occurs and the vertical solid line indicates the time where the sequential procedure points a presence of break in the data.*

4.2. Sequential change-points detection in INGARCH$(1,1)$ *processes*

For the problem of the sequential change-points detection in the INGARCH$(1,1)$ model (14), we consider the following scenarios.

- **Scenario A** : under H_0, $\theta_0^*=(0.5,0.7,0.15)$. Under H_1, θ_0^* changes to $\theta_1^*=(0.5,0.4,0.3)$ at $k^*=1.25n$.
- **Scenario B** : under H_0, $\theta_0^*=(1,0.2,0.3)$. Under H_1, θ_0^* changes to $\theta_1^*=(0.4,0.2,0.3)$ at $k^*=1.25n$.

Note that, the scenario A is close to the fitted model obtained from the number of transactions per minute for the stock Ericsson B, which is known to exhibit $\alpha_1+\beta_1\approx 0.9$ (see for instance Fokianos et al., 2009).

Table 1 reports the empirical levels and powers based on 200 replications for $n=150,300,500$. Some elementary statistics of the empirical detection delay (see (10)) are summarized in Table 2.

TABLE 1. *Empirical levels and powers for sequential change-point detection in INGARCH*$(1,1)$ *processes according to the scenarios A and B.*

		$n=150$	$n=300$	$n=500$
Empirical levels :	Scenario A	0.090	0.085	0.060
	Scenario B	0.085	0.075	0.055
Empirical powers :	Scenario A	0.650	0.840	0.980
	Scenario B	0.630	0.945	0.995

TABLE 2. *Elementary statistics of the empirical detection delay for sequential change-point detection in INGARCH*$(1,1)$ *processes according to the scenarios A and B.*

$\widehat{d_n}$		Mean	SD	Min	Q_1	Med	Q_3	Max
Scenario A	$n = 150$; $k^* = 187$	53.96	27.25	4	34	55	78	109
	$n = 300$; $k^* = 375$	98.30	48.57	11	59	95	133	218
	$n = 500$; $k^* = 625$	102.90	53.67	13	65	94	126	351
Scenario B	$n = 150$; $k^* = 187$	65.77	21.64	17	49	68	84	112
	$n = 300$; $k^* = 375$	90.13	33.33	19	65	87	109	215
	$n = 500$; $k^* = 625$	94.60	29.03	32	74	91	112	206

The results of Table 1 show some distortion in the empirical levels. But it decreases as n increases for approaching the nominal level according to Corollary 3.1. The empirical powers increases with n and approaching one when $n = 500$ for both scenarios A and B. Even if the asymptotic power one has not yet been proved, these results show the efficiency of the procedure to detect change-points under the alternative.
Moreover, recall that the detection delay $\widehat{d_n}$ is the random distance between the break time and the stopping time of the procedure. For example, when $n = 150$ with the break occurred at the time $k^* = 187$; from Table 2, this break is detected on average after a delay of 54 for the scenario A. Note that, even if it appears to be symmetric (from Table 2 and some other simulations not reported here) the distribution of $\widehat{d_n}$ is unknown. As pointed out by Bardet and Kengne (2014), such procedure (that take into account all the historical data in the updated parameter estimate) leads to a detection delay which increases with n ; but one can see that the standard deviation is stabilized for $n \geq 300$. According to the dependence structure of the model and the numerical difficulties to compute the estimate of the parameter (a minimum sample size of 50 is needed to expect the convergence of the numerical algorithm), the results of Table 2 are quite satisfactory.

4.3. *Real data example*

We consider the number of transactions per minute for the stock Ericsson B during July 16, 2002. There are 460 observations which represent trading from 09 : 35 to 17 : 14. These data are displayed in Figure 2. Note that, the number of transactions per minute during July 2, 2002 has been studied by Fokianos et al. (2009, 2013) and are known to be described adequately by the INGARCH$(1,1)$ model.
The INGARCH$(1,1)$ model describes the first 130 observations(from 09 : 35 to 11 : 45) adequately according to the goodness-of-fit test proposed by Fokianos et al. (2013). Therefore, these observations are considered as the historical data and we assume that the parameter of the model does not change during this period. Hence, the monitoring starts at the time $t = 131$ and the detector $\widehat{D}_k$ is computed at each time where a new observation is supposed to be available. $\widehat{D}_k$ is displayed in Figure 3.
Figure 3 shows that the procedure works well for this real data example, according to the detection delay which is reasonably good. Note that, in practice, once the sequential procedure stops and indicates a break in the observations, a retrospective procedure has to be applied to estimate the

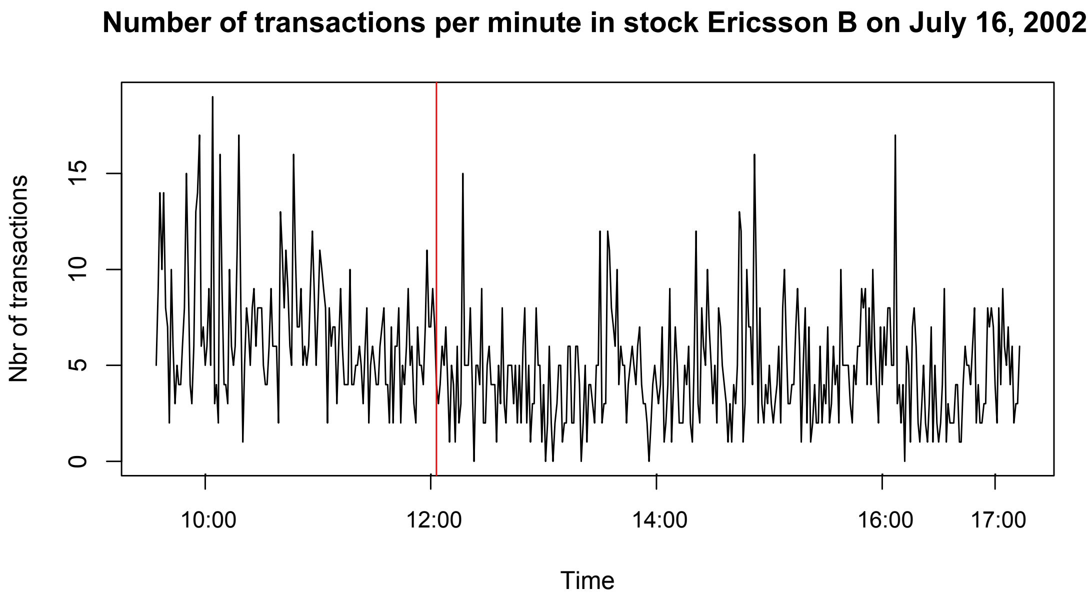

FIGURE 2. *Number of transactions per minute for the stock Ericsson B during July 16, 2002. The vertical line represents the break detected using the retrospective procedure proposed by Doukhan and Kengne (2015).*

real time of change. To end this subsection, let us point out that the sequential procedure could be a good tool for off-line multiple change-points detection, see Bardet and Kengne (2014) for an example of causal time series.

5. Proofs of the main results

In the sequel, C denotes a positive constant which the value may differ from one inequality to another.
Under H_0, the asymptotic covariance matrix of $\widehat{\theta}(T_n)$ is Σ^{-1}. For any $k > n$, denote

$$D_k := \sqrt{n}\left\|\Sigma^{1/2}\left(\widehat{\theta}(T_k) - \widehat{\theta}(T_n)\right)\right\|.$$

The following lemma will be useful.

Lemma 5.1. *Let $T > 1$ (can be equal to ∞). Under the assumptions of Theorem 3.1,*

$$\sup_{n<k<[Tn]+1} \frac{1}{b(k/n)} \|\widehat{D}_k - D_k\| = o_P(1) \ \ as\ n \to \infty.$$

Proof. From Doukhan and Kengne (2015), it hold that $\widehat{\Sigma}_n \xrightarrow[n\to\infty]{a.s.} \Sigma$, $\|\widehat{\theta}(T_n) - \theta_0^*\| = O_P(1/\sqrt{n})$

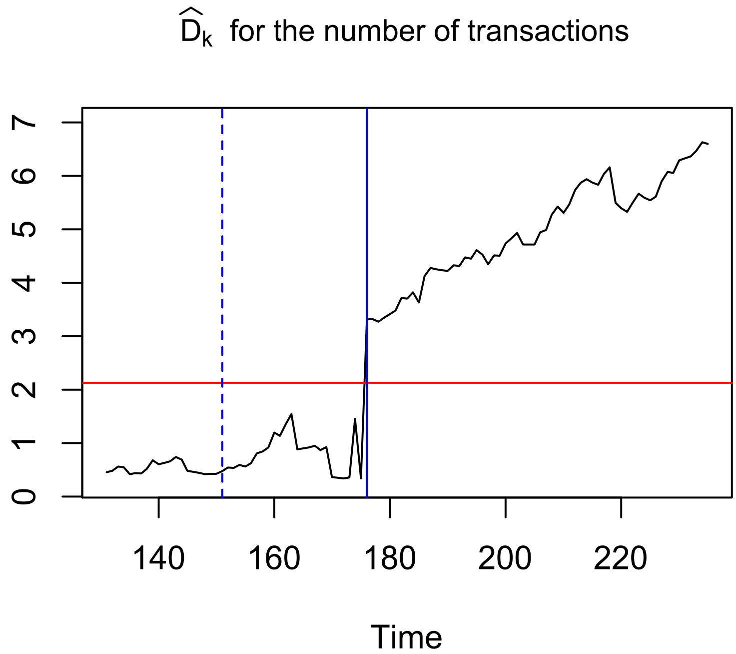

FIGURE 3. *Realizations of the statistics $(\widehat{D}_k)_{131\leq k\leq 235}$ for number of transactions per minute in the stock Ericsson B during July 16, 2002 ; the historical data considered are the first 130 observations. The horizontal solid line represents the limit of the critical region, the vertical dotted line indicates the break that has been detected using the retrospective procedure of Doukhan and Kengne (2015) and the vertical solid line indicates the stopping time of the sequential procedure.*

and for $k > n$, $\|\widehat{\theta}(T_k) - \theta_0^*\| = O_P(1/\sqrt{k})$ as $n \to \infty$. Hence

$$\sup_{n<k<[Tn]+1} \frac{1}{b(k/n)} \|\widehat{D}_k - D_k\| = \frac{1}{\underset{1\leq s\leq T}{\mathrm{Inf}}\, b(s)} \sup_{n<k<[Tn]+1} \sqrt{n}\Big\|\Big(\widehat{\Sigma}_n^{1/2} - \Sigma^{1/2}\Big)\Big(\widehat{\theta}(T_k) - \widehat{\theta}(T_n)\Big)\Big\|$$
$$\leq C \sup_{n<k<[Tn]+1} \sqrt{n}\|\widehat{\Sigma}_n^{1/2} - \Sigma^{1/2}\|\|\widehat{\theta}(T_k) - \theta_0^*\|$$
$$+ C\sqrt{n}\|\widehat{\Sigma}_n^{1/2} - \Sigma^{1/2}\|\|\widehat{\theta}(T_n) - \theta_0^*\|$$
$$= o_P(1) + o_P(1) = o_P(1).$$

■

Proof of Theorem 3.1

The following proof is for the closed-end procedure ($T < \infty$). The result for open-end ($T = \infty$) procedure is established along similar lines.
According to (8) and Lemma 5.1, it suffices to show that

$$\sup_{n<k\leq[Tn]+1} \frac{D_k}{b(k/n)} \xrightarrow[n\to+\infty]{\mathscr{D}} \sup_{1\leq s\leq T} \frac{\|W_d(s) - sW_d(1))\|}{s\, b(s)}.$$

For $k > n$, from the proof of Theorem 4.1 of Doukhan and Kengne (2015), it hold that

$$\Sigma(\widehat{\theta}(T_k) - \theta_0^*) = \frac{1}{k}\frac{\partial}{\partial\theta}L(T_k, \theta_0^*) + o_P(\frac{1}{\sqrt{k}}) \ \text{ and } \ \Sigma(\widehat{\theta}(T_n) - \theta_0^*) = \frac{1}{n}\frac{\partial}{\partial\theta}L(T_n, \theta_0^*) + o_P(\frac{1}{\sqrt{n}}).$$

Hence,

$$\Sigma(\widehat{\theta}(T_k) - \widehat{\theta}(T_n)) = \frac{1}{k}\Big(\frac{\partial}{\partial\theta}L(T_k, \theta_0^*) - \frac{k}{n}\frac{\partial}{\partial\theta}L(T_n, \theta_0^*)\Big) + o_P(\frac{1}{\sqrt{n}}).$$

The latter inequality holds uniformly in $k > n$ by using $o_P(\frac{1}{\sqrt{k}}) + o_P(\frac{1}{\sqrt{n}}) = o_P(\frac{1}{\sqrt{n}})$. Therefore,

$$\begin{aligned}\sqrt{n}\Sigma^{1/2}(\widehat{\theta}(T_k) - \widehat{\theta}(T_n)) &= \frac{n}{k}\frac{1}{\sqrt{n}}\Sigma^{-1/2}\Big(\frac{\partial}{\partial\theta}L(T_k, \theta_0^*) - \frac{k}{n}\frac{\partial}{\partial\theta}L(T_n, \theta_0^*)\Big) + o_P(1)\\ &= \frac{n}{k}\frac{1}{\sqrt{n}}\Sigma^{-1/2}\Big(\sum_{t=1}^{k}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*) - \frac{k}{n}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*)\Big) + o_P(1). \qquad (15)\end{aligned}$$

According to Doukhan and Kengne (2015), $\big(\ell_t(\theta_0^*), \mathscr{F}_t\big)_{t\in\mathbb{Z}}$ is a stationary ergodic square integrable martingale difference sequence with covariance matrix Σ. Hence, from Cramér-Wold device (cf. Billingsley, 1968), it holds that, for any $T > 1$,

$$\frac{1}{\sqrt{n}}\sum_{t=1}^{[ns]}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*) \xrightarrow[n\to+\infty]{\mathscr{D}[1,T]} W_{d,\Sigma}$$

where $W_{d,\Sigma}$ is a zero mean d-dimensional Gaussian process satisfying $E\big(W_{d,\Sigma}(s)'W_{d,\Sigma}(\tau)\big) = \min(s,\tau)\Sigma$. Hence

$$\frac{n}{[ns]}\frac{1}{\sqrt{n}}\Sigma^{-1/2}\Big(\sum_{t=1}^{[ns]}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*) - \frac{[ns]}{n}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*)\Big) \xrightarrow[n\to+\infty]{\mathscr{D}[1,T]} \frac{W_d(s) - sW_d(1)}{s} \qquad (16)$$

where W_d is a d-dimensional standard Brownian motion. Thus, from (15) and (16), it follows that

$$\begin{aligned}\sup_{n<k\le[Tn]+1}\frac{D_k}{b(k/n)} &= \sup_{n<k\le[Tn]+1}\frac{1}{b(k/n)}\frac{n}{k}\Big\|\frac{1}{\sqrt{n}}\Sigma^{-1/2}\Big(\sum_{t=1}^{k}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*) - \frac{k}{n}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*)\Big)\Big\| + o_P(1)\\ &= \sup_{1<s\le T}\frac{1}{b([ns]/n)}\frac{n}{[ns]}\Big\|\frac{1}{\sqrt{n}}\Sigma^{-1/2}\Big(\sum_{t=1}^{[ns]}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*) - \frac{[ns]}{n}\sum_{t=1}^{n}\frac{\partial}{\partial\theta}\ell_t(\theta_0^*)\Big)\Big\| + o_P(1)\\ &\xrightarrow[n\to+\infty]{\mathscr{D}} \sup_{1\le s\le T}\frac{\|W_d(s) - sW_d(1)\|}{sb(s)}.\end{aligned}$$

■

Proof of Corollary 3.1

The following proof is for the closed-end procedure ($T < \infty$). The result for open-end ($T = \infty$) procedure is established along similar lines.

If $b(s) \equiv c$ is constant, then from Theorem 3.1 it follows that

$$\lim_{n\to\infty} P\{\tau(n) < \infty\} = P\Big\{\sup_{1\le s\le T}\frac{\|W_d(s) - sW_d(1))\|}{s} > c\Big\}.$$

Hence, it suffices to show that

$$\sup_{1\le s\le T} \frac{\|W_d(s) - sW_d(1))\|}{s} \overset{\mathscr{D}}{=} \sqrt{(T-1)/T} \sup_{0\le s\le 1} \|W_d(s)\|.$$

By computing the covariance matrix, one can verify that

$$\Big\{\frac{W_d(s) - sW_d(1))}{s},\ s \ge 1\Big\} \overset{\mathscr{D}}{=} \Big\{W_d(\frac{s-1}{s}),\ s\ge 1\Big\}.$$

Thus,

$$\begin{aligned}\sup_{1\le s\le T} \frac{\|W_d(s) - sW_d(1))\|}{s} &\overset{\mathscr{D}}{=} \sup_{1\le s\le T} \Big\|W_d(\frac{s-1}{s})\Big\| \\ &= \sup_{0\le \tau\le 1} \Big\|W_d\Big(\frac{T-1}{T}\tau\Big)\Big\| \\ &\overset{\mathscr{D}}{=} \sup_{0\le \tau\le 1} \sqrt{(T-1)/T}\|W_d(\tau)\|.\end{aligned}$$

■

References

Aue, A., Hörmann, S., Horváth, L., Reimherr, M., et al. (2009a). Break detection in the covariance structure of multivariate time series models. *The Annals of Statistics*, 37(6B):4046–4087.

Aue, A., Horváth, L., Hušková, M., and Kokoszka, P. (2006). Change-point monitoring in linear models. *The Econometrics Journal*, 9(3):373–403.

Aue, A., Horváth, L., and Reimherr, M. L. (2009b). Delay times of sequential procedures for multiple time series regression models. *Journal of Econometrics*, 149(2):174–190.

Bardet, J.-M. and Kengne, W. (2014). Monitoring procedure for parameter change in causal time series. *Journal of Multivariate Analysis*, 125:204–221.

Bardet, J.-M., Kengne, W., Wintenberger, O., et al. (2012). Multiple breaks detection in general causal time series using penalized quasi-likelihood. *Electronic Journal of Statistics*, 6:435–477.

Basseville, M., Nikiforov, I. V., et al. (1993). *Detection of abrupt changes: theory and application*, volume 104. Prentice Hall Englewood Cliffs.

Berkes, I., Gombay, E., Horváth, L., and Kokoszka, P. (2004). Sequential change-point detection in garch (p, q) models. *Econometric Theory*, 20(06):1140–1167.

Billingsley, P. (1968). Convergence of probability measures.

Chu, C.-S. J., Stinchcombe, M., and White, H. (1996). Monitoring structural change. *Econometrica: Journal of the Econometric Society*, pages 1045–1065.

Csörgö, M. and Horváth, L. (1997). *Limit theorems in change-point analysis*. Wiley New York.

Doukhan, P., Fokianos, K., and Tjøstheim, D. (2012). On weak dependence conditions for poisson autoregressions. *Statistics & Probability Letters*, 82(5):942–948.

Doukhan, P. and Kengne, W. (2015). Inference and testing for structural change in general poisson autoregressive models. *Electronic Journal of Statistics*, 9:1267–1314.

Doukhan, P. and Wintenberger, O. (2008). Weakly dependent chains with infinite memory. *Stochastic Processes and their Applications*, 118(11):1997–2013.

Ferland, R., Latour, A., and Oraichi, D. (2006). Integer-valued garch process. *Journal of Time Series Analysis*, 27(6):923–942.

Fokianos, K. and Fried, R. (2010). Interventions in ingarch processes. *Journal of Time Series Analysis*, 31(3):210–225.

Fokianos, K. and Fried, R. (2012). Interventions in log-linear poisson autoregression. *Statistical Modelling*, 12(4):299–322.

Fokianos, K., Gombay, E., and Hussein, A. (2014). Retrospective change detection for binary time series models. *Journal of Statistical Planning and Inference*, 145:102–112.

Fokianos, K., Neumann, M. H., et al. (2013). A goodness-of-fit test for poisson count processes. *Electronic Journal of Statistics*, 7:793–819.

Fokianos, K., Rahbek, A., and Tjøstheim, D. (2009). Poisson autoregression. *Journal of the American Statistical Association*, 104(488):1430–1439.

Fokianos, K. and Tjøstheim, D. (2012). Nonlinear poisson autoregression. *Annals of the Institute of Statistical Mathematics*, 64(6):1205–1225.

Fryzlewicz, P. and Subba Rao, S. (2013). Multiple-change-point detection for auto-regressive conditional heteroscedastic processes. *Journal of the Royal Statistical Society: series B (statistical methodology).*

Gombay, E. and Serban, D. (2009). Monitoring parameter change in time series models. *Journal of Multivariate Analysis*, 100(4):715–725.

Horváth, L., Hušková, M., Kokoszka, P., and Steinebach, J. (2004). Monitoring changes in linear models. *Journal of Statistical Planning and Inference*, 126(1):225–251.

Horvath, L., Kokoszka, P., and Steinebach, J. (2007). On sequential detection of parameter changes in linear regression. *Statistics & probability letters*, 77(9):885–895.

Kang, J. and Lee, S. (2009). Parameter change test for random coefficient integer-valued autoregressive processes with application to polio data analysis. *Journal of Time Series Analysis*, 30(2):239–258.

Kang, J. and Lee, S. (2014). Parameter change test for poisson autoregressive models. *Scandinavian Journal of Statistics*, 41(4):1136–1152.

Kengne, W. C. (2012). Testing for parameter constancy in general causal time-series models. *Journal of Time Series Analysis*, 33(3):503–518.

Leisch, F., Hornik, K., and Kuan, C.-M. (2000). Monitoring structural changes with the generalized fluctuation test. *Econometric Theory*, 16(06):835–854.

Na, O., Lee, Y., and Lee, S. (2011). Monitoring parameter change in time series models. *Statistical Methods & Applications*, 20(2):171–199.

Neumann, M. H. et al. (2011). Absolute regularity and ergodicity of poisson count processes. *Bernoulli*, 17(4):1268–1284.

Weiß, C. H. (2009). Modelling time series of counts with overdispersion. *Statistical Methods and Applications*, 18(4):507–519.

Zeileis, A., Leisch, F., Kleiber, C., and Hornik, K. (2005). Monitoring structural change in dynamic econometric models. *Journal of Applied Econometrics*, 20(1):99–121.

Zhu, F. and Wang, D. (2010). Diagnostic checking integer-valued arch (p) models using conditional residual autocorrelations. *Computational Statistics & Data Analysis*, 54(2):496–508.