

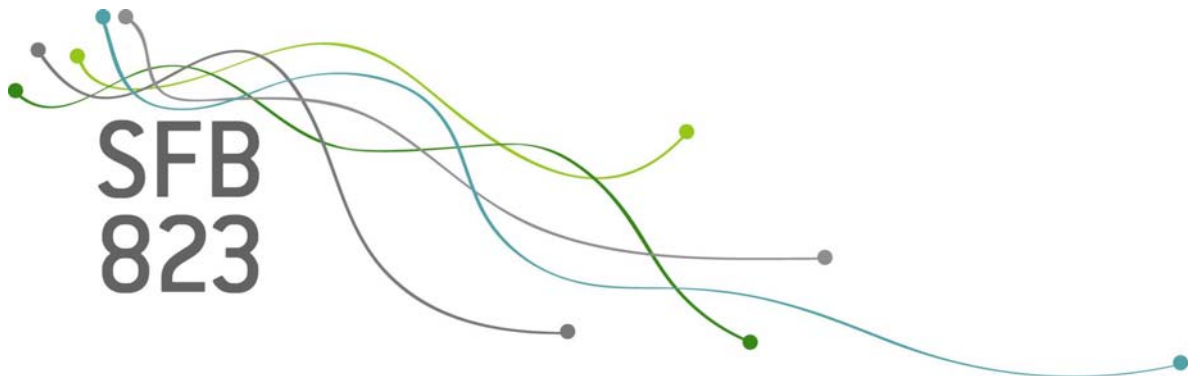
SFB
823

Modelling interventions in INGARCH processes

Tobias Liboschik, Pascal Kerschke,
Konstantinos Fokianos, Roland Fried

Nr. 3/2013

Discussion Paper



Modelling interventions in INGARCH processes

Tobias Liboschik^{a,*}, Pascal Kerschke^a, Konstantinos Fokianos^b, Roland Fried^a

^a*Technische Universität Dortmund, Department of Statistics, Germany*

^b*University of Cyprus, Department of Mathematics and Statistics, Cyprus*

Abstract

We study different approaches to describe intervention effects within the framework of integer-valued GARCH (INGARCH) models for time series of counts. Fokianos and Fried (J. Time Ser. Anal. 2010, 31: 210–225) treat a model where an intervention affects the non-observable underlying mean process at the time point of its occurrence and additionally the whole process thereafter via its dynamics. As an alternative, we consider a model where an intervention directly affects the observation at the time point of its occurrence, but not the underlying mean, and then also enters the dynamics of the process. While the former definition describes an internal change, the latter can be understood as an external effect on the observations due to e.g. immigration. For our alternative model we develop conditional likelihood estimation and, based on this, tests and detection procedures for intervention effects. Both models are compared analytically and using simulated and real data examples. We study the effect of misspecification on the fitted intervention model. Special attention is paid to computational issues.

Keywords: Time series of counts, generalised linear models, change-point detection, level shifts, transient shifts, outliers

1. Introduction

Time series of counts naturally appear in various fields whenever a number of events per time period is observed over time. Examples for the wide range of practical applications are the monthly number of cases in epidemiology, the number of stock market transactions per minute in finance and the number of photon arrivals per microsecond measured in a biological experiment.

In many applications unusual external effects or measurement errors can lead to sudden or gradual changes in the structure of the data or to singular observations of a different nature than the others. Following Box and Tiao (1975) we use the term intervention for

*Corresponding author

Email addresses: liboschik@statistik.tu-dortmund.de (Tobias Liboschik), pascal.kerschke@tu-dortmund.de (Pascal Kerschke), fokianos@ucy.ac.cy (Konstantinos Fokianos), fried@statistik.tu-dortmund.de (Roland Fried)

all kinds of unusual effects disturbing the ordinary pattern of the data, including structural changes and all types of outliers. One goal of an intervention analysis can be to examine the effect of known interventions, for example to judge whether a policy change had the intended impact (for practical applications see for example Box and Tiao, 1975). Another possible goal is to search for unknown intervention effects and to find explanations for them *a posteriori*. Based on the belief that intervention effects are meaningful, they are often included in the model by deterministic covariates (e.g. Box and Tiao, 1975; Abraham and Box, 1979). This approach is fundamentally different from a robust approach, where methods are expected to simply ignore the effect of aberrant observations. The main focus of this paper is to model various types of interventions in so-called INGARCH processes, which form one particular model class for time series of counts and are defined later on.

Count data time series have some features which make it difficult to detect and model interventions. Restriction to non-negative values leads to a usually positively skewed marginal distribution with a considerable probability for large positive values, so that upward outliers are difficult to identify. Detection of a downward change towards zero is, at least for small means, hardly possible anyway. In addition, an intervention at a single point in time may affect not just one but several subsequent observations due to serial dependence. And lastly, models for intervention effects need to be preserve the discreteness of the observations.

For modelling real-valued time series one often assumes a simple (and often linear) autoregressive model for an observation conditionally on its past, like in the popular class of autoregressive moving average (ARMA) models and its extensions. However, for time series of counts this would not ensure the observations to be non-negative integers. A natural extension solving this problem is to use a generalised linear model (GLM) for the observations conditionally on the past information, choosing a distribution suitable for count data and an appropriate link function. This approach of time series following GLMs is pursued by Kedem and Fokianos (2002), among others. An important special case of this class of models is the autoregressive conditional Poisson (ACP) model studied by Heinen (2003). Ferland et al. (2006) further studied these processes and call them integer-valued GARCH (INGARCH) processes, because of the analogy to generalised autoregressive conditional heteroscedastic (GARCH) processes. These non-linear processes use the identity link function for the conditional mean and assume the observations, conditionally on its past, to follow a Poisson distribution. A process $\{Y_t\}_{t \in \mathbb{N}}$ is called INGARCH(p, q) process of order p and $q \in \mathbb{N}_0$, if

$$Y_t | \mathcal{F}_{t-1}^{Y, \lambda} \sim \text{Poisson}(\lambda_t), \quad (1)$$

$$\lambda_t = \beta_0 + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{j=1}^q \alpha_j \lambda_{t-j}, \quad (2)$$

with $\beta_0 > 0$ an intercept, $\beta_1, \dots, \beta_p, \alpha_1, \dots, \alpha_q \geq 0$ parameters describing the dependence structure and the σ -field $\mathcal{F}_{t-1}^{Y, \lambda} = \sigma(\{Y_{1-p}, \dots, Y_{t-1}, \lambda_{1-q}, \dots, \lambda_0\})$ containing the information about the past of the process up to time $t-1$. The conditional mean at a certain time depends on past observations and past values of the conditional mean. This feedback mechanism also describes overdispersion, i.e. the variance is greater than the mean, which is often observed

for time series of counts. Despite the dependence on unobserved past conditional means, this model is observation- and not parameter-driven in the classification of Cox (1981), as these conditional means in turn depend on past observations.

In this paper we introduce a new model to describe intervention effects within the class of INGARCH processes and compare it to an existing intervention model proposed by Fokianos and Fried (2010). Our new model is able to model interventions which enter the dynamics of the process in a different way than with the existing intervention model. This allows to describe an external effect on the observed value rather than an internal change of the underlying state of the process, which can be more plausible in some applications. Section 2 defines both intervention models and compares them analytically. Section 3 presents joint maximum likelihood estimation of the ordinary model parameters and the intervention effects and studies its properties by simulations. Section 4 presents asymptotic procedures to test for single interventions at a given time or to detect one or multiple interventions whose positions are unknown. We verify these procedures by simulations. Furthermore, we briefly discuss the problem of misspecification of the intervention model. Section 5 applies some of the methods to a real data example. Section 6 concludes with some final remarks.

2. Intervention models

As a first intervention model in the framework of INGARCH(p, q) processes, Fokianos and Fried (2010) propose to define a process $\{Z_t\}_{t \in \mathbb{N}}$ with an intervention at time τ by

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

where $\{X_t\} = \{\delta^{t-\tau} I_{[\tau, \infty)}(t)\}$ is a deterministic process describing the intervention effect, $\nu > -\beta_0$ (to ensure $\kappa_t > 0$) is the intervention size, the predefined constant $\delta \in [0, 1]$ specifies the type of intervention and all other model parameters are as before. The information about the past of the process is denoted by $\mathcal{F}_{t-1}^{Z, \kappa} = \sigma(\{Z_{1-p}, \dots, Z_{t-1}, \kappa_{1-q}, \dots, \kappa_0\})$. Their intervention model covers three types of interventions with different choices of δ :

- Spiky outlier (SO) for $\delta = 0$
- Transient shift (TS) for $\delta \in (0, 1)$
- Level shift (LS) for $\delta = 1$

Simulated examples of all three types of interventions and without an intervention are shown in Figure 1. We used different sizes ν for the three types, because for a fixed intervention size an intervention effect is harder to detect if δ is lower. The intervention effect is added to the underlying conditional mean process $\{\kappa_t\}$ and not to the observations itself. Through the autoregression on past observations and past conditional means the intervention effect enters the actual dynamics of the process. The parameter δ regulates how quickly the direct

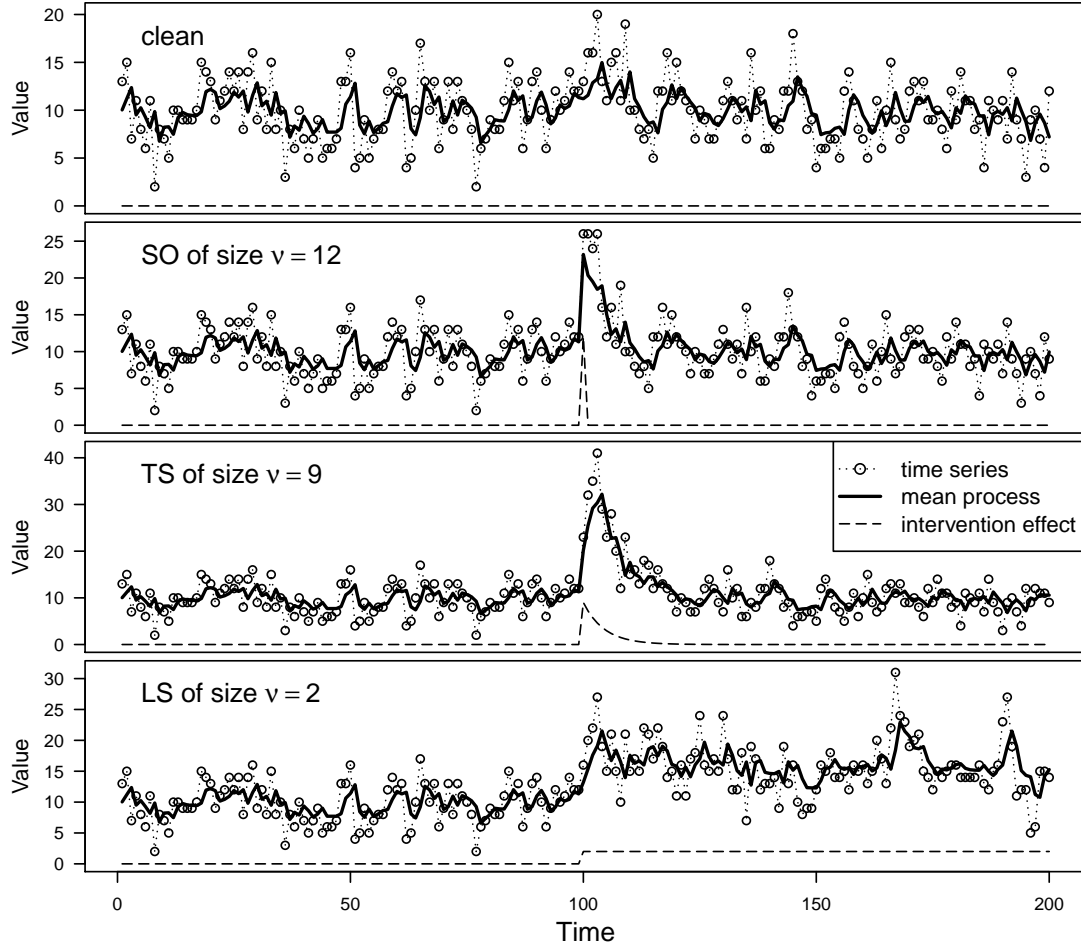


Figure 1: Examples for simulated INGARCH(1,1) time series with parameters $\beta_0 = 3$, $\beta_1 = 0.4$ and $\alpha_1 = 0.3$ with an intervention of the given type and size at time $\tau = 100$ according to model (i) (TS with $\delta=0.8$).

impact of the intervention on future values of the mean vanishes, whereas its indirect impact is governed by the dependence structure of the process.

We study an alternative intervention model for INGARCH(p,q) processes $\{Z_t\}$ following the definition

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

with $\{X_t\}$, the regular INGARCH parameters and the additional intervention parameters as before. The difference between both intervention models becomes more obvious, if we

rewrite the equation for the conditional mean κ_t of intervention model (ii) to

$$\kappa_t = \beta_0 + \sum_{i=1}^p \beta_i Z_{t-i} + \sum_{j=1}^q \alpha_j (\kappa_{t-j} - \underbrace{\nu X_{t-j}}_{\text{not for model (i)}}) + \nu X_t, \quad (3)$$

which is apart from the labelled term identical to the conditional mean equation of intervention model (i). In the alternative intervention model (ii) the intervention effect is not propagated via the feedback mechanism of the conditional mean but only via the contaminated observations. From the definition of intervention model (i) it is not obvious that it is sufficient to define $\mathcal{F}_{t-1}^{Z,\kappa} = \sigma(\{Z_{1-p}, \dots, Z_{t-1}, \kappa_{1-q}, \dots, \kappa_0\})$ as for the first intervention model, because $\mathcal{F}_{t-1}^{Z,\kappa}$ seems to contain only information on past values of κ_t and not of λ_t . However, from Equation (3) we see that both intervention models can be written in a unified way and are fully specified by their observable process $\{Z_t\}$ and their unobservable conditional mean process $\{\kappa_t\}$ (and of course also their deterministic intervention process $\{X_t\}$).

One could gain more insight into the difference between the intervention models by considering another representation, which is also useful for cleaning the processes from intervention effects (see Section 4). For intervention model (i) Fokianos and Fried (2010) show that it is possible to decompose $\{Z_t\}$ into an intervention-free process $\{Y_t\}$ and a contamination process $\{C_t\}$, which are mutually independent, such that $Z_t = Y_t + C_t$. The intervention-free process $\{Y_t\}$ is an INGARCH(p, q) process with conditional mean process $\{\lambda_t\}$ and the same parameters as $\{Z_t\}$. The contamination process $\{C_t\}$ underlies the same dependences as $\{Z_t\}$ and can be described by

$$C_t | \mathcal{F}'_{t-1} \sim \text{Poisson}(\mu_t),$$

$$\mu_t = \sum_{i=1}^p \beta_i C_{t-i} + \underbrace{\sum_{j=1}^q \alpha_j \mu_{t-j}}_{\text{not for model (ii)}} + \nu \delta^{t-\tau} I_{[\tau, \infty)}(t), \quad (4)$$

with $\{\mu_t\}$ its conditional mean process and $\kappa_t = \lambda_t + \mu_t$. It is easily shown that a similar decomposition holds for intervention model (ii) without the labelled summand. Naturally μ_t and thus also C_t are zero for $t < \tau$.

In both intervention models the intervention affects the conditional mean from time τ on. However, in intervention model (ii) its effect does not enter the dynamics of the process directly but only via the observations. This can be seen more clearly in Equation (4): Only for intervention model (i) the intervention effect propagates directly via the conditional mean μ_t . The difference can also be seen from Equation (3), where for intervention model (ii) the feedback mechanism is adjusted by the intervention effect. The larger the feedback parameters α_i are, the greater is the difference between both models. For $\alpha_1 = \dots = \alpha_q = 0$ both models are equivalent.

The simulated example of a transient shift in Figure 2 illustrates the difference between the intervention models. After the intervention the conditional mean returns to its undisturbed level more quickly for intervention model (ii) than for model (i).

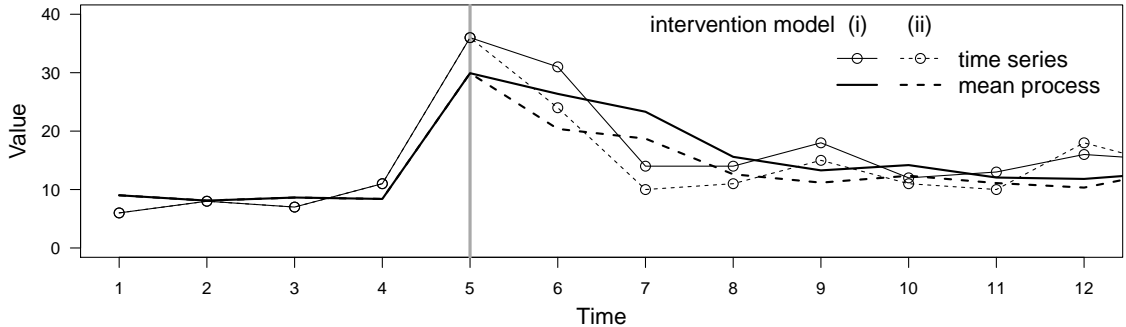


Figure 2: Comparison of the intervention models for a simulated INGARCH(1,1) processes with parameters $\beta_0 = 3$, $\beta_1 = 0.4$ and $\alpha_1 = 0.3$ with a transient shift ($\delta = 0.8$) of size $\nu = 20$ at time $\tau = 5$.

An intervention following model (ii) can be interpreted as a sudden external effect, whilst an intervention following model (i) can be rather seen as an internal change of the data generating process. We will therefore refer to model (i) as the “internal” intervention model and to model (ii) as the “external” one. For illustration consider the weekly number of registered cases of a disease spreading human to human and by contaminated nutrition. A level shift could be for instance caused by a change of the reporting requirements, an enhanced detection method or just a new and more aggressive type of pathogen. A spiky outlier in the external intervention model could be caused by the return of an exceptionally high number of infected persons from another country, which will affect the future number of cases only by human to human spread. In contrast, a spiky outlier from the internal intervention model could be caused by a large quantity of contaminated food resulting in much more infections than usual. This will affect the future number of cases not only by human to human spread of the additionally infected patients, but also by a larger number of pathogens circulating for example in food processing establishments.

Note that purely additive outliers which do not enter the dynamics of the process cannot be handled satisfactory with the likelihood approach chosen for this work. Fried et al. (2012) apply a Bayesian approach to detect additive outliers in INGARCH processes but point out that several outliers of the same size are needed for reliable estimation without prior knowledge.

3. Estimation

Estimation of the alternative intervention model (ii) can be done according to Fokianos and Fried (2010) with some modifications. Fokianos and Fried treat the intervention as a time-dependent covariate process $\{X_t\}$ and jointly estimate the vector of unknown model parameters $\theta = (\beta_0, \beta_1, \dots, \beta_p, \alpha_1, \dots, \alpha_q, \nu)'$ by conditional maximum likelihood (CML) along the lines of Ferland et al. (2006) and Fokianos et al. (2009), who treat intervention-free INGARCH(p, q) processes. We present estimation formulas valid for both intervention models and point out where these differ from each other. Note that the time τ when the intervention occurs as well as the parameter δ specifying the type of intervention are treated

as known for estimation. To ensure stationarity of the process we assume $\boldsymbol{\theta} \in \Theta$ with

$$\Theta = \left\{ \boldsymbol{\theta} \in \mathbb{R}^{p+q+2} \left| \beta_0 > 0, \beta_1, \dots, \beta_p, \alpha_1, \dots, \alpha_q \geq 0, \sum_{i=1}^p \beta_i + \sum_{j=1}^q \alpha_j < 1, \nu > -\beta_0 \right. \right\}.$$

The following equations are conditional on the unobserved past $\mathcal{F}_0^{Z, \kappa}$ of the process. We will discuss later on how to deal with this in practice. For an observed time series $\mathbf{z} = (z_1, \dots, z_n)'$ the conditional log-likelihood function $\ell : \Theta \rightarrow \mathbb{R}$ is up to a constant given by

$$\ell(\boldsymbol{\theta}; \mathbf{z}) = \sum_{t=1}^n (z_t \ln(\kappa_t(\boldsymbol{\theta})) - \kappa_t(\boldsymbol{\theta})),$$

where the conditional mean is regarded as a function $\kappa_t : \Theta \rightarrow \mathbb{R}$ and thus denoted by $\kappa_t(\boldsymbol{\theta})$ for all t . The conditional score function $S_{n\tau} : \Theta \rightarrow \mathbb{R}^{p+q+2}$ is given by

$$S_{n\tau}(\boldsymbol{\theta}; \mathbf{z}) = \frac{\partial \ell(\boldsymbol{\theta}; \mathbf{z})}{\partial \boldsymbol{\theta}} = \sum_{t=1}^n \left(\frac{z_t}{\kappa_t(\boldsymbol{\theta})} - 1 \right) \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}},$$

with the vector of partial derivatives of $\kappa_t(\boldsymbol{\theta})$ given by

$$\frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left(\frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \beta_0}, \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \beta_1}, \dots, \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \beta_p}, \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \alpha_1}, \dots, \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \alpha_q}, \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \nu} \right)',$$

where

$$\begin{aligned} \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \beta_0} &= 1 + \sum_{j=1}^q \alpha_j \frac{\partial \kappa_{t-j}(\boldsymbol{\theta})}{\partial \beta_0}, \\ \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \beta_k} &= y_{t-k} + \sum_{j=1}^q \alpha_j \frac{\partial \kappa_{t-j}(\boldsymbol{\theta})}{\partial \beta_k}, \quad k = 1, \dots, p, \\ \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \alpha_l} &= \kappa_{t-l}(\boldsymbol{\theta}) + \sum_{j=1}^q \alpha_j \frac{\partial \kappa_{t-j}(\boldsymbol{\theta})}{\partial \alpha_l} \underbrace{-\nu \delta^{t-l-\tau} I_{[\tau, \infty)}(t-l)}_{\text{not for model (i)}}, \quad l = 1, \dots, q, \quad \text{and} \\ \frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \nu} &= \underbrace{\sum_{j=1}^q \alpha_j \frac{\partial \kappa_{t-j}(\boldsymbol{\theta})}{\partial \nu}}_{\text{not for model (ii)}} + \delta^{t-\tau} I_{[\tau, \infty)}(t) \end{aligned}$$

are calculated recursively. Finally, the conditional information matrix is given by

$$\mathcal{I}_{n\tau}(\boldsymbol{\theta}) = \sum_{t=1}^n \frac{1}{\kappa_t(\boldsymbol{\theta})} \left(\frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \kappa_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)'$$

The conditional maximum likelihood (CML) estimator $\hat{\boldsymbol{\theta}}_n$ of $\boldsymbol{\theta}$ is the value $\hat{\boldsymbol{\theta}}_n \in \Theta$ solving the system of equations $S_{n\tau}(\hat{\boldsymbol{\theta}}_n; \mathbf{z}) = \mathbf{0}$, provided the existence of such a solution. Finding this solution is equivalent to solving the non-linear constrained optimisation problem

$$\hat{\boldsymbol{\theta}}_n = \arg \max_{\boldsymbol{\theta} \in \Theta} \ell(\boldsymbol{\theta}; \mathbf{z}), \quad (5)$$

i.e. maximising the log-likelihood function over the parameter space Θ . The parameter restriction $\boldsymbol{\theta} \in \Theta$ can in fact be formulated as linear inequality constraints, which means we can find a matrix $\mathbf{U} \sim (p+q+2) \times (p+q+1)$ and a vector $\mathbf{c} \sim (p+q+2) \times 1$, such that $\Theta = \{\boldsymbol{\theta} | \mathbf{U}\boldsymbol{\theta} \geq \mathbf{c}\}$. To enforce the strict inequalities the concerned constraints are tightened by an arbitrarily small constant $\xi > 0$, chosen to be $\xi = 10^{-6}$ in our computations. In our software implementation we solve maximisation problem (5) numerically using the function `constrOptim` of the statistical environment R (R Development Core Team, 2012). This function applies an algorithm described by Lange (1999, Chapter 14), which essentially enforces the constraints by adding a barrier value to the objective function and then employs an algorithm for unconstrained optimisation of this new objective function, iterating these two steps if necessary. We chose the quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm for the latter task of unconstrained optimisation.

The procedure finding the CML estimator depends on the past $\mathcal{F}_0^{Z,\kappa}$ of the process previous to the observed sample, which is of course unknown in practice. For models with order $q > 0$ we would not only need p but all (infinitely many) unobserved pre-sample observations. In our implementation we initialise z_{1-p}, \dots, z_0 and $\kappa_{1-q}(\boldsymbol{\theta}), \dots, \kappa_0(\boldsymbol{\theta})$ with the marginal mean

$$\mu(\boldsymbol{\theta}) = \beta_0 / \left(1 - \sum_{i=1}^p \beta_i - \sum_{j=1}^q \alpha_j \right) \quad (6)$$

of an intervention-free INGARCH(p, q) process (cf. Heinen, 2003). We initialise $\partial \kappa_{1-q}(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}$, \dots , $\partial \kappa_0(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}$ with $\partial \mu(\boldsymbol{\theta}) / \partial \boldsymbol{\theta}$, the vector of partial derivatives of the marginal mean.

3.1. Initial estimation

To solve the non-linear optimisation problem stated above one needs to provide a starting value for the parameter vector $\boldsymbol{\theta}$. In this work we use estimators under the assumption of no intervention, whose adequacy in the presence of an intervention we study later. Initial estimators for $\boldsymbol{\theta}$ can easily be obtained by employing a representation of an INGARCH(p, q) process $\{Y_t\}$ as an ARMA($\max\{p, q\}, p$) process with the same second-order properties, which is for $p = q = 1$ given by Ferland et al. (2006). For arbitrary p and q with $k := \max\{p, q\}$ this representation is, after straightforward calculations, given by

$$(Y_t - \underbrace{\mu(\boldsymbol{\theta})}_{=: \eta}) - \sum_{i=1}^k \underbrace{(\beta_i + \alpha_i)}_{=: \psi_i} (Y_{t-i} - \mu(\boldsymbol{\theta})) = \varepsilon_t + \sum_{i=1}^q \underbrace{(-\alpha_i)}_{=: \phi_i} \varepsilon_{t-i},$$

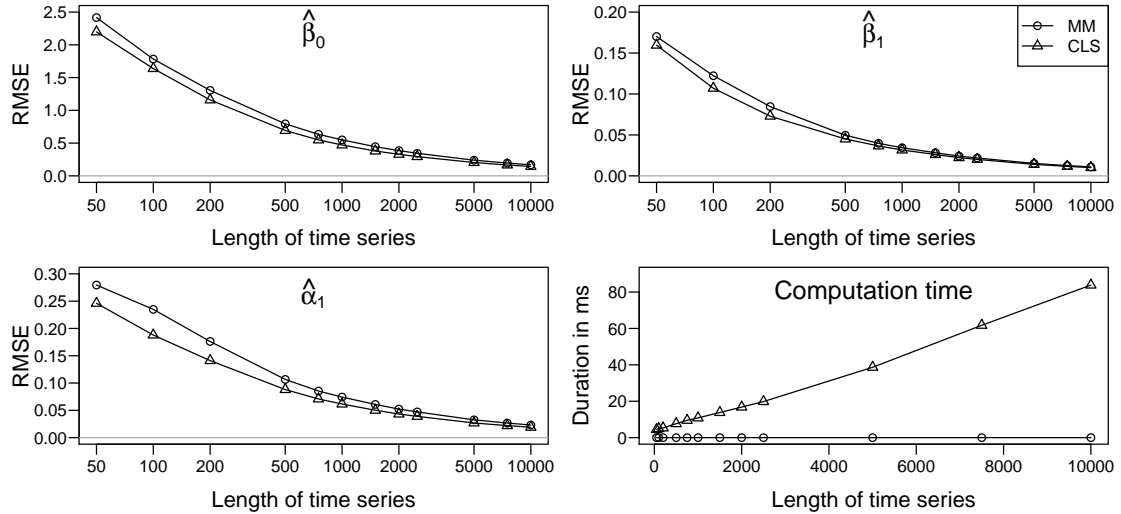


Figure 3: Root mean square error (RMSE) and computation time of different initial parameter estimators for a simulated intervention-free INGARCH(1,1) process averaged over 5000 repetitions.

where $\beta_i := 0$ for $i > p$, $\alpha_i := 0$ for $i > q$ and $\{\varepsilon_t\}$ is a white noise process. Given the autoregressive parameters ψ_i and the moving average parameters ϕ_i of the ARMA representation of $\{Y_t\}$, the INGARCH parameters are obtained by $\alpha_i = -\phi_i$ and $\beta_i = \psi_i + \phi_i$. We get β_0 from $\beta = \eta(1 - \sum_{i=1}^p \beta_i - \sum_{j=1}^q \alpha_j)$ using Equation (6). With these formulas INGARCH estimates $\hat{\beta}_0$, $\hat{\beta}_i$ and $\hat{\alpha}_i$ are obtained from ARMA estimates $\hat{\eta}$, $\hat{\psi}_i$ and $\hat{\phi}_i$. We compare two kinds of initial estimators, which differ in the way the ARMA parameters are estimated. The first kind uses a conditional least squares (CLS) fit, the second kind uses the method of moments (MM). For the latter we could only derive formulas for $p, q \leq 1$, for higher order models we use an ARMA(1,1) representation and set the initial estimates of higher order parameters to zero.

For the chosen constrained optimisation algorithm the initial estimation needs to lay in the interior of the parameter space. This is ensured by gentle modification of the initial estimation when necessary. We use the following procedure to obtain valid starting values β_0^* , β_i^* and α_i^* from parameter estimates $\hat{\beta}_0$, $\hat{\alpha}_i$ and $\hat{\beta}_i$, $i = 1, \dots, k$, found by an ARMA fit:

1. Set $\beta_i^* := \min\{\hat{\beta}_i, \varepsilon\}$ and $\alpha_i^* := \min\{\hat{\alpha}_i, \varepsilon\}$, $i = 1, \dots, k$.
2. If $\gamma := \sum_{i=1}^k \beta_i^* + \sum_{i=1}^k \alpha_i^* > 1 - \xi - \varepsilon$, shrink each β_i^* , α_i^* by the factor $(1 - \xi - \varepsilon)/\gamma$.
3. Set $\beta_0^* := \hat{\eta}(1 - \gamma)$, with $\hat{\eta}$ the estimated ARMA intercept.

A small constant $\varepsilon > 0$ ensures that the initial value lies inside the parameter space Θ and not on its boundaries. It is chosen to be $\varepsilon = 10^{-6}$ in our computations. The shrinkage factor in step 2 is chosen such that the sum of the parameters equals $1 - \xi - \varepsilon$ after shrinkage. The choice of β_0^* in step 3 ensures that the marginal mean remains unchanged for the modified starting values.

For the simulations in this work we use an INGARCH(1,1) process with parameters $\beta_0 = 3$, $\beta_1 = 0.4$ and $\alpha_1 = 0.3$, unless stated otherwise. Without an intervention this

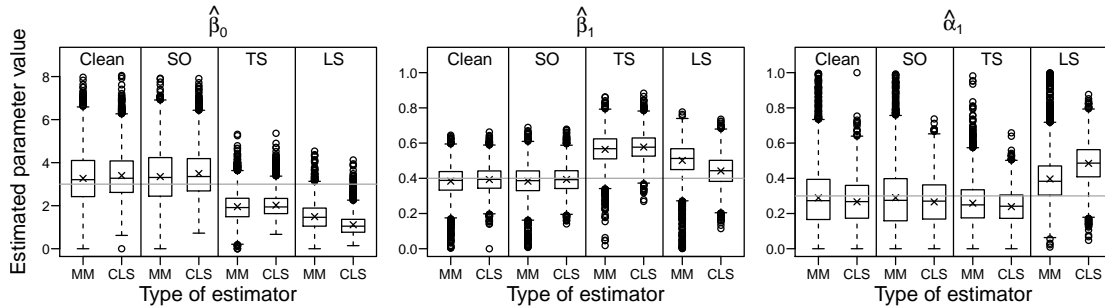


Figure 4: Initial parameter estimates of a simulated INGARCH(1,1) process with $n = 200$ observations and an intervention in the middle (Clean without intervention, SO with intervention of size $\nu = 24$, TS with $\nu = 18$, LS with $\nu = 3$). The true value is marked by a grey line, the sample mean by a cross. Simulation results are based on 5000 repetitions.

process has a marginal mean of 10 and a marginal variance of about 13.14 and thus exhibits moderate overdispersion. The autocorrelation function is 0.47 for lag 1 and decays to a value below 0.01 after 12 lags. Following Fokianos and Fried (2010) the parameter δ of a transient shift is set to 0.8 for simulation as well as for estimation, if not explicitly stated otherwise. All shown simulation results are based on 1000 repetitions, unless stated otherwise.

The results of the simulations shown in Figure 3 suggest that all considered initial parameter estimators are mean square consistent, although the convergence of the root mean square error to zero is fairly slow. Particularly for smaller sample sizes the CLS estimator has a considerably lower RMSE than the MM estimator.

The computation time grows linearly in the sample size for both estimators, but is much shorter and grows slower for the MM estimator than for the CLS estimator. However, even though the MM estimator is more than ten times faster for time series with 200 or more observations, the computation times of both initial estimators are negligible compared to the final CML estimation.

In this work we are particularly interested in the behaviour of the estimators in the presence of interventions. The simulation presented in Figure 4 shows that with a TS or LS both estimators become heavily biased, whilst the singular event of a SO does not really disturb them. If the time of possible interventions is known, the initial estimation could be done using only an intervention-free part of the time series. This approach is certainly not always possible, because in many applications the times of possible interventions are not known or the intervention-free part of the time series is too short for a sensible initial estimation. However, our simulations presented in Figures 5 and 6 show that the final CML estimation of β_0 , β_1 and α_1 works well, although the initial estimation is biased by an intervention. With respect to its behaviour in the presence of an intervention none of the two initial estimation methods seems to be superior.

Considering all simulation results we suggest CLS estimation for initialisation, because its finite sample properties are better than for MM estimation and its longer computation time is negligible compared to those for the final estimation. All following simulations find starting values for the final CML estimation by CLS estimation based on the complete time

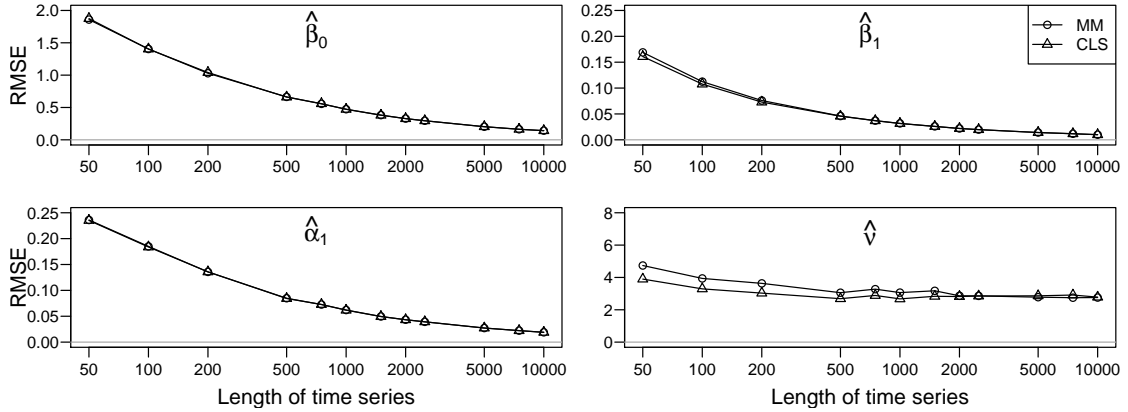


Figure 5: Root mean square error (RMSE) for the final CML estimator with different initial estimations (see legend) for a simulated INGARCH(1,1) process with TS of size $\nu = 9$ in the centre of $n = 200$ observations.

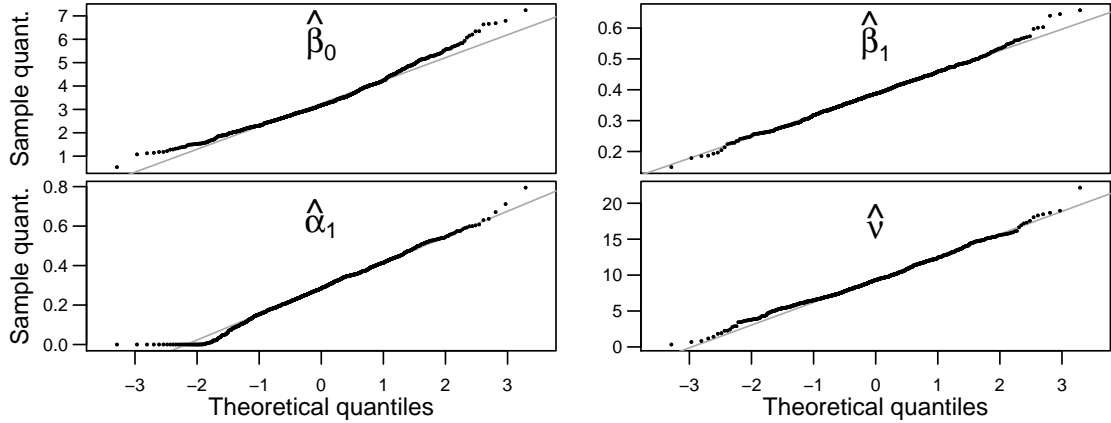


Figure 6: Normal QQ-plots of the CML estimation of an INGARCH(1,1) process according to model (ii) with a transient shift of size $\nu = 9$ in the center of $n = 200$ observations.

series.

3.2. Properties

Fokianos and Fried (2010) assume asymptotic normality of the CML estimator for their intervention model (i) and verify this by simulations. They do not provide a rigorous proof but outline the essential arguments necessary for it. We presume that the same also holds for the alternative intervention model (ii), more precisely that

$$\sqrt{n} \left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0 \right) \xrightarrow{d} N_{p+q+2} \left(\mathbf{0}, \mathcal{I}_{n\tau}(\hat{\boldsymbol{\theta}}_n)^{-1} \right)$$

for $n \rightarrow \infty$ and $\boldsymbol{\theta}_0$ the true parameter value. This would imply weak consistency of $\hat{\boldsymbol{\theta}}_n$, supposed that the information matrix converges in probability to a non-random limit.

Figure 6 and our simulations for plenty of other parameter settings and sample sizes support our conjecture of asymptotic normality, but expose that the convergence is rather

slow if the true parameter is close to the boundary of the parameter space. In this case the estimator is skewed for small sample sizes. In the example shown in Figure 6 there are many estimations of α_1 very close to its lower parameter constraint zero, which is an artefact of the constrained optimisation. This problem disappears if either the true α_1 is further away from the boundary of the parameter space Θ or the sample size n becomes larger and thus the variance of the estimations decreases. In this example with $\alpha_1 = 0.3$ not very close to zero $n = 500$ observations are enough to solve this problem, but in the same situation with $\alpha_1 = 0.1$ one would need several thousand observations. The same problem occurs for the estimation of $\beta_0 = 3$, where $n = 5000$ observations are needed until there is no serious deviation from normality in the QQ-plots any more. The described problem is not inherent in the external intervention model (ii), but arises in the same way for intervention model (i), as we have confirmed by simulations not shown here.

Another problem we discovered in our simulations occurs for situations with $\delta < 1$ only. In the case of a level shift (i.e. $\delta = 1$) all observations after time τ carry roughly the same amount of information about the unknown intervention size ν , whereas for $\delta < 1$ this amount becomes smaller for observations longer after the time of intervention τ . Hence, estimation of ν is difficult in those situations, which results in a log-likelihood function which is extremely flat in the direction of ν . In our simulations we observed in a considerable number of cases that the algorithm used for optimisation had not really moved away from its starting value, although the log-likelihood function was not even close to its maximum there. This happens even more often when the sample size is large. In most cases we can overcome this problem by choosing a very strict stopping rule for the BFGS algorithm, which of course leads to vastly longer computation times but reasonable results. The stopping criterion of the BFGS algorithm implemented in R is determined by the relative convergence tolerance, which is set to a value of 10^{-11} for our computations. This is a much lower value than its default, leading to a very strict stopping rule.

The final CML estimation of the INGARCH parameters β_0 , β_1 and α_1 turns out to be mean square consistent in our simulations under intervention model (ii), as illustrated in Figure 5 for the case of a TS. For a consistent estimation of the intervention size ν we found no considerable empirical evidence. With growing sample size the root mean square error decays only very slowly and is far away from zero even for a sample size of 10 000 observations. This is also the case in a simulation for intervention model (i) not shown here. For a LS with a value of $\delta = 1$ there is enough information about the parameter ν in the data so that the root mean square error converges to zero sufficiently fast in simulations for both models.

4. Testing for intervention effects

In this section we briefly present the procedures Fokianos and Fried (2010) proposed for the internal intervention model (i), which can be carried over to the external intervention model (ii) using the respective conditional score function and information matrix presented in Section 3. We verify by simulations similarly to those presented by Fokianos and Fried (2010) that these procedures are also valid for the external intervention model (ii).

$\tau =$		0.25n			0.5n			0.75n			Type
		1.0	5.0	10.0	Significance level			1.0	5.0	10.0	
		1.0	5.0	10.0	1.0	5.0	10.0	1.0	5.0	10.0	
$n =$	200	1.3	4.9	10.3	0.8	4.0	8.8	1.3	4.8	9.7	SO
	500	0.3	5.4	10.9	0.7	4.4	9.0	1.2	4.1	9.2	
	1000	0.7	3.3	8.1	1.5	6.4	11.1	0.7	4.4	8.2	
	5000	1.9	5.3	9.6	0.9	5.7	10.2	0.8	5.0	10.3	
	200	1.4	5.2	10.4	0.9	5.0	10.1	1.5	5.7	11.5	TS
	500	0.6	4.2	9.2	1.3	4.9	10.3	1.5	6.8	12.1	
	1000	0.5	4.5	10.1	0.7	6.0	10.0	0.9	4.9	10.8	
	5000	1.2	5.4	10.2	0.9	5.3	10.9	1.2	4.6	9.8	
	200	1.6	5.6	11.3	0.8	6.4	12.9	1.3	6.5	12.2	LS
	500	1.5	6.0	10.6	1.2	5.0	10.6	1.8	5.7	11.0	
	1000	0.7	4.5	9.1	0.8	5.1	8.7	0.9	5.6	10.1	
	5000	0.8	3.3	7.4	0.5	4.3	10.2	1.0	5.5	10.8	

Table 1: Size (in percent) of the test for an intervention of a given type at known time from model (ii) under the null hypothesis of no intervention. We varied the type of intervention we tested for, the position of the outlier τ and the sample size n .

4.1. Intervention of known type at known time

In a first step Fokianos and Fried (2010) give a test for the presence of an intervention effect of a given type (i.e. δ known and fixed) occurring at a known time τ . They propose a score test, because it only requires fitting under the null hypothesis of no intervention unlike likelihood ratio or Wald tests. This is desirable for situations with unknown type and time of intervention, where fitting for each type and time of intervention would be necessary otherwise. The null hypothesis $H_0 : \nu = 0$ is tested against the alternative $H_1 : \nu \neq 0$. Let $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\eta}}', 0)'$, with $\tilde{\boldsymbol{\eta}} = (\tilde{\beta}_0, \tilde{\beta}_1, \dots, \tilde{\beta}_p, \tilde{\alpha}_1, \dots, \tilde{\alpha}_q)'$ the CML estimator under the restricted, intervention-free model (2). The score test statistic is given by

$$T(\tau) = [S_{n\tau}(\tilde{\boldsymbol{\theta}}; \mathbf{Z})]' \mathcal{I}_{n\tau}^{-1}(\tilde{\boldsymbol{\theta}}) [S_{n\tau}(\tilde{\boldsymbol{\theta}}; \mathbf{Z})]. \quad (7)$$

Note that the score vector $S_{n\tau}$ and the information matrix $\mathcal{I}_{n\tau}$ depend on δ and τ , what can be seen from the corresponding equations in Section 3. Under the null hypothesis $H_0 : \nu = 0$ the test statistic (7) converges in distribution to a chi-square distribution with one degree of freedom, as $n \rightarrow \infty$, i.e.

$$T(\tau) \xrightarrow{d} \chi_1^2.$$

Hence, we obtain an asymptotic test for $H_0 : \nu = 0$ at a given significance level α , if we reject the null hypothesis when the observed value of T_τ exceeds the $(1 - \alpha)$ -quantile of the χ_1^2 -distribution.

We examine the finite sample behaviour of this test procedure for intervention model (ii) by simulations. Table 1 gives the averaged realised significance levels under the null hypothesis. These results resemble the ones for intervention model (i) by Fokianos and Fried (2010) (for a different parameter setting), and do not reveal any unusually large deviations

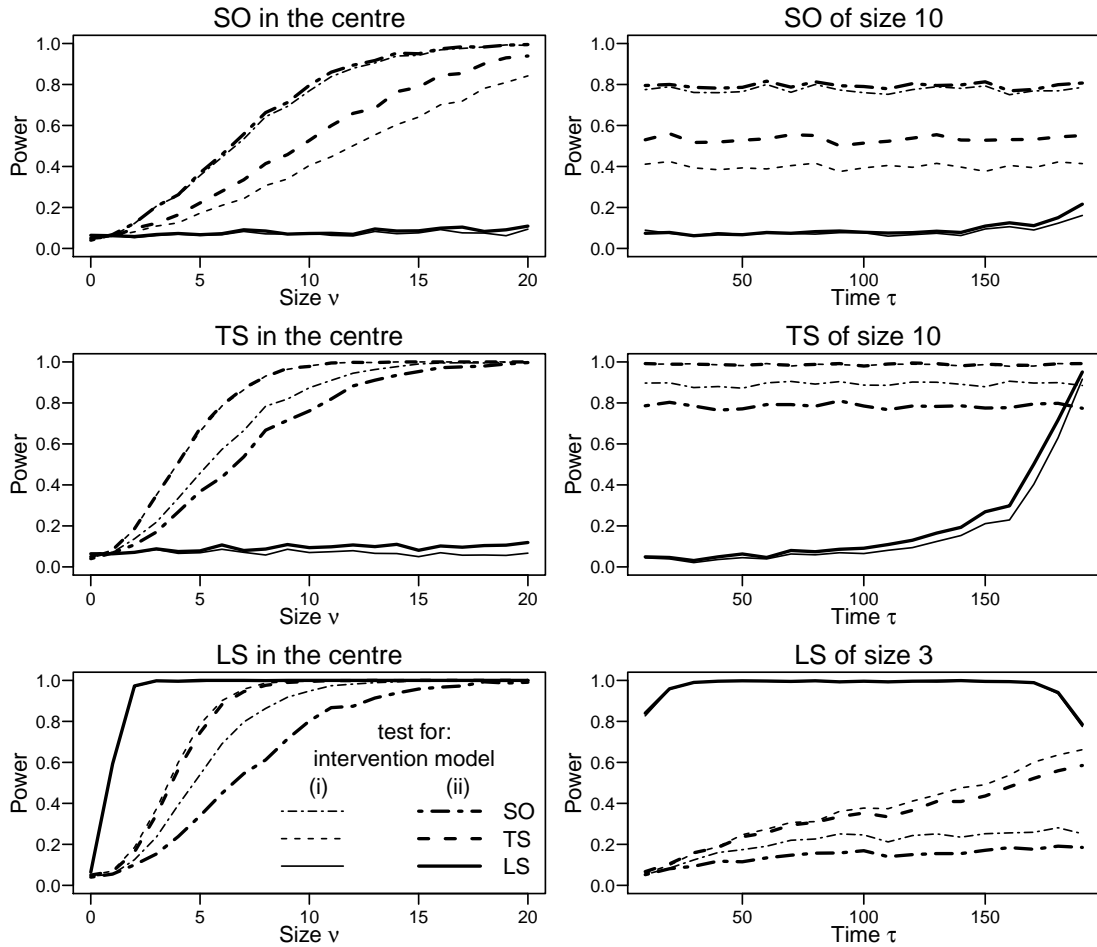


Figure 7: Power of the test for an intervention of a given type at known time from a given intervention model (see legend at the bottom). The data are simulated under the alternative of an intervention of type SO (top), TS (middle) and LS (bottom) from model (ii). In the left column the position of the intervention is fixed in the centre of the time series and the size ν of the intervention varies. In the right column the time τ of the intervention varies and the size of the intervention is fixed. The time series are of length $n = 200$.

of the achieved from the nominal significance levels 1%, 5% and 10% for various sample sizes, types and positions of interventions. This supports that the chi-square approximation of the test statistic is adequate also for intervention model (ii).

The bold lines in Figure 7 give the average simulated power of the test for an intervention following model (ii). From the left column we see that naturally interventions of larger size are detected more often than smaller ones. We also see that a level shift (even of a much lower size) is easier to detect than a transient shift, which is in turn easier to detect than a spiky outlier. The reason for this is that for larger δ more information about the intervention is carried by the observations following the intervention. In a situation with a certain type of intervention the test for this respective type of intervention has the highest power. This will be important for classifying the type of an intervention when it is not known, as it is

done in Section 4.3. The right column shows that for interventions at the end of a time series it is more difficult to distinguish a level shift from a transient shift or spiky outlier. This is because the time series cannot return to its undisturbed level any more. Note that for a level shift of the same size as the spiky outlier or transient shift this effect would be much more obvious than from the situation shown in our figures, where the LS has only less than a third of the size of SO and TS.

Again, our results resemble those of Fokianos and Fried (2010) for intervention model (i) and a different parameter setting. For better comparability we also run own simulations based on intervention model (i) for simulation and testing, using our standard parameter setting as before. Its results are shown as thin lines in Figure 8. These results are qualitatively the same, but interventions from model (i) are in general easier to detect than those from model (ii) when using the corresponding test. This can be explained by the stronger influence an intervention has on subsequent observations for intervention model (i) than for model (ii), which leads to a greater amount of information about the intervention contained in the data. However, as we are not primarily interested in models which are easy to estimate, but in ones which adequately describe real data situations, this does not disqualify the alternative intervention model (ii). We discuss the problem of (mis)specifying the intervention model in Section 4.4.

4.2. Intervention of known type at unknown time

For detection of an intervention of a given type occurring at unknown time Fokianos and Fried (2010) propose the test statistic

$$T^* = \max_{\tau \in D} T(\tau),$$

i.e. the maximum of the test statistics for an intervention at each possible time. Denote $\tau^* = \arg \max_{\tau \in D} T(\tau)$. Note that *a priori* knowledge about the time where an intervention might occur could be included by restricting the maximisation to a predefined smaller set D of values for τ , which is else chosen to be $D = \{2, \dots, n\}$. The p-value of a test for the hypothesis $H_0 : \nu = 0 \forall \tau \in D$ is approximated by the following parametric bootstrap procedure proposed by Fokianos and Fried (2010):

1. Generate B bootstrap replicates from an INGARCH model with parameter vector $\tilde{\theta}$, the estimation under H_0 defined in the previous section.
2. For each bootstrap replicate $b = 1, \dots, B$ compute the test statistic T_b^* .
3. The p-value is given by the number of bootstrapped test statistics exceeding the original test statistic, divided by $B + 1$.

Our simulations presented in Figure 9 (a) support that this is a valid testing procedure also for intervention model (ii). The p-values show a reasonable approximation to the uniform distribution, the achieved significance levels do not seem to deviate systematically from the respective nominal significance levels. The positions of the erroneously detected interventions in Figure 9 (b) show no peculiarities for the test on SO and TS. The test on a

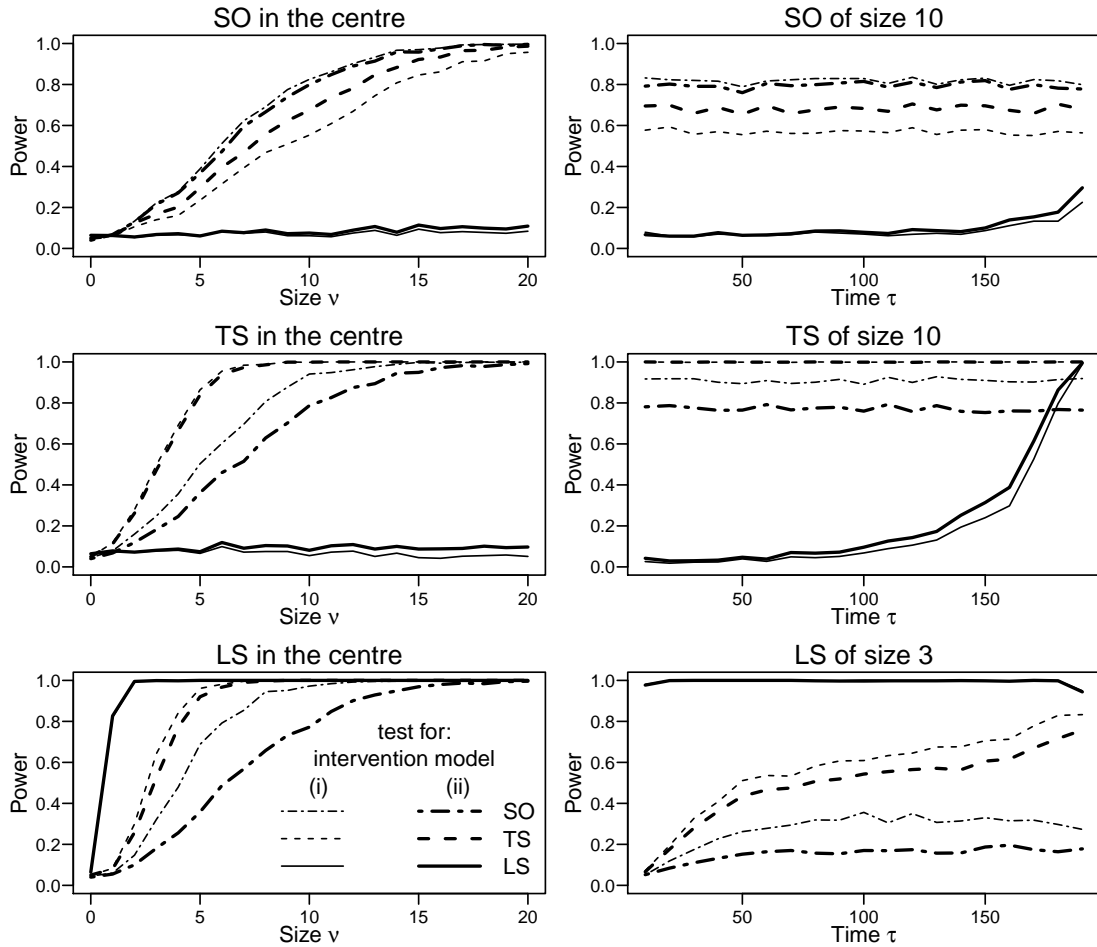


Figure 8: The same simulations as shown in Figure 7 but with interventions simulated from model (i) instead of model (ii).

LS detects more interventions at the beginning and at the end of the time series, where one randomly occurring extreme observation can hardly be distinguished from a LS.

To investigate the power of the test we consider the simulations shown in Figure 10 (a). A test on a SO needs a size of about $\nu = 15$ to reject a simulated SO in more than 50% of the cases (see left segment in the left plot). For a test on a LS a size of $\nu = 2$ of a simulated LS is needed for this (see right segment in the right plot). The timing of the intervention shown in Figure 10 (b) is very accurate for SO and TS (over 85% of the cases are detected correctly) and a little bit less accurate for the LS (about 40% detected correctly, the other cases are up to ten observations away from the true value).

Again our simulations on size and power of the test for intervention model (ii) are similar to those of Fokianos and Fried (2010) for intervention model (i).

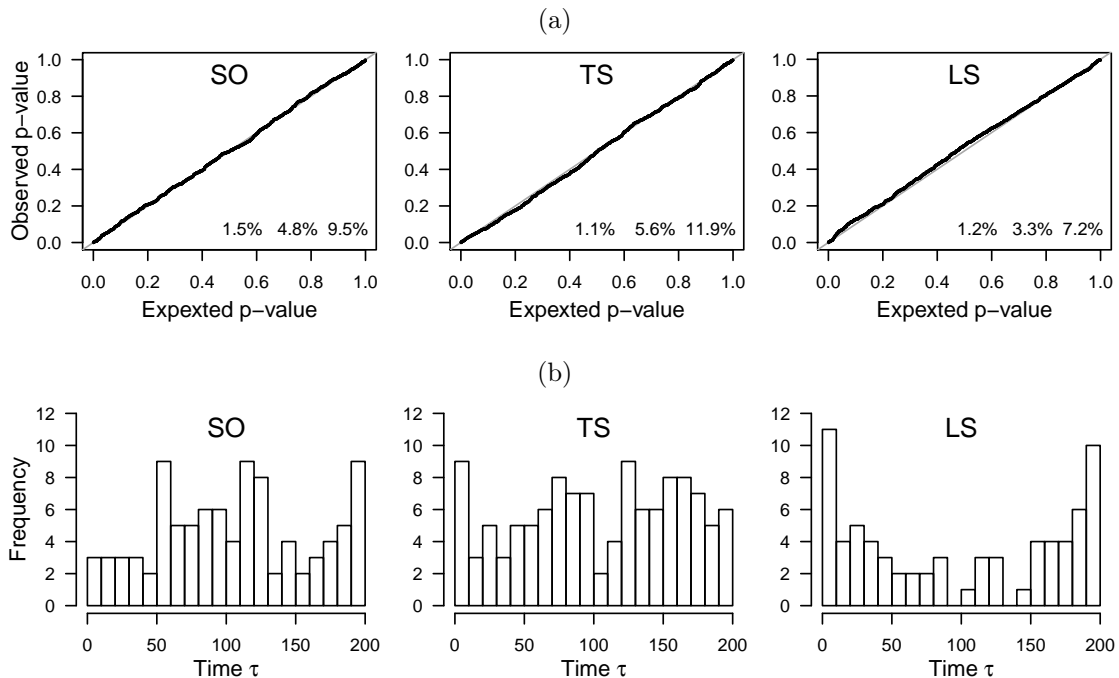


Figure 9: (a) Size of the test for an intervention of given type (see title) at unknown time from intervention model (ii) using $B = 500$ bootstrap samples. The data are simulated under the null hypothesis of no intervention. A value above the bisecting line indicates that the test is conservative for this particular size. The values given at the bottom are the achieved significance levels for 1%, 5% and 10% respectively. The time series are of length $n = 200$.
(b) Detected times of intervention τ^* for tests with a p-value below 10% in the simulations above.

4.3. Multiple interventions of unknown type at unknown time

A common situation in practical applications is that not only the position but also the type of intervention is unknown. From Figure 10 (a) we get an idea that it is possible to classify the type of an intervention by the minimum p-value of all (three) tests. On average the p-values of the tests for the matching type are lower than the p-values of the tests on the other types. Of course more than three values of δ could be considered, but it would be certainly hard to distinguish between more than three types. In case of equal p-values Fokianos and Fried (2010) recommend to prefer the type with the highest value of δ .

Furthermore, there might be more than one intervention in a time series. Fokianos and Fried (2010) propose a stepwise procedure to detect, classify and eliminate multiple intervention effects in a time series. The elimination of intervention effects is based on the decomposition presented in Section 2. As for the other tests we can adopt their iterative procedure for the alternative intervention model (ii) by using the modified equations from Section 2. We refer to their paper for details and just show some simulation results for the alternative intervention model (ii).

We applied the iterative detection procedure to simulated time series with two transient shifts. The first intervention at time 50 is of size $\nu = 12$ and has a value of $\delta = 0.7$, the second one at time 151 has $\nu = 9$ and $\delta = 0.9$. Each of the interventions is found in

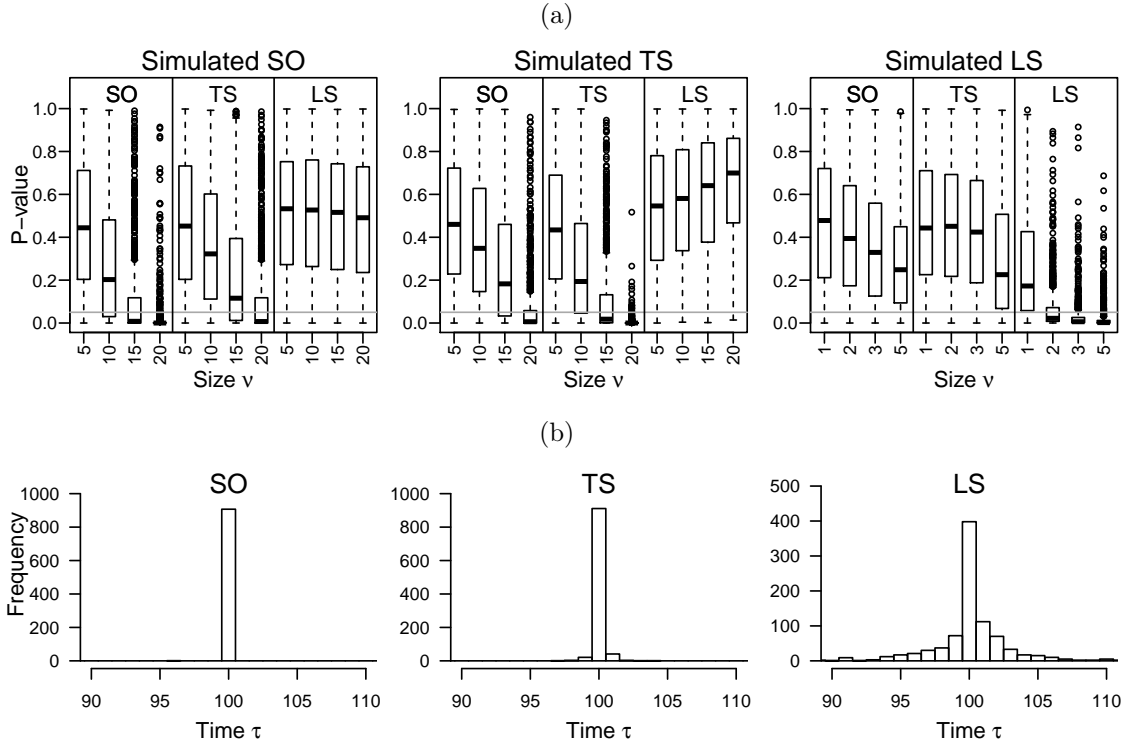


Figure 10: (a) Power of the test for an intervention of given type (see segments within each plot) at unknown time from intervention model (ii) using $B = 500$ bootstrap samples. The data are simulated under the alternative of an intervention of given type (see title) from model (ii) in the centre of a time series of length $n = 200$. We vary the size of the intervention (see horizontal axis).

(b) Detected times of intervention τ^* for tests with a p-value below 10% in the simulations above. The simulated intervention is of the same type which is also tested for (see title) and is of size 20 for SO and TS and of size 5 for LS.

about 60% of the 1000 repetitions. In 38% of the cases we found both interventions, in about 33% we only found both and no others. The detected times are shown in Figure 11. Although we tested for interventions with $\delta = 0.8$ and thus slightly misspecified their type, our results are satisfactory. There are no systematic simulations of Fokianos and Fried (2010) for intervention model (i), but their simulated examples point in the same direction.

4.4. Misspecification

Up to now, we introduced an alternative intervention model to the one proposed by Fokianos and Fried (2010). We found that the alternative intervention model (ii) is a bit harder to estimate. However, it is not primarily important that a model fits to simulated data from exactly this model but that it also fits to real data. For a given time series with intervention effects it is presumably a too ambitious task to find out which intervention model is more adequate, as both differ only little. Nevertheless it is interesting to study the effect of misspecification of the intervention model.

Consider the situation with an intervention of a given type at a given time. We look at the thin lines in Figure 7 for the test on an intervention of model (i) and compare them with

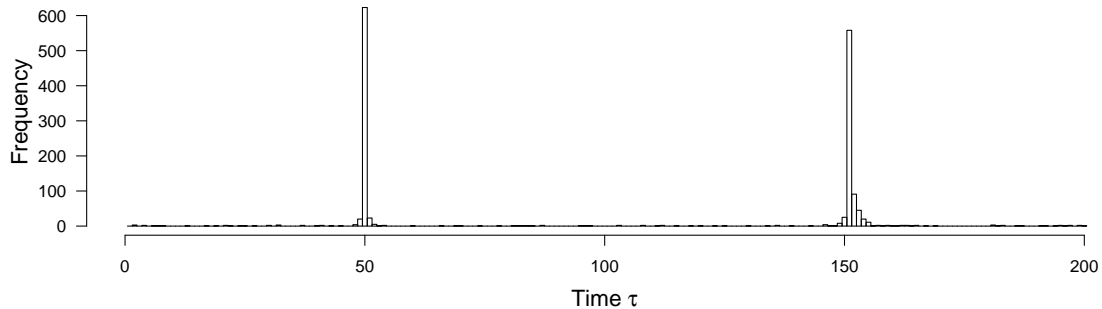


Figure 11: Times of interventions found by the iterative detection procedure for model (ii) using $B = 500$ bootstrap samples. The data are simulated with two transient shifts from model (ii) at time 50 ($\delta = 0.7$, size $\nu = 12$) and at time 151 ($\delta = 0.9$, size $\nu = 9$). The time series are of length $n = 200$.

the thick ones of the same kind, which stand for the test based on model (ii) from which the data are simulated. The test using a different model as for simulation has an only slightly lower power. The same is true for the converse situation as shown in Figure 8. Hence, there is some robustness against misspecification of the intervention model.

Another interesting finding from Figures 7 and 8 is that for example a transient shift from model (ii) is somewhat similar to a spiky outlier from model (i). In general an intervention effect from model (ii) resembles one from model (i) with a slightly lower value of δ and vice versa. This is in line with the explanations given in Section 2.

5. Real data application

We study a time series of campylobacteriosis infection cases first studied by Ferland et al. (2006). An infection with the *Campylobacter* bacterium may be caused mainly by contaminated aliment but also by contact with infected animals. Human to human spread is possible particularly for infants. The incubation period is typically 2-5 days. Infected patients often do not show any symptoms and are potentially infectious via their excrement for 2-4 weeks on average. The number of cases is higher in warm seasons. The original time series with $n = 140$ is shown in grey in Figure 12. Ferland et al. (2006) fit an INGARCH(1,0) model with an additional autoregression on the conditional mean for lag 13 to account for seasonality. For simplicity we consider an INGARCH(1,1) model without seasonality instead. We do not have any prior information on the occurrence of possible interventions, so we apply the iterative detection procedure from Section 4.3 to the data.

We first search for interventions following intervention model (ii). In the first step, the bootstrapped p-values for all considered types (SO, TS with $\delta = 0.8$ and LS) are zero. The procedure decides in favour of the LS with an estimated size of 4.600 found at time 84, because it has the highest value of δ . The detected intervention effect is removed from the time series. In a second step, both the test on a SO and on a TS are significant at time 100, the same position already found in the first step for these types of intervention. We decide in favour of the SO with an estimated size of 16.360, because it has a lower p-value than the TS. Again, the time series is cleaned by the detected intervention effect. In a third step, no

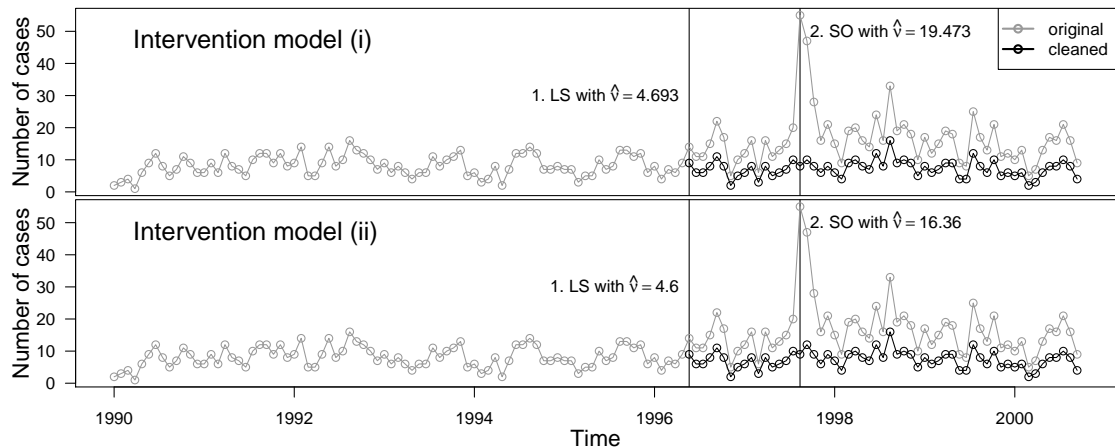


Figure 12: Number of campylobacterosis infections (reported every 28 days) in the north of Québec in Canada. The original time series is displayed in grey, the modified time series cleaned by the procedure from Section 4.3 in black.

further interventions with a p-value lower than 5% are detected. The estimated INGARCH parameters for the cleaned time series are $\hat{\beta}_0 = 4.102$, $\hat{\beta}_1 = 0.373$ and $\hat{\alpha}_1 = 0.098$. The cleaned time series and both interventions are shown in Figure 12 (bottom).

We also applied this procedure for intervention model (i) and found the same types of interventions at the same times with only slightly different estimated intervention sizes. These results are presented in Figure 12 (top). The average of the squared Pearson residuals for this model is 1.006, compared to 1.015 for intervention model (ii). Hence the internal intervention model (i) fits in this sense a little better to the data than the external one.

6. Discussion

In this work we study a new variant of an intervention model for INGARCH processes and adopt a maximum likelihood approach for detection and estimation of intervention effects proposed by Fokianos and Fried (2010) to this new model. Simulations support that the presented procedures work quite reliably. In comparison to the intervention model studied by Fokianos and Fried (2010) the new model has less influence on subsequent observations after the occurrence of an intervention. An intervention concentrates its effect more on the time of occurrence and influences the future less than in the model of Fokianos and Fried (2010), since e.g. a spiky outlier can be interpreted as a change of the observation but not the underlying mean at the time of its occurrence, see Equation (ii).

We found some robustness against misspecification of the intervention model. Thus one can expect a reasonable performance for this kind of intervention effects applying either of the two models. On the other hand this finding suggests that a test to discriminate both intervention models could only have a very low power and is probably not worthwhile.

One undesirable feature of the procedures for detection of interventions are the high computational costs. For the example in the previous section with only $n = 140$ observations each step of the iterative detection procedure runs about 39 minutes for intervention

model (i) and even about 52 minutes for model (ii) on a single processor unit (Intel Xeon CPU with 2.83 GHz). The reason for the long computation time is the parametric bootstrap. Fokianos and Fried (2012) reduced the computation time of the bootstrap considerably by not estimating the parameter vector for each bootstrap replicate but using the true value used for its generation instead. They show that this modification results in quite conservative procedures. With parallelised computation of the bootstrap we have greatly shortened its duration using multiple processors. This makes our approach feasible for moderately long time series but not yet for very long time series.

Acknowledgements

The financial support of the Deutsche Forschungsgemeinschaft (SFB 823, “Statistical modelling of nonlinear dynamic processes”) is gratefully acknowledged. The authors thank Professor A. Latour for kindly providing the campylobacteriosis data.

References

- Abraham, B. and Box, G. E. P. (1979). Bayesian analysis of some outlier problems in time series. *Biometrika*, 66(2):229–236.
- Box, G. E. P. and Tiao, G. C. (1975). Intervention Analysis with Applications to Economic and Environmental Problems. *Journal of the American Statistical Association*, 70(349):70–79.
- Cox, D. R. (1981). Statistical Analysis of Time Series: Some Recent Developments. *Scandinavian Journal of Statistics*, 8(2):93–115.
- 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., Rahbek, A., and Tjøstheim, D. (2009). Poisson Autoregression. *Journal of the American Statistical Association*, 104(488):1430–1439.
- Fried, R., Aguesop, I., Bornkamp, B., Fokianos, K., Fruth, J., and Ickstadt, Katja, K. (2012). Bayesian outlier detection in INGARCH time series. *TU Dortmund, SFB 823 Discussion Paper*, 26/12. <http://hdl.handle.net/2003/29508>.
- Heinen, A. (2003). Modelling Time Series Count Data: An Autoregressive Conditional Poisson Model. *CORE discussion paper*, 62. <http://dx.doi.org/10.2139/ssrn.1117187>.
- Kedem, B. and Fokianos, K. (2002). *Regression models for time series analysis*. Wiley series in probability and statistics. Wiley-Interscience, Hoboken.
- Lange, K. (1999). *Numerical analysis for statisticians*. Statistics and computing. Springer, New York.
- R Development Core Team (2012). R – A Language and Environment for Statistical Computing. <http://www.r-project.org>.

