Technical Report

11/2009

Interventions in ingarch processes

Konstantinos Fokianos, Roland Fried

Technische Universität Dortmund

SFB 475
Komplexitätsreduktion in multivariaten Datenstrukturen

Sonderforschungsbereich 475
Universität Dortmund
44221 Dortmund
Germany

INTERVENTIONS IN INGARCH PROCESSES

BY KONSTANTINOS FOKIANOS AND ROLAND FRIED

University of Cyprus and TU Dortmund University

June 12, 2009

Abstract

We study the problem of intervention effects generating various types of outliers in a linear count time series model. This model belongs to the class of observation driven models and extends the class of Gaussian linear time series models within the exponential family framework. Studies about effects of covariates and interventions for count time series models have largely fallen behind due to the fact that the underlying process, whose behavior determines the dynamics of the observed process, is not observed. We suggest a computationally feasible approach to these problems, focusing especially on the detection and estimation of sudden shifts and outliers. To identify successfully such unusual events we employ the maximum of score tests, whose critical values in finite samples are determined by parametric bootstrap. The usefulness of the proposed methods is illustrated using simulated and real data examples.

Keywords. parametric bootstrap; generalized linear models; observation driven models; level shifts; transient shifts; outliers.

1. INTRODUCTION

This paper investigates the problem of modeling unusual events in integer-valued GARCH (INGARCH) models, introduced by Ferland et al. (2006) and studied further by Fokianos et al. (2008). An advantage of these models for modeling count time series is the inclusion of a feedback mechanism which yields parsimony–a similar idea to the GARCH model; Bollerslev (1986). In addition, stationarity and geometric ergodicity is guaranteed by simple conditions on the parameters and conditional maximum likelihood model fitting is implemented in a straightforward manner. We will show that techniques for estimation and detection of different types of changes (intervention effects) can be developed within the framework of INGARCH models. In the context of Gaussian linear time series, these questions have been investigated by several authors including Fox (1972), Box and Tiao (1975), Tsay (1986), Chang et al. (1988), Chen and Liu (1993) and Justel, Peña and Tsay (2001), among others. However, to the best of our knowledge, such studies for integer valued dependent data are missing, although their development is important for inference and diagnostics.

Figure 1 motivates much of the subsequent discussion. It shows the number of cases of campylobacterosis infections from January 1990 to the end of October 2000 in the north of the Province of Québec, Canada. These data were recorded every 28 days for a total number of 13 times per year. Ferland et al. (2006) model the time-varying level and variability of these count data by an INGARCH process model. Apparently, the plot illustrates that both the variation and the level of the data increase at the end of the time series. Additionally, there are two possibly outlying values about the time point $t = 100$. It is natural to ask whether these fluctuations can be explained by the INGARCH model or whether the model fit can be improved substantially by the inclusion of such singular effects. In the latter case it can be examined whether the extra variability can be explained by singular real phenomena. the possible causes of the extra variability can be examined. Hence, a method which allows detection of interventions and estimation of their size is needed so that structural changes can be identified successfully. Important steps to achieve this goal are the following, see Chen and Liu (1993):

1. A suitable model for accommodating interventions in count time series data,
2. derivation of test procedures for their successful detection,
3. implementation of joint maximum likelihood estimation of model parameters and outlier sizes
4. and correction of the observed series for the detected interventions.

We address all these issues and give possible directions for further developments of the methodology.

Models for time series of counts have been considered by several authors–see MacDonald and Zucchini (1997) and Kedem and Fokianos (2002, Ch. 4), for instance. Denote by $\{Y_t, t = 1, 2, \dots, n\}$ a count time series. It is usually assumed that the response Y_t, given the information up to time t, is conditionally Poisson distributed with mean λ_t and many of the existing modeling approaches are based upon regressing $\log \lambda_t$– the so-called canonical link parameter–on past values of the response and/or covariates. These models are called observation driven models following Cox (1981). Fokianos and Kedem (2004) show that these models fall within the broad class of time series following generalized linear models and their analysis is based on likelihood inference, see also Zeger and Qaqish (1988), Li (1994), Davis et al. (2003) and Jung et al. (2006), among other authors. However, methods for outlier identification in count time series are missing and the notion

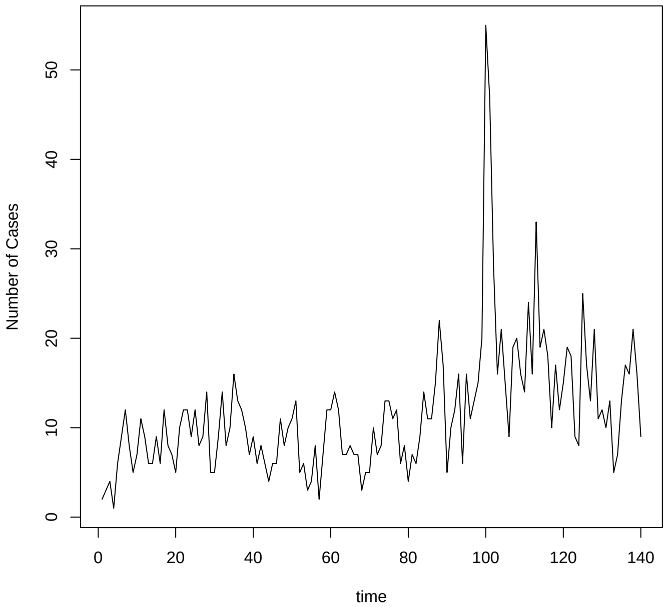


Figure 1: Number of cases of campylobacterosis infections from January 1990 to the end of October 2000.

of intervention (Box and Tiao (1975))) has been only vaguely addressed, see Kedem and Fokianos (2002, Sec. 4.5.2), for instance. A fundamental problem with these models is that the observed process is governed by an underlying hidden process which causes outlier (or any other) modeling to be challenging. We argue that it is reasonable to introduce intervention effects by means of the unobserved (hidden) process.

To develop the theory, we follow Ferland et al. (2006) and focus on the integer-valued GARCH class of models, though our approach can be generalized to other settings. An integer-valued GARCH process $\{Y_t\}$ of orders p and q, abbreviated by INGARCH(p, q), is defined through the following relationships

$$\begin{aligned} Y_t|\mathcal{F}_{t-1}^Y &\sim \text{Poisson}(\lambda_t), \\ \lambda_t &= \beta_0 + \sum_{i=1}^{q} \beta_i Y_{t-i} + \sum_{j=1}^{p} \alpha_j \lambda_{t-j}, \end{aligned} \tag{1}$$

for $t \geq 1$ and an intercept $\beta_0 > 0$, regression parameters $\beta_i > 0$, $i = 1, \ldots, q$, and $\alpha_j > 0$, $j = 1, \ldots, p$. The dynamics of the process are modeled via the conditional mean $\lambda_t = \mathrm{E}(Y_t|\mathcal{F}_{t-1}^Y)$ of Y_t, which is a function of the whole information $\mathcal{F}_{t-1}^Y$ up to time $t-1$ and of the unknown regression parameters. Here $\mathcal{F}_t^Y$ stands for the σ–field generated by $\{Y_{1-q}, \ldots, Y_t, \lambda_{1-p}, \ldots, \lambda_0\}$. A stationary solution of (1) with mean $\beta_0/(1 - \sum_{i=1}^p \alpha_i - \sum_{j=1}^q \beta_j)$ exists provided that $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$. Similar models, in which $\{\lambda_t\}$ is regressed on past values of the observed process and past values of $\{\lambda_t\}$ itself, have been studied before by Rydberg and Shephard (2000) and Streett (2000) and more recently by Fokianos et al. (2008). Apparently model (1) has a close resemblance to the GARCH(p, q)-process popularized by Bollerslev (1986) since the mean of the

Poisson distribution equals its variance. Our focus is on the INGARCH(1,1) model since like in the case of GARCH models it is the simplest interesting variant and yet sufficiently flexible for approximating many realistic dependence structures.

Since INGARCH models are closely related to the common GARCH models it is natural to consider intervention modelling within the GARCH framework. The effects of outliers within the context of GARCH models have been investigated by Van Dijk et al. (1999) and Carnero et al. (2006) among other authors. Since a GARCH model can be represented as an ARMA model, it is tempting to introduce different outlier effects directly to the observations $\{Y_t\}$ following the linear time series methodology, see Charles and Darnè (2005), for instance. Adding a constant to a Gaussian variable yields a shifted random variable which is again Gaussian, but the same is not anymore true for a Poisson distributed random variable. In fact, the resulting random variable is not even integer valued if the shift is not an integer. We resolve these problems by introducing intervention effects through the unobserved process $\{\lambda_t\}$.

In the following, the problem of detection and estimation of intervention effects is discussed within the class of INGARCH$(1,1)$ models. Section 2 proposes definitions for different types of interventions in INGARCH processes. It is argued that it is more sensible to introduce intervention effects through the unobserved process $\{\lambda_t\}$. The proposed approach is quite general and can be employed in other settings dealing with integer valued dependent data. Section 3 develops joint estimation of regression parameters and intervention effects within the framework of maximum likelihood. Section 4 suggests score tests for the detection and identification of changes at known time points and investigates their power. Section 5 modifies the approach to detect changes at unknown time points by employing a parametric bootstrap procedure. Section 6 outlines an iterative procedure for detection of multiple interventions and applies it to real and simulated data. The work concludes with some comments on further research in this area.

2. INTERVENTION EFFECTS

In general, different types of intervention effects on time series data are classified according to whether their influence is concentrated on a single or a few data points, or whether they affect the whole process from some specific time $t = \tau$ on. In classical linear time series methodology an intervention effect is included in the observation equation by employing a sequence of deterministic covariates $\{X_t\}$ of the form

$$X_t = \xi(\mathcal{B}) I_t(\tau) \tag{2}$$

for $t \geq 1$, where $\xi(\mathcal{B})$ is a polynomial operator to be defined below, $\mathcal{B}$ is the shift operator such that $\mathcal{B}^i X_t = X_{t-i}$ and $I_t(\tau)$ is an indicator function, with $I_t(\tau) = 1$ if $t = \tau$, and $I_t(\tau) = 0$ if $t \neq \tau$. The choice of the operator $\xi(\mathcal{B})$ determines the kind of intervention effect: additive outlier (A0), transient shift (TS), level shift (LS) and innovational outlier (IO). Since INGARCH models are not defined in terms of innovations, we focus on the first three types of interventions. By setting

$$\xi(\mathcal{B}) = \quad (1 - \delta \mathcal{B})^{-1},\ \delta \in [0, 1], \tag{3}$$

AO and LS type of interventions correspond to $\delta = 0$ and $\delta = 1$, respectively. For a TS, the value of δ is typically chosen as a predefined constant $\delta \in \{0.7, 0.8, 0.9\}$. In the context of linear time series, the aforementioned specification allows for easily interpretable results because the outlier process enters the observation equation as a covariate, see Tsay (1986, eq. 1.2) and Chen and Liu (1993, eq.2).

For models like (1), whose behavior is determined by a latent process, a formal linear structure as in the case of Gaussian linear time series model does not hold any more and interpretation of interventions is a more delicate issue. We argue that detectable and meaningful intervention effects in the context of model (1) can be defined by

$$\begin{aligned} Z_t|\mathcal{F}_{t-1} &\sim \text{Poisson}(\kappa_t), \\ \kappa_t &= \beta_0 + \sum_{i=1}^{q} \beta_i Z_{t-i} + \sum_{j=1}^{p} \alpha_j \kappa_{t-j} + \nu X_t, \end{aligned} \tag{4}$$

for $t \geq 1$, where ν is the size of the intervention effect, $\{X_t\}$ is defined in (2) and $\mathcal{F}_t$ is the σ–field generated by $\{Z_{1-q}, \dots, Z_t, \kappa_{1-p}, \dots, \kappa_0\}$. In other words, the mean process $\{\lambda_t\}$ of (1) is transformed to another mean process, called $\{\kappa_t\}$, by adding a single intervention effect starting at time τ. The idea of introducing such effects via the underlying hidden process is quite general and can be employed in the context of other models for time series of counts, like log–linear autoregressive models. However, there is need for some model dependent adjustments regarding interpretation of interventions.

The main idea that model (4) brings forward is that instead of the "clean" INGARCH process $\{Y_t\}$ we observe the contaminated process $\{Z_t\}$, which includes the effect of an intervention at time τ. In case of $\nu = 0$ and equal starting values for $\{\lambda_t\}$ and $\{\kappa_t\}$ as well as for $\{Y_t\}$ and $\{Z_t\}$, model (4) reduces to (1), because $\{\lambda_t\}$ and $\{\kappa_t\}$ are derived from the same recursions with identical initialization. For $\nu > 0$, the observed process $\{Z_t\}$ can be thought as the sum of two independent processes: an intervention free time series $\{Y_t\}$ that follows (1) and a sequence $\{C_t\}$ of Poisson random variables with mean depending on both ν and the choice of $\xi(\mathcal{B})$. In other words

$$Z_t = Y_t + C_t \,. \tag{5}$$

For any value of ν, model (4) implies that at time $t = \tau + h$ the intercept of the regression equation for the process $\{\kappa_t\}$ becomes $\beta_0 + \nu\xi_h$ instead of β_0, where ξ_h is the coefficient of the term $\mathcal{B}^h$ in $\xi(\mathcal{B})$. Restricting $\nu > -\beta_0$ in (4) implies that the contaminated process $\{\kappa_t\}$ does not become negative because (3) shows that none of the coefficients of $\xi(\mathcal{B})$ takes values larger than one. This restriction is not severe since we can only speak of negative (or better downward) changes if β_0 has a large positive value. In this case the mean of the observed process $\{Y_t\}$ stays away from zero. For small values of β_0 we can hardly call any value a downward outlier since every small non negative integer occurs with substantial probability unless $\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j \approx 1$.

A few illuminating calculations for the INGARCH(1,1) model show the consequences of introducing intervention effects by means of the unobserved process. First recall model (1) and notice that by successive substitutions we obtain that

$$\lambda_t = \beta_0 \frac{1-\alpha_1^t}{1-\alpha_1} + \alpha_1^t \lambda_0 + \beta_1 \sum_{i=0}^{t-1} \alpha_1^i Y_{t-(i+1)}.$$

Similarly,

$$\kappa_t = \beta_0 \frac{1-\alpha_1^t}{1-\alpha_1} + \alpha_1^t \lambda_0 + \beta_1 \sum_{i=0}^{t-1} \alpha_1^i Z_{t-(i+1)} + \nu \sum_{i=0}^{t} \alpha_1^i X_{t-i},$$

or equivalently

$$\kappa_t = \lambda_t + \sum_{i=0}^{t-1} \alpha_1^i C_{t-(i+1)} + \nu \sum_{i=0}^{t} \alpha_1^i X_{t-i}, \tag{6}$$

assuming that $\kappa_0 = \lambda_0$ and using (5). When $\nu = 0$, $\kappa_t = \lambda_t$ for all t. A transient shift, see (3), yields the contaminated latent process

$$\kappa_t = \begin{cases} \lambda_t, & t < \tau, \\ \lambda_t + \sum_{i=0}^{t-1} \alpha_1^i C_{t-(i+1)} + \nu \sum_{j=0}^{t} \alpha_1^j \delta^{t-j-\tau}, & t \geq \tau. \end{cases} \tag{7}$$

where $\delta \in (0,1)$. Note that a transient shift with $\delta \in (0,1)$ implies the presence of a decaying shift whose effect becomes gradually smaller as time grows since $\sum_{j=0}^{t} \alpha_1^j \delta^{t-j-\tau} = \delta^{t-\tau} \dfrac{1-(\alpha_1/\delta)^{t+1}}{1-\alpha_1/\delta} = \delta^{-\tau} \dfrac{\delta^{t+1}-\alpha_1^{t+1}}{\delta-\alpha_1}$. The size of the effect decreases faster for smaller values of δ and α_1. Figure 2(b) illustrates this point. In the case of $\delta = 0$, we obtain that

$$\kappa_t = \begin{cases} \lambda_t, & t < \tau, \\ \lambda_t + \sum_{i=0}^{t-1} \alpha_1^i C_{t-(i+1)} + \nu \alpha_1^{t-\tau}, & t \geq \tau. \end{cases} \tag{8}$$

Figure 2(a) illustrates this situation and it should be noted that this form of intervention effect still influences the process from time τ on, but to a rapidly decaying extent provided that α_1 is not close to unity. Accordingly, we call this a spiky outlier, and abbreviate it by SO. On the other extreme, for $\delta = 1$ a level shift type of intervention is observed, corresponding to a permanent change in the mean (and the variance) of the process; see Figure 2(c). Equation (6) becomes

$$\kappa_t = \begin{cases} \lambda_t, & t < \tau, \\ \lambda_t + \sum_{i=0}^{t-1} \alpha_1^i C_{t-(i+1)} + \nu \dfrac{1-\alpha_1^{t-\tau+1}}{1-\alpha_1}, & t \geq \tau. \end{cases} \tag{9}$$

As a further remark, note that any INGARCH(1,1) model has identical second order properties as the following ARMA(1,1) model

$$Y_t - \mu - (\beta_1 + \alpha_1)(Y_{t-1} - \mu) = e_t - \alpha_1 e_{t-1},$$

where $\{e_t\}$ is a white noise sequence with $\mu = \beta_0/(1-\beta_1-\alpha_1)$ and $\sigma^2 = \mu$, see Ferland et al. (2006, Cor. 2). Hence it is tempting to define outliers by the corresponding ARMA representation, along the lines of Chen and Liu (1993). Notice that these authors introduce outlier effects to the process $\{Y_t\}$ by simple addition, while we modify the model for the underlying (conditional) mean process $\{\lambda_t\}$. In case of a Gaussian process, both approaches are equivalent. Thus, our proposal indeed means an extension of the model of Chen and Liu (1993) to Poisson time series. The other kind of extension, in which the outliers are added directly to the observations, has been considered in the context of GARCH models by Charles and Darnè (2005), among others, who employ the ARMA(1,1) representation of a GARCH(1,1) model to define additive and innovative

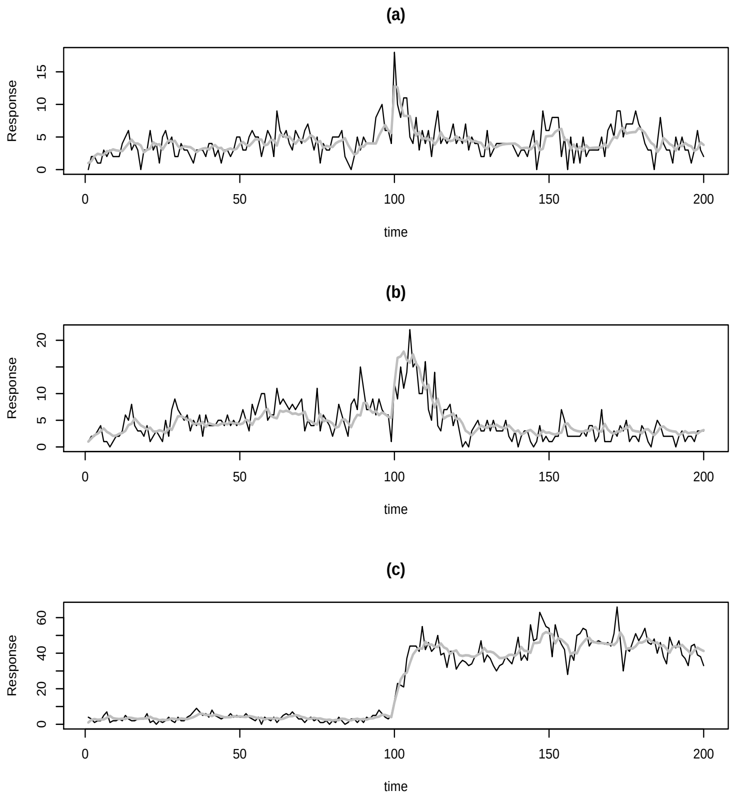


Figure 2: Effects of different types of outliers of size $\nu = 8$ at time point $\tau = 100$ on a realization of an INGARCH(1,1) model generated with $\beta_0 = 1$, $\beta_1 = \alpha_1 = 0.3$ and $n = 200$. The black line shows the observed time series $\{Z_t\}$ and the grey line the underlying mean process $\{\kappa_t\}$. (a) Spiky outlier. (b) Transient shift with $\delta = 0.8$. (c) Level shift.

outliers. Transferring this idea to the INGARCH model means that instead of the outlier free process $\{Y_t\}$, the observations are generated by

$$Z_t = Y_t + \nu X_t.$$

Such a definition leads to some complications for the case of count data. All types of outliers require ν to be an integer. In case of a level shift, the observations $\{Z_t\}$ are no longer Poisson distributed for $t \geq \tau$, but follow a shifted Poisson distribution instead. For the TS case, further modifications are needed for the observations to take integer values. Therefore, it seems more natural to define outliers in count time series via the conditional mean evolution of model (4). The main assumption here is that abrupt changes or interventions directly influence the conditional mean, and the result in turn affects the observed series $\{Z_t\}$. Since the INGARCH model is not defined in terms of innovations, we do not distinguish between additive and innovational outliers. Indeed, spiky outliers (the above case of $\delta = 0$) influence the future of the time series according to the dynamics of the process, and hence bear some analogy to the innovational outliers in classical ARMA modeling.

3. MAXIMUM LIKELIHOOD ESTIMATION

Maximum likelihood inference for model (1) has been discussed by Ferland et al. (2006) and Fokianos et al. (2008). Along these lines, joint estimation of model parameters and outlier effects can be carried out by viewing (4) as a regression model that includes the time dependent covariate process defined by (2).

Accordingly, the conditional likelihood function of the observed data $z_1, \ldots, z_n$ given $z_0, \ldots, z_{1-q}$ and $\kappa_0, \ldots, \kappa_{1-p}$ for model (4) is given by

$$L(\theta) = \prod_{t=1}^{n} \frac{\kappa_t(\theta)^{z_t} e^{-\kappa_t(\theta)}}{z_t!} ,$$

where $\theta = (\beta_0, \beta_1, \ldots, \beta_q, \alpha_1, \ldots, \alpha_p, \nu)'$ is the vector of unknown model parameters and $\kappa_t(\theta)$ is given by equation (4). Therefore, the log-likelihood function is equal to

$$\ell(\theta) = \sum_{t=1}^{n} (z_t \ln \kappa_t(\theta) - \kappa_t(\theta)) = \sum_{t=1}^{n} \ell_t(\theta),$$

up to a constant. Differentiation shows that the score function is given by the $(p+q+2)$–dimensional vector

$$S_{n\tau}(\theta) = \frac{\partial \ell(\theta)}{\partial \theta} = \sum_{t=1}^{n} \frac{\partial \ell_t(\theta)}{\partial \theta} = \sum_{t=1}^{n} \left(\frac{z_t}{\kappa_t(\theta)} - 1 \right) \frac{\partial \kappa_t(\theta)}{\partial \theta}, \tag{10}$$

where

$$\begin{aligned}
\frac{\partial \kappa_t(\theta)}{\partial \beta_0} &= 1 + \sum_{j=1}^{p} \alpha_j \frac{\partial \kappa_{t-j}(\theta)}{\partial \beta_0} \\
\frac{\partial \kappa_t(\theta)}{\partial \beta_i} &= z_{t-i} + \sum_{j=1}^{p} \alpha_j \frac{\partial \kappa_{t-j}(\theta)}{\partial \beta_i}, \quad i = 1, \ldots, q, \\
\frac{\partial \kappa_t(\theta)}{\partial \alpha_i} &= \kappa_{t-i}(\theta) + \sum_{j=1}^{p} \alpha_j \frac{\partial \kappa_{t-j}(\theta)}{\partial \alpha_i}, \quad i = 1, \ldots, p, \\
\frac{\partial \kappa_t(\theta)}{\partial \nu} &= \sum_{j=1}^{p} \alpha_j \frac{\partial \kappa_{t-j}(\theta)}{\partial \nu} + \xi(\mathcal{B}) I_t(\tau)
\end{aligned}$$

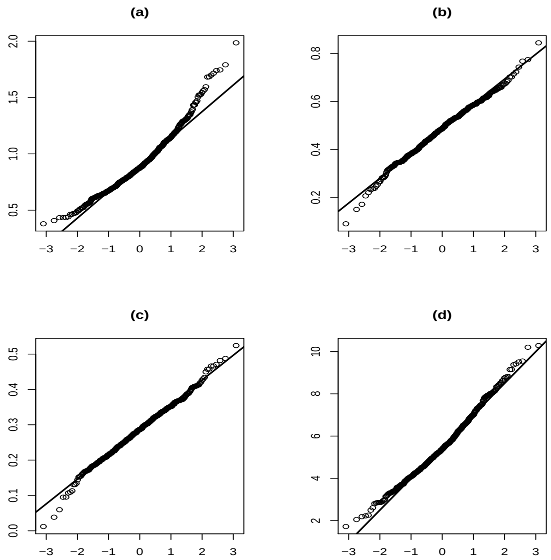
by using (2). The notation $S_{n\tau}(\theta)$ is used to indicate that the score depends upon τ – the time that an intervention has taken effect. The solution of the equations $S_{n\tau}(\theta) = 0$, provided that it exists, yields the conditional maximum likelihood estimate $\hat{\theta}$ of θ. In addition, the conditional information matrix for θ is given by

$$G_{n\tau}(\theta) = \sum_{t=1}^{n} \mathrm{Cov} \left[\frac{\partial \ell_t(\theta)}{\partial \theta} \Big| \mathcal{F}_{t-1} \right] = \sum_{t=1}^{n} \frac{1}{\kappa_t(\theta)} \left(\frac{\partial \kappa_t(\theta)}{\partial \theta} \right) \left(\frac{\partial \kappa_t(\theta)}{\partial \theta} \right)' . \tag{11}$$

Note that we assume the intervention time τ to be known here. The common case in which τ is not known a-priori can be treated by testing for interventions at all time points and choosing the value of τ maximizing the standardized test statistic, see Section 5.

In case of the INGARCH(1,1) model, we obtain that θ is a four dimensional vector of unknown parameters and the score function is given by (10) by modifying accordingly the quantities $\partial \kappa_t / \partial \theta$. Large sample properties of $\hat{\theta}$ are studied by proving joint ergodicity of the process $\{(Z_t, \kappa_t)\}$ and finiteness of its moments. This is the kind of conditions needed to obtain asymptotic normality for the parameter estimates. More specifically, it is required that the score function is asymptotically normally distributed and the Hessian matrix converges in probability to a non random limit. In addition, the third derivatives of the likelihood function have to be uniformly bounded, Fokianos et al. (2008). Figure 3 supports the claim of asymptotic normality of $\hat{\theta}$ if a TS type of outlier is included in model (4). Note that the distribution of the estimator of the intercept is moderately skewed– a phenomenon occurred in the GARCH(1,1) model as well.

Figure 3: QQ-plots of (a) $\hat{\beta}_0$, (b) $\hat{\alpha}_1$, (c) $\hat{\beta}_1$ and (d) $\hat{\nu}$ estimated from a realization of an INGARCH(1,1) model with a TS generated by (4). The parameter values are $\beta_0 = 1$, $\beta_1 = 0.3$, $\alpha_1 = 0.5$, $\nu = 5$ and $n = 200$. The time of intervention is $\tau = 120$ and $\delta = 0.8$, see (3). The plots are based on 500 simulations.



Calculation of the maximum likelihood estimators is carried out by numerical optimization of the log–likelihood function. The optimization is accomplished by employing a quasi–Newton method–the so called BFGS method–implemented in the function `constrOptim` of the `R` statistical language, R Development Core Team (2004). For the basic case of the INGARCH(1,1) model, initial estimates for the optimization are obtained by conditional least squares using the ARMA(1,1) representation.

4. Testing for an intervention effect at a known point in time

We consider testing for the presence of an intervention effect within the INGARCH model in the case that the type and the time of the outlier are known. In the next section we extend our approach to the situation in which both the type and the time of the outlier are unknown. We utilize the score test because its application requires model fitting only under the null hypothesis of no intervention. This allows us to perform individual tests for each type of intervention effect at each time point simultaneously, fitting the model only once.

Alternatively, Wald tests or likelihood ratio tests could also be employed, but their implementation in the situation of unknown type and time of intervention requires separate model fitting for each type of intervention at each time point. Therefore, both Wald type tests and likelihood ratio tests increase the computational burden substantially, especially for long time series.

Consider a certain type of intervention at a specific time point τ and test the hypothesis $\mathrm{H}_0^{(\tau)} : \nu = 0$ against the alternative $\mathrm{H}_1^{(\tau)} : \nu \neq 0$ in model (4). The corresponding score test statistic is given by

$$T_n(\tau) = S_{n\tau}^{'}(\tilde{\beta}_0, \ldots, \tilde{\alpha}_p, 0) G_{n\tau}^{-1}(\tilde{\beta}_0, \ldots, \tilde{\alpha}_p, 0) S_{n\tau}(\tilde{\beta}_0, \ldots, \tilde{\alpha}_p, 0), \tag{12}$$

where $S_{n\tau}(\tilde{\beta}_0, \ldots, \tilde{\alpha}_p, 0)$ and $G_{n\tau}(\tilde{\beta}_0, \ldots, \tilde{\alpha}_p, 0)$ are the score function (10) and the conditional information matrix (11), respectively, evaluated at $(\tilde{\beta}_0, \ldots, \tilde{\alpha}_p, 0)^{'}$–the maximum likelihood estimators computed under model (1). Then we have the following result, which follows directly from Basawa (1991).

Lemma 1 Suppose that model (4) holds and let θ_0 denote the vector of the true parameter values. Assume the following two conditions:

1)

$$\frac{1}{n} G_{n\tau}(\theta_0) \rightarrow G(\theta_0)$$

in probability, as $n \rightarrow \infty$, where $G(\theta_0)$ is a $(p + q + 2) \times (p + q + 2)$ positive definite matrix.

2)

$$\frac{1}{\sqrt{n}} S_{n\tau}(\theta_0) \rightarrow \mathcal{N}_{p+q+2}\left(0, G^{-1}(\theta_0)\right)$$

in distribution, as $n \rightarrow \infty$, where $\mathcal{N}_d$ denotes a d-variate normal distribution.

Then, under the null hypothesis $\mathrm{H}_0^{(\tau)} : \nu = 0$

$$T_n(\tau) \rightarrow \chi_1^2,$$

in distribution, as $n \rightarrow \infty$, where χ_1^2 is the chi-square distribution with one degree of freedom.

The form of the score test (12) depends upon the type of the outlier considered–see (3). For the general INGARCH(p, q) the first two conditions of Lemma 1 are assumed to hold since the region of ergodicity for such processes is still unknown. For the INGARCH(1,1) case, Fokianos et al. (2008) show that the condition $0 < \alpha_1 + \beta_1 < 1$ implies consistency and asymptotic normality of the maximum likelihood estimators.

Recall that under $\mathrm{H}_0^{(\tau)} : \nu = 0$ model (4) reduces to model (1). Therefore we have the following result for the special case of the INGARCH(1,1) model. Its proof is based on the Lemmas 3.1 and 3.2 of Fokianos et al. (2008) and on Basawa (1991). Note that δ is fixed, for example $\delta = 0.8$, when testing for a TS since otherwise this parameter is not identifiable under the null hypothesis.

Lemma 2 If model (4) is true with $0 < \alpha_1 + \beta_1 < 1$, then under the null hypothesis $\mathrm{H}_0^{(\tau)} : \nu = 0$ it holds

$$T_n(\tau) \rightarrow \chi_1^2.$$

Lemma 2 allows derivation of critical values for an asymptotic test of the null hypothesis of no intervention against the specific alternative $\mathrm{H}_1^{(\tau)}$ of an intervention of a given type at a fixed time τ: we reject the null hypothesis at a given significance level α, if the value of $T_n(\tau)$ is larger than the $(1-\alpha)$-quantile of the χ_1^2-distribution. Table 1 shows a few simulation results for examining the adequacy of the chi–square approximation. We compare the output with the corresponding 90%-, 95%- and 99%-percentiles of the chi–square distribution with one degree of freedom. In all cases the results indicate a close agreement between the achieved and nominal significance levels.

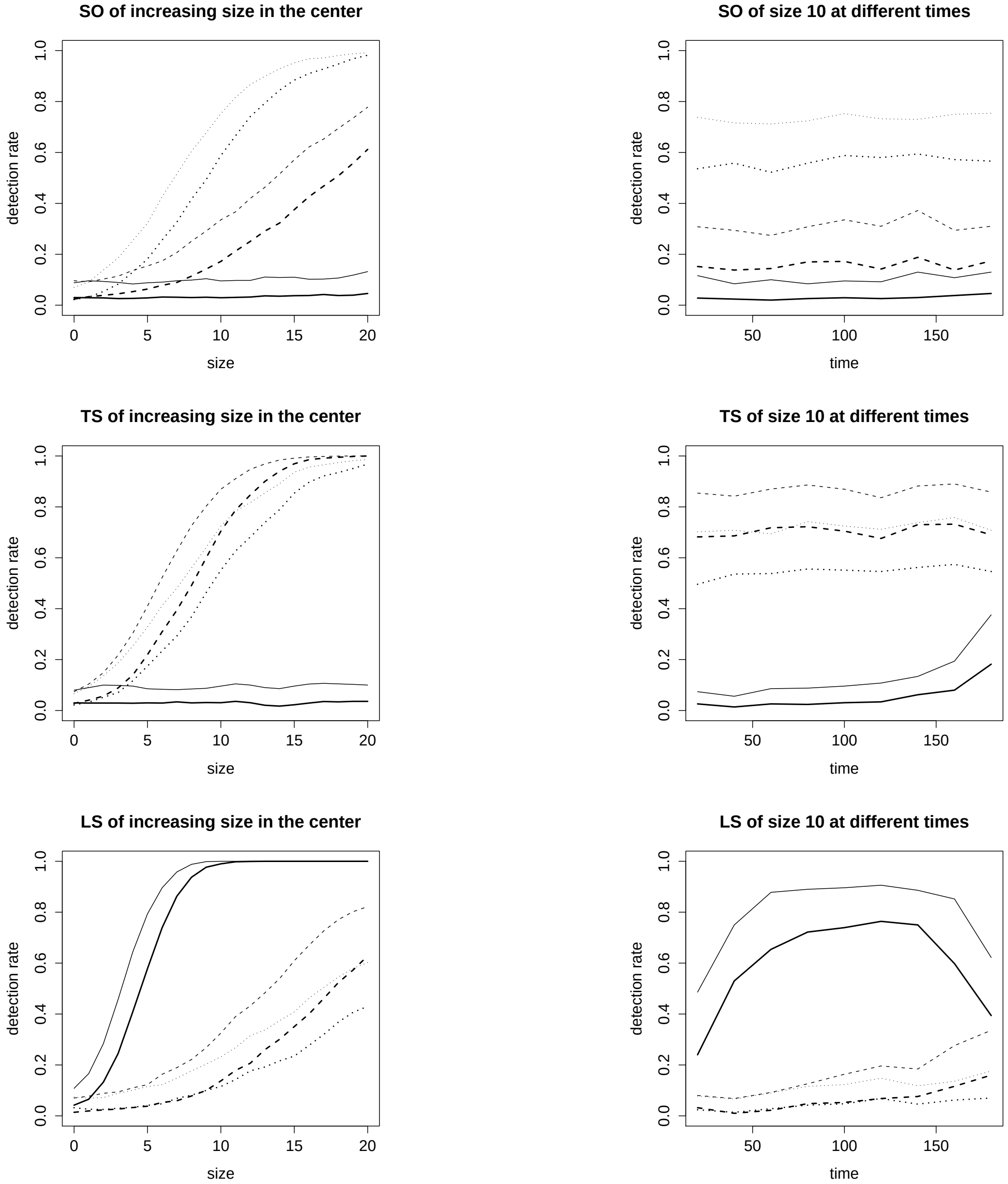
Table 1: Achieved significance levels (in percent) of the score test statistic $T_n(\tau)$ for different sample sizes and different time of interventions. The data have been generated by an INGARCH(1,1) model with $\beta_0 = 0.8$, $\alpha_1 = 0.5$ and $\beta_1 = 0.3$. For the TS type statistic $\delta = 0.8$. The results are based on 1000 simulations each.

	$\tau = 0.25n$			$\tau = 0.50n$			$\tau = 0.75n$			
Type	1	5	10	1	5	10	1	5	10	n
SO	1	5.2	10.2	1	5.6	9.8	1.2	4.9	9	
TS	1.4	5.4	9.9	1.3	4.7	9.6	0.5	4.9	9.6	200
LS	0.7	4.2	8.5	1	3.1	8.2	1.1	3.3	8.1	
SO	0.8	4.9	10.1	0.9	5.2	10.3	0.8	3.6	8.2	
TS	1.4	4.6	9.2	1.1	4.5	8.5	1.4	4.6	8.3	500
LS	0.6	3.4	7	1.1	4.5	9	0.7	3.9	8.3	
SO	0.8	4	8.8	0.6	3.8	8.6	1.3	4.6	9.5	
TS	0.7	4.1	9.1	0.8	5.2	10.7	1.7	6.0	10.6	1000
LS	0.9	3.9	8.5	0.8	5.6	11.3	0.4	4.8	9.4	

Figure 4 illustrates the power of the tests as a function of the size and the time of the intervention, for different types of interventions. It is clear that the power of the tests increases rapidly along with the size. The power of the tests for SO and TS is somewhat affected by the time of occurrence of the outlier. LS types of interventions are more difficult to be detected especially if they occur either at the beginning or at the end of the series. In such a case, we lack information on the time before or after the shift.

We also examine the sensitivity of the tests against misspecification of the type of intervention, because, occasionally, there is information about the time point of a special event, but not about its effect. While it appears that a LS is rarely detected in case of an SO or TS, the test for SO is often significant in case of a TS and vice versa. Nevertheless, the test for the correct type of intervention seems to take typically more significant values so that we can classify the type of the outlier according to the magnitude of the p-values of the tests. Note that an LS at the end of the series can be confused with a TS if we do not know the type of the outlier. In general the tests are somewhat oversized. As a rough guideline we state that testing for any type of outlier at a given time point by performing the tests for the different types at an 1% level of significance each yields roughly to a total 5% level. These results have been confirmed by simulations for other parameter settings.

Figure 4: Detection rates for the different types of interventions as a function of the size (left) and the time (right) of the intervention: SO (top), TS (center) and LS (bottom). Detection rates of the test for SO (dotted), for TS (dashed) and for LS (solid) at 5% (thin) and at 1% significance level (bold line). Each curve is based on 500 simulation runs from an INGARCH(1,1) model with $\beta_0 = 5$, $\beta_1 = \alpha_1 = 0.3$.



5. Testing for an intervention effect at an unknown point in time

To detect an intervention of a certain type at any time point, we maximize the test statistic (12) with respect to τ and reject the hypothesis $\mathrm{H}_0 : \nu = 0$ if

$$\tilde{T}_n = \max_{\tau} T_n(\tau) \tag{13}$$

is large. We performed extensive simulations to derive critical values for the maximum score test statistic $\tilde{T}_n$ under different parameter settings. However, it turned out that the empirical percentiles of $\tilde{T}_n$ varied substantially for different parameter configurations, even in the case of long series consisting of $n = 500$ observations. In other words, suitable critical values depend upon the unknown underlying parameter. A solution to this problem could have been a grid search over the parameter space to derive maximum percentiles as critical values. Besides the large computational costs of this approach, a further drawback would be a substantial loss of power for a broad range of parameter values. Moreover, we found the percentiles to depend on the type of outlier–a fact that can be explained by the different degree of dependence among the test statistics for the individual time points. Therefore, we apply a simple parametric bootstrap for deriving critical values of (13).

Towards this goal, we fit an INGARCH model to the observed time series assuming that there are no outliers. Then we generate a large number of, say, $B = 500$ bootstrap replicates from the fitted INGARCH model with the same parameters as those estimated for the observed real data. The maximum test statistics (13) are calculated both for the real and the bootstrapped data: if the real data do not contain outliers, we expect the corresponding value of $\tilde{T}_n$ to be comparable to those for the bootstrap series.

We present a few exemplary simulation results to check whether the number of bootstrap test statistics which are larger than the value of $\tilde{T}_n$ for the real data can be viewed as a random draw from a discrete uniform distribution on $0, 1, \ldots, B$. 'Real' time series of length $n = 200$ have been generated from an INGARCH(1,1) model with parameters $(\beta_0, \alpha_1, \beta_1) = (5, 0.3, 0.3)$ and the parameters are estimated assuming that there are no outliers. Then we generate $B = 500$ bootstrap replicates from an INGARCH(1,1) with the estimated parameters. The maximum test statistics $\tilde{T}_n$ are calculated for the original and for the B bootstrap series. From this we obtain the number N of bootstrap replicates for which $\tilde{T}_n$ is larger than for the original data. This is repeated 100 times and the resulting counts are depicted in Figure 5. These histograms show adequate approximation to the uniform distribution, and therefore we can transform N to a p-value, dividing it by $B + 1$.

Next we generate time series of length $n = 200$ with a level shift of increasing size $\nu = 0, 1, \ldots, 5$ at time $\tau = 100$ from the same INGARCH(1,1) model as before. Figure 6 depicts boxplots of the number N of bootstrap replicates for which $\tilde{T}_n$ is larger than for the original data, as a function of the shift size. Apparently, the distribution of N is only slightly affected by the shift in case of SO and TS, while it rapidly concentrates almost all of the mass on very small values when testing for an LS. Thus, we expect to obtain good power of the test and also reliable classification results based on a comparison of the p-values for the different types of outliers.

This is confirmed in Figure 7, which illustrates the power of the tests for different types of interventions. The type of the intervention is classified according to the minimal p-value, with preference being given to interventions with larger value of δ in case of equality. The reason for this preference is the tendency of the test for SO to be also significant when testing for either LS or TS, and of the test for TS if there is an LS,

Figure 5: Histograms of the number of bootstrap replicates (out of $B = 500$) for which $\tilde{T}_n$ takes a larger value than for the original data, obtained from 100 repetitions: SO (left), TS (middle) and LS (right).

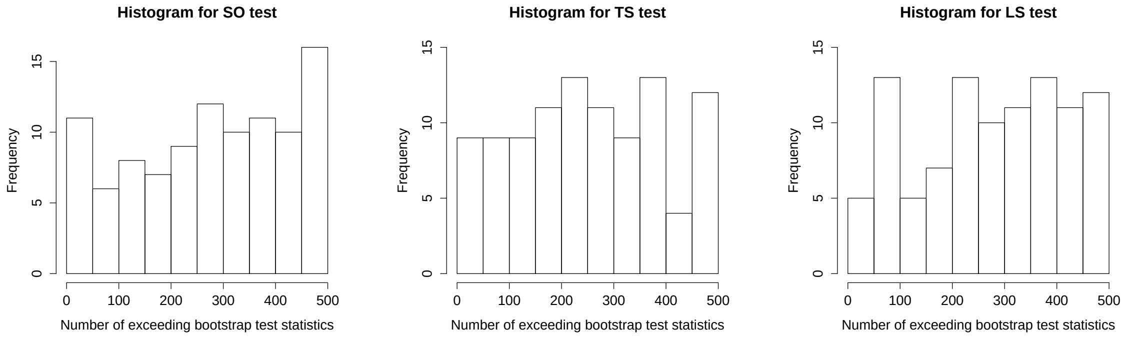
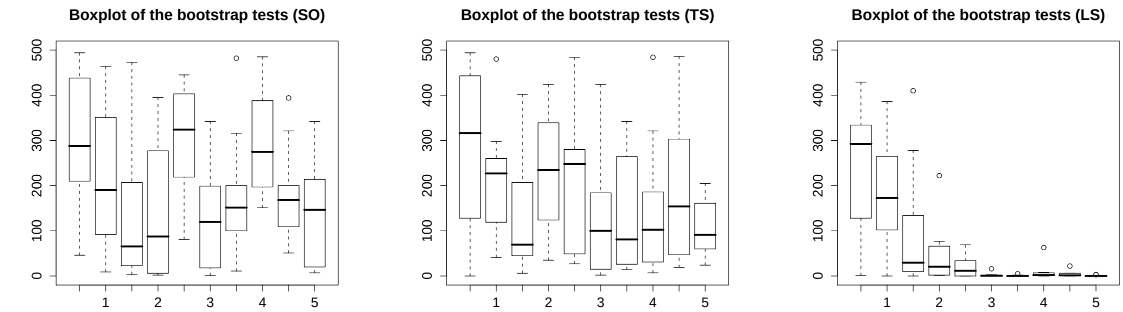


Figure 6: Boxplots of the number of bootstrap replicates (out of $B = 500$) for which $\tilde{T}_n$ takes a larger value than for the original data in case of an LS of increasing size, obtained from 30 repetitions each: test for SO (left), TS (middle) and LS (right).



while the reverse applies less frequently. Apparently the power of the test for LS increases to 1 with increasing size of the shift, and there is some confusion with TS or SO only if the shift is small. We also illustrate the results for another more challenging situation of time series with length $n = 100$ from an INGARCH(1,1) model with a smaller marginal mean. For this example the parameters of the INGARCH model are given by $(\beta_0, \alpha_1, \beta_1) = (2, 0.3, 0.3)$ and the data contains an intervention at $\tau = 50$. Observe that the power of all tests for the correct type of intervention approaches 1, except for SO as it is occasionally mistaken as a TS. This is in part due to the above preference ordering. A remedy might be to increase the number B of bootstrap replicates to improve the classification.

6. Iterative detection of intervention effects

Since there can be more than one intervention in a time series, procedures for iterative detection of multiple outliers and data cleaning have been suggested by Tsay (1986) and Chen and Liu (1993), among others. Based on the previous findings, we suggest the following approach for stepwise detection, classification and elimination of multiple intervention effects, setting $Z_t^\star = Z_t$, $t = 1, \ldots, n$, for initialization:

1. Fit an INGARCH(p, q) model (1) to the data $\{Z_t^\star\}$.
2. Test for a single intervention of any type at any time point by employing (4) and using the maximum of the score test statistics as described in Section 5.
3. If there is no significant result, then stop; the data $Z_1^\star, \ldots, Z_n^\star$ are considered as clean. Otherwise:
   (a) Fit a contaminated INGARCH(p, q) model (4) by choosing $\xi(\mathcal{B})$ according to the type of intervention identified in the previous step. Let $\hat{\nu}$ be the estimated size of the intervention effect and τ its point in time.
   (b) Estimate the effect of the intervention on the observation $Z_t^\star$ by the rounded value
   $$\hat{C}_t = \left[\frac{\hat{\mu}_t}{\hat{\kappa}_t} Z_t\right]$$
   where $\hat{\kappa}_t$ is obtained from equation (4) by plugging in the estimates of the model parameters and
   $$\hat{\mu}_t = \sum_{i=1}^{q} \hat{\beta}_i \hat{C}_{t-i} + \sum_{j=1}^{p} \hat{\alpha}_j \hat{\mu}_{t-j} + \hat{\nu} X_t \ , t = \tau, \tau+1, \ldots,$$
   with $\hat{C}_t = \hat{\mu}_t = 0$ for $t < \tau$.
   (c) Correct the time series for the estimated intervention effects,
   $$Z_t^\star = Z_t^\star - \hat{C}_t, \ t \geq \tau,$$
   and return to step 1.

The iterative procedure is continued until no further interventions are detected in step 2.

It is not possible to exactly eliminate intervention effects from the series since $\hat{\nu}$ is the effect to the unobserved mean process at the time $t = \tau$ of its occurrence. We argue that the correction in step 3 (c) is adequate if we have identified the correct type of intervention and point in time τ. Let $\{Y_t\}$ be an uncontaminated INGARCH(p, q) process generated from model (1), and let $\{C_t\}$ be a contaminating process which is independent from $\{Y_t\}$ with,

$$\begin{aligned} C_t|\mathcal{F}_{t-1}^C &\sim \text{Poisson}(\mu_t), \\ \mu_t &= \sum_{i=1}^{q} \beta_i C_{t-i} + \sum_{j=1}^{p} \alpha_j \mu_{t-j} + \nu X_t, \end{aligned} \tag{14}$$

for $t \geq \tau$, with $\mu_1 = \ldots = \mu_{\tau-1} = 0$ and $\mathcal{F}_t^C$ the σ-field generated from $\{C_{1-q}, \ldots, C_t, \mu_{1-p}, \ldots, \mu_0\}$. Denoting $\mathcal{F}_t^{(Y,C)} = \mathcal{F}_t^Y \vee \mathcal{F}_t^C$, the σ-field generated from $\mathcal{F}_t^Y$ and $\mathcal{F}_t^C$ and assuming independence between $\{Y_t\}$ and $\{C_t\}$, we get for $Z_t = Y_t + C_t$ that

$$\begin{aligned} Z_t|\mathcal{F}_{t-1}^{(Y,C)} &\sim \text{Poi}(\kappa_t) \\ \kappa_t = \lambda_t + \mu_t &= \sum_{i=1}^{q} \beta_i Z_{t-i} + \sum_{j=1}^{p} \alpha_j(\lambda_{t-j} + \mu_{t-j}) + \nu X_t \,. \end{aligned}$$

Note that Z_t depends on the past only via $Z_{1-q}, \ldots, Z_{t-1}, \lambda_{1-p}, \ldots, \lambda_0$, i.e. it is conditionally independent from $\mathcal{F}_{t-1}^{(Y,C)}$ given these variables. Thus, it possesses the same stochastic properties as the contaminated process $\{Z_t\}$ generated from (4). Accordingly, we clean the observed time series $\{Z_t\}$ by subtracting a prediction $\hat{C}_t$ of C_t from Z_t. We use the conditional expectation of C_t given $\mathcal{F}_{t-1}^{(Y,C)}$ and Z_t for this, which is

$$\begin{aligned} E\left(C_t|Z_t = z, \mathcal{F}_{t-1}^{(Y,C)}\right) &= \sum_{i=0}^{z} iP\left(C_t = i|Z_t = z, \mathcal{F}_{t-1}^{(Y,C)}\right) = \sum_{i=0}^{z} i\frac{P\left(C_t = i, Y_t = z-i|\mathcal{F}_{t-1}^{(Y,C)}\right)}{P\left(Z_t = z|\mathcal{F}_{t-1}^{(Y,C)}\right)} \\ &= \sum_{i=0}^{z} i\frac{(\mu_t^i e^{-\mu_t}/i!) \cdot \lambda_t^{z-i} e^{-\lambda_t}/(z-i)!}{\kappa_t^z e^{-\kappa_t}/z!} = \sum_{i=0}^{z} i\binom{z}{i}\left(\frac{\mu_t}{\kappa_t}\right)^i \left(\frac{\lambda_t}{\kappa_t}\right)^{z-i} = \frac{\mu_t}{\kappa_t} z \,. \end{aligned}$$

Into the resulting formula $E\left(C_t|Z_t, \mathcal{F}_{t-1}^{(Y,C)}\right) = \mu_t Z_t/\kappa_t$ we plug in the estimates of μ_t and κ_t, obtained from the recursions (14) and (4), respectively, and using the parameter estimates $\hat{\beta}_0, \ldots, \hat{\beta}_q, \hat{\alpha}_1, \ldots, \hat{\alpha}_p$.

We demonstrate this methodology by analyzing two data examples: a simulated one where we can compare our findings to the ground truth, and the real data described in the introduction.

As a first example we consider a simulated time series of length $n = 200$ generated from an INGARCH(1,1)-model with parameters $\beta_0 = 0.5$, $\alpha_1 = 0.3$ and $\beta_1 = 0.5$, see Figure 8. To check the sensitivity of the results concerning misspecification of the parameter δ regarding transient shifts, we generate data with two transient shifts of the same size $\nu = 10$ at times $\tau_1 = 50$ and $\tau_2 = 150$, using $\delta = 0.7$ and $\delta = 0.9$, respectively. We test the existence of an SO, LS or TS with $\delta = 0.8$ at any time point.

When we fit model (1) to the data assuming that there are no interventions, we obtain the conditional maximum likelihood estimates $(\tilde{\beta}_0, \tilde{\alpha}_1, \tilde{\beta}_1) = (0.522, 0.295, 0.602)$. Then we calculate the test statistics for SO, TS and LS at all time points, see Figure 8. The test statistic for TS assumes significant values both at time $t = 50$ (with a value of 62.5) and at $t = 150$ (with a value of 44.1). However, the test statistics for SO are also significant at the same instances, namely 100.8 and 31.4, respectively. The corresponding bootstrap p-values equal zero except for the SO at time $t = 150$, for which it is 0.006. The test statistics for LS are not significant with a maximum of 7.1 (p-value 0.36). According to the proposed classification rule, we might identify correctly a TS at time $\tau = 150$. However, here we want to investigate the effects of a misclassification and decide incorrectly in favor of a SO at time $\tau_1 = 50$, and estimate its size as 12.783 by joint conditional maximum likelihood, see Section 3.

Iterating the algorithm further, always correcting the data according to step 3(c) of the iterative algorithm, we conclude the analysis by detecting a TS of size 8.122 at time $\tau_2 = 150$ (p-value 0, while the p-values are large for SO and LS). After correcting the data for its effects there are no further significant test statistics in the third iteration. The final parameter estimates $(\hat{\beta}_0, \hat{\alpha}_1, \hat{\beta}_1) = (0.515, 0.310, 0.529)$ are quite close to the true values. The classification of transient shifts TS works quite reliably here in spite of the misspecified value of δ, even though there is some risk of misclassifying a TS for an SO, particularly if a too large value of δ is used in the testing.

Finally we investigate whether there are intervention effects in the time series of campylobacterosis infection cases illustrated in Figure 1. Notice that for the analysis of these data, Ferland et al. (2006) suggest a kind of seasonal model which includes Y_{t-1} and λ_{t-13} as regressors in the equation for λ_t, see (1). We follow this approach and start by fitting the same seasonal model.

Table 2: Iterative parameter estimates and intervention effects for the campylobacterosis data.

Iteration	Parameter Estimates			Outlier		
	$\tilde{\beta}_0$	$\tilde{\alpha}_1$	$\tilde{\beta}_1$	Type	τ	Size
1	2.439	0.196	0.591	LS	84	7.64
	(0.654)	(0.077)	(0.058)			
2	3.681	0.150	0.409	SO	100	22.55
	(1.207)	(0.159)	(0.081)			
3	2.300	0.387	0.323	–		
	(1.005)	(0.162)	(0.090)			

After fitting the seasonal INGARCH model we test for the different types of interventions using the bootstrap versions of the maximum score tests described in Section 5. In the first iteration, the bootstrap p-values equal zero for spiky (SO) and transient shift (TS) types of outliers at time $t = 100$ (recall that δ is set to 0.8) and also for a level shift (LS) at time $t = 84$. According to the above classification rule we decide in favor of the LS, estimate its size as 7.64 and eliminate its effect from the time series. The top right plot of Figure 9 shows the corrected series.

The we fit the seasonal model to the cleaned data and test for the presence of further intervention effects. Now we detect an SO of size 22.55 at time 100 with a p-value of 0 (the test statistic for SO is 38.66 and thus much larger than the one for TS here, which is 25.82 and would also have been significant). After cleaning the data from the effects of an SO at time 100 and refitting the seasonal model to the cleaned data all p-values become large. Therefore we conclude the analysis with the two mentioned interventions identified.

Table 2 summarizes the results of the iterations. Fitting the full model with the two interventions to the original data, we conclude with the following enlarged INGARCH(1,1) model for the numbers of campylobacterosis infections:

$$
\begin{aligned}
Z_t \mid \mathcal{F}_{t-1} &\sim \text{Poisson}(\kappa_t) \\
\kappa_t &= 3.584(.761) + 0.230(.081)\kappa_{t-13} + 0.352(.054)Z_{t-1} \\
&\quad +2.930(.887)I(t \geq 84) + 41.645(7.302)I(t = 100), \qquad (15)
\end{aligned}
$$

for $t \geq 1$, where in parentheses are the corresponding standard errors of the regression coefficients. Based on model (15) and setting $\hat{\kappa}_t = \kappa_t(\hat{\theta})$, both predictions and data are plotted in Figure 9, illustrating that the method successfully accomodates unusual observations and fits the data quite well. As a closing remark, the mean square error of the residuals for a model which does not include intervention effects is equal to 2.309. The corresponding value for (15) equals 1.313 which clearly indicates the improved fit.

Alternatively, we could have decided in favor of an SO at time 100 in the first step since the value of its test statistic is the largest among all score tests, namely 91.0 versus 82.1 for a TS and 33.2 for a LS. These differences appear to be large as compared to the differences between the test statistics for the bootstrap replicates, since the 95% percentiles of these are 16.1 (SO), 14.6 (TS) and 12.5 (LS) while the maxima are 22.9 (SO), 22.1 (TS) and 25.3 (LS), respectively. Then we would have estimated the size of the SO as 36.34. Refitting the INGARCH model and testing for the presence of further interventions, we would have identified another SO at time 125 in the second step, with an estimated size of 15.01 (p-value 0.006, which is smaller than those for an TS at time 111 and an LS at time 83, namely 0.014 and 0.028). In the third iteration a TS of size 14.75 would have been identified at $\tau_3 = 111$ (p-value 0.00, while the one for an SO at time 111 is 0.010). The possible LS at time 83 would have lost its significance then (test statistic 9.25, p-value 0.237), i.e. we would have detected several intervention effects with the same direction instead of an LS.

7. DISCUSSION

This work analyzes a model for estimation and detection of intervention effects in count time series by both theory and simulation. A model for the data is necessary in this context since otherwise it is not clear how distant an observation needs to be from the rest of the data to be called an outlier, see Davies and Gather (1993). The results indicate that the suggested modeling approach is reasonable. In particular, the proposed score test statistics have been proven useful to detect different types of interventions, although the tests for a certain intervention at a given point in time seem to be oversized. Masking and swamping effects due to multiple outliers apparently pose less difficulties than they do in the case of linear time series models like ARMA (see Peña (2000)). This can be explained by the estimation of the mean underlying the INGARCH process except for the estimates of the model parameters depending only on the past of time t. Gather et al. (2002) suggest a similar approach to reduce masking and swamping effects in outlier detection for ARMA models.

We feel that model (4) suggests more problems and further research into different directions. For instance, as a generalization with m intervention effects we fit the model

$$\begin{aligned} Z_t|\mathcal{F}_{t-1} &\sim \text{Poisson}(\kappa_t), \quad t \geq 1 \\ \kappa_t &= \beta_0 + \sum_{i=1}^{q} \beta_i Z_{t-i} + \sum_{j=1}^{p} \alpha_j \kappa_{t-j} + \sum_{k=1}^{m} \nu_k \xi_k(\mathcal{B}) I_t(\tau_k), \end{aligned}$$

for joint estimation of intervention effects and model parameters. Some caution though should be exercised since the properties of such processes are not known. To develop inference based on this model, we have proposed an iterative strategy along the lines of Chen and Liu (1993). Discrimination between the different interventions merits further investigation. Based on our results we conjecture that the bootstrap approach presented here works well at least for moderately large intervention effects which are not very close to the start or the end of the series. Otherwise, tests for different interventions can take similarly small p-values. The behavior of the extreme tails of the test statistics in finite samples thus deserves further investigation.

We have provided computationally tractable solutions for intervention effects influencing the future of the process according to its dynamics. The case of a single additive outlier due to e.g. a measurement artifact has been neglected here since joint parameter estimation is not straightforward within our framework and since count time series arising e.g. in epidemiology are controlled well, rendering the occurrence of AOs unlikely. The investigation of such outliers is a topic for further research.

The approach taken in this work is quite general and can be used as a guidance for other dependent count data or more general time series following generalized linear models. For instance, a model which has been often used in the literature for the analysis of count time series is the so called log–linear model, see Zeger and Qaqish (1988), for instance. The proposed method can be reformulated to accommodate such models after suitable modifications.

Acknowledgements

The authors thank Professor A. Latour for kindly providing the campylobacterosis data. The second author gratefully acknowledges financial support of the Deutsche Forschungsgemeinschaft (SFB 475, ”Reduction of complexity in multivariate data structures”).

References

Basawa, I. V. (1991). Generalized score tests for composite hypotheses. In *Estimating Functions*, pp. 121–131. New York: Oxford Univ. Press.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327.

Box, G. E. P. & Tiao, G. C. (1975). Intervention analysis with applications to economics and environmental problems. *Journal of the American Statistical Association* **70**, 70–79.

Carnero, A. M., Peña, D. & Ruiz, E. (2006). Effect of outliers on the identification and estimation of GARCH models. *Journal of Time Series Analysis* **28**, 471–497.

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

Charles, A. & Darnè, O. (2005). Outliers and GARCH models in financial data. *Economic Letters* **86**, 347–352.

Chen, C. & Liu, L.-M. (1993). Joint estimation of model parameters and outlier effects in time series. *Journal of the American Statistical Association* **88**, 284–297.

Cox, D. R. (1981). Statistical analysis of time series: Some recent developments. *Scandinavian Journal of Statistics* **8**, 93–115.

Davies, P.L. & Gather, U. (1993). The identification of multiple outliers. Invited paper with discussion and reply. *J. Am. Statist. Assoc.* **88**, 782–792.

Davis, R. A., Dunsmuir, W. T. M. & Streett, S. B. (2003). Observation-driven models for Poisson counts. *Biometrika* **90**, 777–790.

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

Fokianos, K. & Kedem, B. (2004). Partial likelihood inference for time series following generalized linear models. *Journal of Time Series Analysis* **25**, 173–197.

Fokianos, K., Rahbek, A. & Tjøstheim, D. (2009). Poisson autoregression. *Journal of the American Statistical Association* to appear.

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

Gather, U., Bauer, M. & Fried, R. (2002). The identification of multiple outliers in online monitoring data. *Estadística* **54**, 289–338.

Jung, R. C., Kukuk, M. & Liesenfeld, R. (2006). Time series of count data: modeling, estimation and diagnostics. *Computational Statistics & Data Analysis* **51**, 2350–2364.

Justel, A., Peña, D. & Tsay, R.S. (2001). Detection of outlier patches in autoregressive time series. *Statistica Sinica* **11**, 651–673.

Kedem, B. & Fokianos, K. (2002). *Regression Models for Time Series Analysis*. Hoboken, NJ: Wiley.

Li, W. K. (1994). Time series models based on generalized linear models: some further results. *Biometrics* **50**, 506–511.

MacDonald, I. L. & Zucchini, W. (1997). *Hidden Markov and Other Models for Discrete–valued Time Series*. London: Chapman & Hall.

Peña, D. (2000). Outliers, influential observations and missing data. In *A course in time series analysis.* Peña, D., Tiao, G.C., Tsay, R.S. (eds.) Wiley, New York.

R Development Core Team (2004). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. 3-900051-07-0.

Rydberg, T. H. & Shephard, N. (2000). A modeling framework for the prices and times of trades on the New York stock exchange. In W. J. Fitzgerlad, R. L. Smith, A. T. Walden, and P. C. Young (Eds.), *Nonlinear and Nonstationary Signal Processing*, pp. 217–246. Cambridge: Isaac Newton Institute and Cambridge University Press.

Streett, S. (2000). *Some observation driven models for time series of counts*. Ph. D. thesis, Colorado State University, Department of Statistics.

Tsay, R. S. (1986). Time series model specification in the presence of outliers. *Journal of the American Statistical Association* **81**, 132–141.

Van Dijk, D., Franses, P. H. & Lucas, A. (1999). Testing for ARCH in the presense of additive outlier. *Journal of Applied Econometrics* **14**, 539–562.

Zeger, S. L. & Qaqish, B. (1988). Markov regression models for time series: a quasi-likelihood approach. *Biometrics* **44**, 1019–1031.

Figure 7: Power of the bootstrap tests with classification based on the minimal p-value for intervention effects of increasing size: level shift at $\tau = 100$ in a series of length $n = 200$ with parameters $(\beta_0, \alpha_1, \beta_1) = (5, 0.3, 0.3)$ (top left), level shift at $\tau = 50$ in a series of length $n = 100$ with parameters $(\beta_0, \alpha_1, \beta_1) = (2, 0.3, 0.3)$ (top right), transient shift at $\tau = 50$ in a series of length $n = 100$ with parameters $(\beta_0, \alpha_1, \beta_1) = (2, 0.3, 0.3)$ (bottom left) and spiky outlier at $\tau = 50$ in a series of length $n = 100$ with parameters $(\beta_0, \alpha_1, \beta_1) = (2, 0.3, 0.3)$ (bottom left). Test for SO (dotted), TS (dashed) and LS (solid).

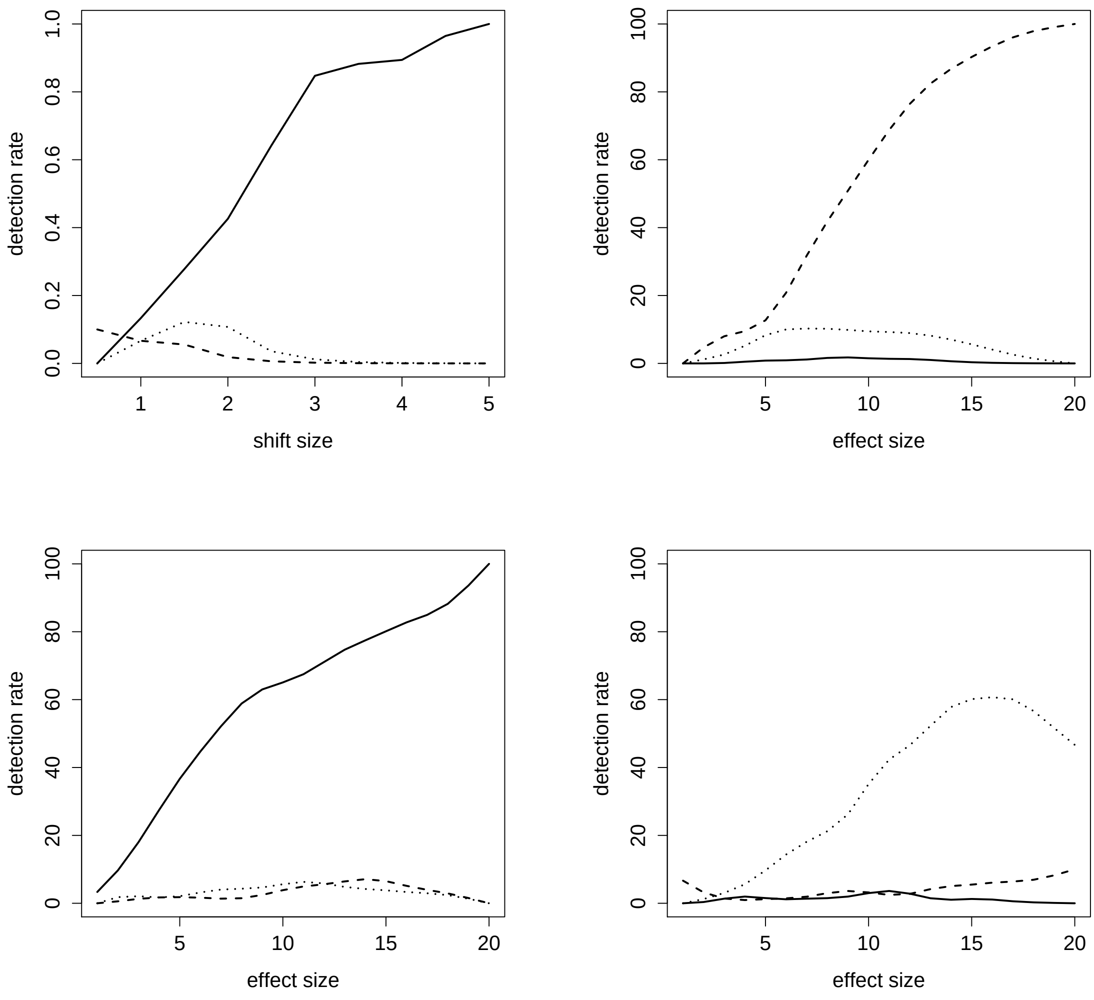


Figure 8: Simulated time series with two transient outliers at times $\tau_1 = 50$ and $\tau_2 = 150$ (dots) as well as the test statistics for SO (dotted line), TS (dashed line) and LS (solid line), all divided by 2 for illustration.

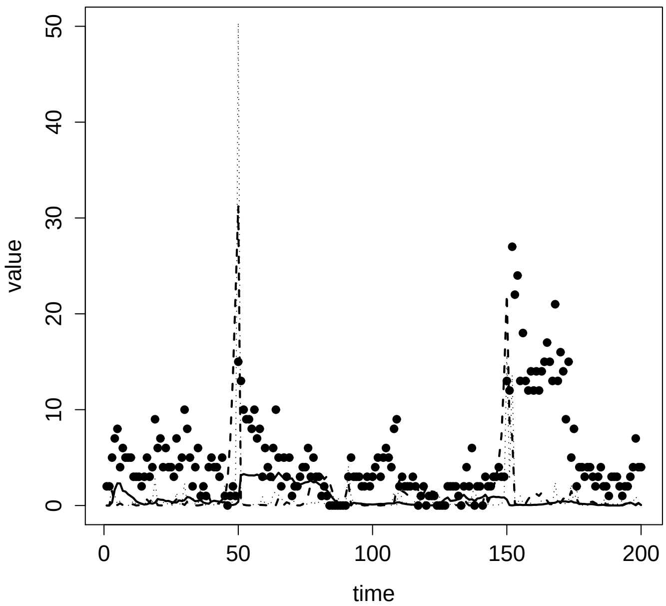


Figure 9: Iterations for correcting the campylobacterosis data: time series (dashed line), SO (light grey line), TS (dark grey line), and LS test statistics (bold black line): first (top left), second (top right) and third iteration (bottom left) and fitted model (bottom right, model with interventions black, model without interventions grey).