Optimal Alarm Systems for FIAPARCH Processes

Conceição Costa, Manuel G. Scotto* and Isabel Pereira

Departamento de Matemática

Research Unit 'Matemática e Aplicações'

Universidade de Aveiro

Campo Universitário de Santiago

3810-193 Aveiro, Portugal

Abstract

In this work, an optimal alarm system is developed to predict whether a financial time series modeled via Fractionally Integrated Asymmetric Power ARCH (FIAPARCH) models, up/downcrosses some particular level and give an alarm whenever this crossing is predicted. The paper presents classical and Bayesian methodology for producing optimal alarm systems. Both methodologies are illustrated and their performance compared through a simulation study. The work finishes with an empirical application to a set of data concerning daily returns of the S. Paulo Stock Market.

Keywords: FIAPARCH processes, Optimal alarm systems, Econometrics

*Manuel G. Scotto, Departamento de Matemática, Universidade de Aveiro. Campo Universitário de Santiago, 3810-193 Aveiro, Portugal. Tel.: +351 234370670; fax: +351 234370670. Email: mscotto@ua.pt.

1 Introduction

Recently, it has been recognized the potential of optimal alarm systems in detecting and warning the occurrence of catastrophes or some other related events; see for example Monteiro et al. (2008) and the references therein. Conceptually, the simplest way of constructing an alarm system is to predict X_{t+h} by a predictor say, $\hat{X}_{t+h}$, which is usually chosen so that the mean square error is minimized, providing

$$\hat{X}_{t+h} = E[X_{t+h}|X_s, -\infty < s \leq t].$$

An alarm is given every time the predictor exceeds some critical level. This alarm system, however, does not have a good performance on the ability to detect the events, locate them accurately in time and give as few false alarms as possible. Lindgren (1980,1985) and de Maré (1980) set the principles for the construction of optimal alarm systems and obtained some basic results regarding the optimal prediction of level crossings. Svensson et al. (1996) applied these principles in the prediction of level crossings in the sea levels of the Baltic sea. It is worth to mention that the alarm system introduced by Lindgren and de Maré, ignores the sampling variation of the model parameters. Given heed to this issue, Amaral Turkman and Turkman (1990) suggested a Bayesian approach and particular calculations were carried out for an autoregressive model of order one. Further extensions and generalizations were proposed by Antunes et al. (2003) and more recently by Monteiro et al. (2008).

The spectrum of applications of optimal alarm systems is wide and yet to be explored. One major area of applications is environmental economics. Atmospheric concentrations of air pollutants like ozone, carbon monoxide or sulfur dioxide constitute time series that can be analyzed under the perspective of the upcrossings of some critical levels, usually related with public health (e.g., Smith et al. 2000; Koop and Tale 2004; Tobias and Scotto 2005). Another area of potential applications is econometrics and in particular in risk management. In this case, the implementation of probabilistic models for the assessment of market risks or credit risks is mandatory. Examples can be found in the forecasting of financial risk of lending to costumers

(Thomas 2000), the arrivals forecast of guests at hotels (Weatherford and Kimes 2003) and in forecasting daily stock volatility, which has direct implications in option pricing, asset allocation or value-at-risk (Fuentes et al. 2009). All the above referred references, however, are not directly applicable to calculate in advance the probability of future up/downcrossings. Is in this context that the implementation of an alarm system reveals to be useful. A related interesting problem, which has not been addressed yet, is to develop optimal alarm systems for financial time series. This article aims to give a contribution towards this direction.

The analysis of financial time series like log-return series of foreign exchange rates, stock indices or share prices, has revealed some common features: sample means not significantly different from zero, sample variances of the order 10^{-4} or smaller and sample distributions roughly symmetric in its center, sharply peaked around zero but with a tendency to negative asymmetry. In particular, it has usually been found that the conditional volatility of stocks responds asymmetrically to positive versus negative shocks: volatility tends to rise higher in response to negative shocks as opposed to positive shocks, which is known as the leverage effect. Another common feature of series of log-returns is that the sample autocorrelation function is negligible at all lags, (except perhaps for the first) but the sample autocorrelation functions for the absolute values or the squares of the log-returns are different from zero for a large number of lags and stay almost constant and positive for large lags. This last feature, is known, in this context, as long memory or long range dependency. Several models have been proposed ir order to describe this stylized facts about log-return series. We mention here the ARCH models, introduced by Engle (1982) and some of the subsequent generalizations: GARCH, (Bollerslev 1986), EGARCH (Dellaportas et al. 2000), APARCH (Ding et al. 1993), FIGARCH (Baillie et al. 1996) and FIAPARCH (Tse 1998). Some of these models will be later addressed in Section 2. For a survey of ARCH-type models see Teräsvirta (2009).

The rest of the paper is organized as follows: in Section 2, basic theoretical concepts related to optimal alarm systems are presented. Furthermore, an optimal alarm system for FIAPARCH processes is implemented. Expressions for some alarm characteristics of the alarm system are given. Estimation of the model FIAPARCH$(1, d, 1)$ by classical and Bayesian methodology is covered in Section 3. In Section 4, the results are illustrated through a simulation study. A real-data example is given in Section 5.

2 Optimal alarm systems and their application to FIAPARCH processes

Let $\{X_t,\, t \geq 1\}$ be a discrete time stochastic process. The time sequel $\{1, 2, \ldots, t-1, t, t+1, \ldots\}$ is divided into three sections, namely the data or informative experience, $D_t = \{X_1, X_2, \ldots, X_{t-q}\}$, the present experiment, $\mathbf{X}_2 = \{X_{t-q+1}, \ldots, X_t\}$ and the future experiment, $\mathbf{X}_3 = \{X_{t+1}, X_{t+2}, \ldots\}$. Any event of interest, say $C_{t,j}$, in the σ-field generated by $\boldsymbol{X}_3$ is defined as a catastrophe. Until further notice a catastrophe will be considered as the upcrossing event $C_{t,j} = \{X_{t+j-1} \leq u < X_{t+j}\}$, for some $j \in \mathbb{N}$. Moreover, any event $A_{t,j}$ in the σ-field generated by $\mathbf{X_2}$, predictor of $C_{t,j}$, will be an alarm region. It is said that an alarm is given at time t, for the catastrophe $C_{t,j}$, if the observed value of $\mathbf{X}_2$ belongs to the alarm region. In addition, the alarm is said to be correct if the event $A_{t,j}$ is followed by the event $C_{t,j}$. Conversely, a false alarm is defined as the occurrence of $A_{t,j}$ without $C_{t,j}$. If an alarm is given when the catastrophe occurs, it is said that the catastrophe is detected. Furthermore, the alarm region $A_{t,j}$ is said to have size $\alpha_{t,j}$ if $\alpha_{t,j} = P(A_{t,j}|D_t)$. The alarm region is optimal of size $\alpha_{t,j}$ if

$$P(A_{t,j}|C_{t,j}, D_t) = \sup_{B \in \sigma_{\mathbf{X_2}}} P(B|C_{t,j}, D_t), \tag{1}$$

with $P(B|D_t) = \alpha_{t,j}$.

Definition 2.1. *An optimal alarm system of size $\{\alpha_{t,j}\}$ is a family of alarm regions $\{A_{t,j}\}$ in time, satisfying (1).*

Lemma 2.1. *The alarm system $\{A_{t,j}\}$ with alarm region given by*

$$A_{t,j} = \{\mathbf{x}_2 \in \mathbb{R}^q : P(C_{t,j}|\mathbf{x}_2, D_t) \geq k_{t,j} P(C_{t,j}|D_t)\},$$

for a fixed $k_{t,j} : P(\boldsymbol{X}_2 \in A_{t,j}|D_t) = \alpha_{t,j}$, *is optimal of size* $\alpha_{t,j}$.

This lemma ensures that the alarm region defined above renders the highest detection probability. Moreover to enhance the fact that the optimal alarm system depends on the choice of $k_{t,j}$, it is important to stress that in view of the fact that $P(C_{t,j}|D_t)$ does not depend on $\mathbf{x}_2$, the alarm region can be rewritten in the form

$$A_{t,j} = \{\mathbf{x}_2 \in \mathbb{N}^q : P(C_{t,j}|\mathbf{x}_2, D_t) \geq k\}, \tag{2}$$

where $k = k_{t,j}P(C_{t,j}|D_t)$ is chosen in some optimal way to accommodate conditions over the following operating characteristics of the alarm system.

Definition 2.2 (Operating characteristics)**.**

1. $P(A_{t,j}|D_t)$ - *Alarm size*
2. $P(C_{t,j}|A_{t,j}, D_t)$ - *Probability of correct alarm*
3. $P(A_{t,j}|C_{t,j}, D_t)$ - *Probability of detecting the event*
4. $P(\overline{C_{t,j}}|A_{t,j}, D_t)$ - *Probability of false alarm*
5. $P(\overline{A_{t,j}}|C_{t,j}, D_t)$ - *Probability of undetected event*

Most models for financial time series used in practice are given in the multiplicative form

$$X_t = \sigma_t Z_t, \tag{3}$$

where $\{Z_t\}$ forms an i.i.d. sequence with zero mean and unit variance, $\{\sigma_t\}$ is a stochastic process such that σ_t and Z_t are independent for fixed t. In general, $\{\sigma_t\}$ and $\{X_t\}$ are assumed to be strictly stationary. Motivation for considering this particular choice of simple multiplicative model comes from the fact that (a) in practice, the direction of price changes is well modeled by the sign of Z_t, whereas σ_t provides a good description of the order of magnitude of this change; and (b) the volatility σ_t^2 represents the conditional variance of X_t given σ_t. This representation expresses the belief that the direction of price changes can not be modeled, only their magnitude (e.g., Mikosch 2003).

Engle (1982) considered the following model for the volatility σ_t in (3)

$$\sigma_t^2 \;=\; \omega + \sum_{i=1}^{p} \alpha_i X_{t-i}^2,$$

where $\omega > 0$ and $\alpha_i \geq 0$, for $i = 1, \ldots, p$. Bollerslev (1986) suggested a generalization of this model leading to the Generalized ARCH model of order (p, q)

$$\sigma_t^2 \;=\; \omega + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

with $\omega > 0$ and $\alpha_i \geq 0$, $\beta_j \geq 0$, for $i = 1, \ldots, p$ and $j = 1, \ldots, q$. Ding et al. (1993) proposed the Asymmetric Power ARCH of order (p, q), in short APARCH(p, q), model defined as

$$\sigma_t^\delta \;=\; \omega + \sum_{i=1}^{p} \alpha_i (|X_{t-i}| - \gamma_i X_{t-i})^\delta + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^\delta,$$

where $\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $\delta \geq 0$ and $-1 < \gamma_i < 1$. This model allows to detect asymmetric responses of the volatility for positive or negative shocks. If $\gamma_i > 0$, negative shocks have stronger impact on volatility than positive shocks, as would be expected in the analysis of financial time series. If $\gamma_i < 0$, the reverse happens. Baillie et al. (1996) proposed the FIGARCH(p, d, q) model in order to accommodate long memory in volatility (accordingly to the most common definition of long memory: autocovariance function, $\gamma(k)$, decaying at the hypergeometric rate k^{2d-1}, with $0 < d < 0.5$). They started by writing the GARCH(p, q) process as an ARMA(m, p) one in X_t^2

$$(1 - \alpha(B) - \beta(B))X_t^2 = \omega + (1 - \beta(B))\nu_t,$$

where $m = \max\{p, q\}$ and $\nu_t = X_t^2 - \sigma_t^2$. When the autoregressive lag polynomial $1 - \alpha(B) - \beta(B)$ contains a unit root, the GARCH(p, q) process is said to be integrated in variance (Engle and Bollerslev 1986). The Integrated GARCH(p, q) or IGARCH(p, q) class of models is given by

$$\phi(B)(1 - B)X_t^2 = \omega + (1 - \beta(B))\nu_t.$$

The Fractionally Integrated GARCH(p, d, q), or FIGARCH(p, d, q) class of models is simply obtained by allowing the differencing operator in the above equation to take non-integer values:

$$\phi(B)(1-B)^d X_t^2 = \omega + (1-\beta(B))\nu_t,$$

with $\beta(B)$ and $\phi(B)$ representing lag polynomials of order p and q, respectively, and the roots of $\phi(z) = 0$ lying outside the unit circle. d is the fractional differencing parameter and the fractional differencing operator is defined by its Maclaurin series expansion,

$$(1-B)^d = 1 - dB - \frac{d(1-d)}{2!}B^2 - \frac{d(1-d)(2-d)}{3!}B^3 - \cdots,$$

with $0 < d < 1$. The FIGARCH(p, d, q) model can be expressed as an ARCH(∞)-process with

$$\sigma_t^2 = \frac{\omega}{1-\beta(B)} + \lambda(B)X_t^2,$$

where $\lambda(B) = \dfrac{\omega}{1-\beta(B)} + [1-(1-\beta(B))^{-1}\phi(B)(1-B)^d]$. For the FIGARCH$(p, d, q)$ model to be well-defined and the conditional variance positive almost surely for all t, all the coefficients in the ARCH(∞) representation must be non-negative. General conditions, however, are difficult to establish. For the FIGARCH$(1, d, 1)$ model, the infinite series coefficients can be obtained recursively [1] and from this recursions it was shown by Bollerslev and Mikkelsen (1996) that the conditions

$$\beta - d \leq \phi \leq \frac{2-d}{3}, \qquad d(\phi - \frac{1-d}{2}) \leq \beta(\phi - \beta + d), \tag{4}$$

are sufficient to ensure non-negativity. In the covariance stationary GARCH(p, q) model, shocks to the conditional variance dissipate exponentially, meaning that the effect of a shock on the forecast of the future conditional variance tends to zero at a fast exponential rate. In the IGARCH(p, q) model, shocks to the conditional variance persist indefinitely, meaning that the shocks remain important for all horizon forecasts. In the FIGARCH(p, d, q) model, the differencing parameter introduces a different behavior: the effect of a shock to the forecast of the future conditional variance is expected to die out at a slow hyperbolic rate. This is the reason why the

[1] With the following expressions: $\lambda_1 = \phi - \beta + d$, $\lambda_i = \beta\lambda_{i-1} + [\frac{i-1-d}{i} - \phi]\delta_{i-1}$ with $\delta_1 = d$ and $\delta_i = \delta_{i-1}\frac{i-1-d}{i}$ for $i \geq 2$.

FIGARCH(p, d, q) process is said to have long memory in volatility. This statement was first proved by Baillie et al (1996) for the case FIGARCH$(1, d, 0)$ model. The statistical properties of the general FIGARCH(p, d, q) process, however, remain unestablished. Namely, stationarity is not a certainty as well as the source of long memory on volatility or even its existence are nowadays controversial.

Tse (1998) modifies the FIGARCH(p, d, q) process to allow for asymmetries, thus originating the Fractionally Integrated Asymmetric Power ARCH of order (p, q), the FIAPARCH(p, d, q) process. Defining $g(X_t) = (|X_t| - \gamma_i X_t)^\delta$ the FIAPARCH(p, d, q) model can be written as

$$\begin{aligned} X_t &= \sigma_t Z_t \\ \sigma_t^\delta &= \frac{\omega}{1-\beta(B)} + [1 - (1-\beta(B))^{-1}\phi(B)(1-B)^d]g(X_t), \end{aligned}$$

with d being the fractional differencing parameter. If $0 < d < 0.5$, long memory is expected to occur. It is very interesting to note that the FIAPARCH representation nests two major classes of ARCH models: the APARCH and the FIGARCH models. When $d = 0$ the process reduces to the APARCH(p, q) model, whereas for $\gamma = 0$ and $\delta = 2$ the process reduces to the FIGARCH(p, d, q) model. The FIGARCH representation includes the GARCH (when $d = 0$) and the IGARCH (when $d = 1$) with the implications in terms of impact of a shock on the forecasts of future conditional variances, as discussed above. Considering all the features involved in this specification, Conrad et al. (2008) point out some advantages of the FIAPARCH(p, d, q) class of models, namely (a) it allows for an asymmetric response of volatility to positive and negative shocks, so being able to traduce the leverage effect, (b) in this particular class of models it is the data that determines the power of returns for which the predictable structure in the volatility pattern is the strongest, and (c) the models are able to accommodate long memory in volatility, depending on the differencing parameter d.

In the case when both $1 - \beta(B)$ and $\phi(B)$ are polynomials of degree 1 and we allow that

$\beta(B) = \beta B$ and $\phi(B) = 1 - \phi B$, the volatility, σ_t, in the FIAPARCH$(1, d, 1)$ takes the form

$$\sigma_t^\delta = \frac{\omega}{1-\beta} + [1 - (1 - \beta B)^{-1}(1 - \phi B)(1 - B)^d](|X_t| - \gamma X_t)^\delta,$$

with

$$\omega > 0, \quad \beta \geq 0, \quad \phi \geq 0, \quad -1 < \gamma < 1, \quad \delta \geq 0, \quad 0 \leq d \leq 1. \tag{5}$$

As previously stated, in the FIGARCH$(1, d, 1)$ model of Baillie et al. (1996) the conditional volatility has an infinite series representation in terms of X_t^2. In the FIAPARCH$(1, d, 1)$ model, X_t^2 is replaced by $g(X_t)$. Nevertheless, the coefficients λ_i remain unaltered. This property leads Tse (1998) to conclude that the effects of the past residuals on the future conditional volatility show the same hyperbolic decay as Baillie et al. (1996) found in the FIGARCH model. Also referring Baillie's conclusions about the FIGARCH process being strictly stationary and ergodic, Tse leaves this issue as an open question for the FIAPARCH process.

The application of the alarm system to the FIAPARCH$(1, d, 1)$ model will be done for the particular case $q = 1$ and $j = 2$ in Lemma 2.1. The event of interest (i.e., the catastrophe) is defined as the upcrossing of some fixed level u two steps ahead, that is

$$C_{t,2} = \{X_{t+1} \leq u < X_{t+2}\}. \tag{6}$$

The alarm region of optimal size $\alpha_{t,2}$ is given by

$$A_{t,2} = \left\{ x_t \in \mathbb{R} : \frac{P(C_{t,2}|x_t, D_t)}{P(C_{t,2}|D_t)} \geq k_{t,2} \right\}.$$

Writing $k = k_{t,2}P(C_{t,2}|D_t)$,

$$A_{t,2} = \{x_t \in \mathbb{R} : P(C_{t,2}|x_t, D_t) \geq k\}. \tag{7}$$

The first step in the construction of the alarm system consists on the calculation of the probability of catastrophe conditional on D_t and x_t, i.e. $P(C_{t,2}|x_t, D_t, \boldsymbol{\theta})$ and $P(C_{t,2}|D_t, \boldsymbol{\theta})$ with

$\boldsymbol{\theta} = (\omega, \beta, \phi, \gamma, \delta, d)$. In doing so, note that

$$\begin{aligned} P(C_{t,2}|x_t, D_t, \boldsymbol{\theta}) &= P(X_{t+1} \leq u < X_{t+2}|x_1, \ldots, x_t, \boldsymbol{\theta}) \\ &= \int_{C_{t,2}} f_{X_{t+1}, X_{t+2}|x_1,\ldots,x_t,\boldsymbol{\theta}}(x_{t+1}, x_{t+2}) dx_{t+1} dx_{t+2}, \end{aligned}$$

with the integration region, $C_{t,2}$, being the catastrophe region as in (6). If $Z_t \sim N(0, 1)$ then

$$P(C_{t,2}|x_t, D_t, \boldsymbol{\theta}) = \int_u^{+\infty} \int_{-\infty}^u \prod_{k=1}^{2} \frac{1}{\sqrt{2\pi\sigma^2_{t+k}}} exp\left(-\frac{x^2_{t+k}}{2\sigma^2_{t+k}}\right) dx_{t+1} dx_{t+2}. \tag{8}$$

Moreover

$$\begin{aligned} P(C_{t,2}|D_t, \boldsymbol{\theta}) &= P(X_{t+1} \leq u < X_{t+2}|x_1, \ldots, x_{t-1}, \boldsymbol{\theta}) \\ &= \int_{C_{t,2}} \int f_{X_t, X_{t+1}, X_{t+2}|x_1,\ldots,x_{t-1},\boldsymbol{\theta}}(x_t, x_{t+1}, x_{t+2}) dx_t dx_{t+1} dx_{t+2}. \end{aligned}$$

Again, by assuming $Z_t \sim N(0, 1)$ it follows that

$$P(C_{t,2}|D_t, \boldsymbol{\theta}) = \int_u^{+\infty} \int_{-\infty}^u \int_{-\infty}^{+\infty} \prod_{k=0}^{2} \frac{1}{\sqrt{2\pi\sigma^2_{t+k}}} exp\left(-\frac{x^2_{t+k}}{2\sigma^2_{t+k}}\right) dx_t dx_{t+1} dx_{t+2}. \tag{9}$$

Having calculated these probabilities it is then possible to determine the alarm region and calculate the alarm characteristics of the alarm system.

1. Alarm size

$$\begin{aligned} \alpha_{t,2} &= P(A_{t,2}|D_t) \\ &= \int_{A_{t,2}} \frac{1}{\sqrt{2\pi\sigma^2_t}} exp\left(-\frac{x^2_t}{2\sigma^2_t}\right) dx_t, \end{aligned}$$

   with $A_{t,2}$ being the alarm region which depends on the value of $k_{t,2}$ chosen.

2. Probability of correct alarm

$$P(C_{t,2}|A_{t,2}, D_t) = \frac{P(C_{t,2} \cap A_{t,2}|D_t)}{P(A_{t,2}|D_t)},$$

where

$$
\begin{aligned}
P(C_{t,2} \cap A_{t,2}|D_t) &= P(X_{t+1} \leq u < X_{t+2} \cap X_t \in A_{t,2}|D_t) \\
&= \int_u^{+\infty} \int_{-\infty}^{u} \int_{A_{t,2}} \prod_{k=0}^{2} \frac{1}{\sqrt{2\pi}\sigma_{t+k}^2} exp\left(-\frac{x_{t+k}^2}{2\sigma_{t+k}^2}\right) dx_t dx_{t+1} dx_{t+2}.
\end{aligned}
$$

Thus

$$
P(C_{t,2}|A_{t,2}, D_t) = \frac{\int_u^{+\infty} \int_{-\infty}^{u} \int_{A_{t,2}} \prod_{k=0}^{2} \frac{1}{\sqrt{2\pi}\sigma_{t+k}^2} exp\left(-\frac{x_{t+k}^2}{2\sigma_{t+k}^2}\right) dx_t dx_{t+1} dx_{t+2}}{\int_{A_{t,2}} \frac{1}{\sqrt{2\pi}\sigma_t^2} exp\left(-\frac{x_t^2}{2\sigma_t^2}\right) dx_t}.
$$

3. Probability of detecting the event

$$
\begin{aligned}
P(A_{t,2}|C_{t,2}, D_t) &= \frac{P(A_{t,2} \cap C_{t,2}|D_t)}{P(C_{t,2}|D_t)} \\
&= \frac{\int_u^{+\infty} \int_{-\infty}^{u} \int_{A_{t,2}} \prod_{k=0}^{2} \frac{1}{\sqrt{2\pi}\sigma_{t+k}^2} exp\left(-\frac{x_{t+k}^2}{2\sigma_{t+k}^2}\right) dx_t dx_{t+1} dx_{t+2}}{\int_u^{+\infty} \int_{-\infty}^{u} \int_{-\infty}^{+\infty} \prod_{k=0}^{2} \frac{1}{\sqrt{2\pi}\sigma_{t+k}^2} exp\left(-\frac{x_{t+k}^2}{2\sigma_{t+k}^2}\right) dx_t dx_{t+1} dx_{t+2}}.
\end{aligned}
$$

3 Estimation procedures

In this section we consider the estimation of the operating characteristics. From the classical framework the method considered is the well-known Quasi-Maximum Likelihood Estimation procedure (QMLE) assuming conditional normality. The QMLE estimates are obtained maximizing the conditional log-likelihood function with respect to $\boldsymbol{\theta} = (\omega, \beta, \phi, \gamma, \delta, d)$, recurring to a routine available within the OxMetrics5 program. The parameter estimates were constrained to be in between the lower and upper bounds in (5). The robust standard errors by Bollerslev and Wooldrige (1992) were also calculated. According with these authors this estimator is generally consistent, has a normal limiting distribution and provides asymptotic standard errors that are valid under non-normality. Nevertheless, the authors state that the QMLE estimator is not asymptotically efficient under non-normality and care should be taken, since as Engle and Gonzalez-Rivera (1991) proved, GARCH estimates are consistent but asymptotically inefficient with the degree of inefficiency increasing with the degree of departure from normality. The impact of violations in conditional normality, however, remains unknown for the FIGARCH and FIAPARCH case. Baillie et al. (1996) suggested that the FIGARCH estimates obtained via QMLE are consistent and asymptotically normal[2]. Furthermore, they also demonstrated the suitability of the QMLE procedure in the estimation of samples with sizes of 1500 and 3000.

From the Bayesian perspective we need to start with a prior distribution for the vector of parameters $\boldsymbol{\theta}$. Assuming independence between all the parameters involved and taking in to account the constrains in (5), the prior distribution of $\boldsymbol{\theta}$, say $h(\boldsymbol{\theta})$, should be proportional to

$$h(\boldsymbol{\theta}) \propto I_{\{\omega>0\}} I_{\{\beta\geqslant 0\}} I_{\{\phi\geqslant 0\}} I_{\{-1<\gamma<1\}} I_{\{\delta\geqslant 0\}} I_{\{0<d<0.5\}}.$$

The posterior distribution $h(\boldsymbol{\theta}|D_t)$ is given by

$$\begin{aligned} h(\boldsymbol{\theta}|D_t) &\propto L(D_t|\boldsymbol{\theta})h(\boldsymbol{\theta}) \\ &\propto \prod_{n=2}^{t-1} \frac{1}{\sqrt{2\pi}\sigma_n} exp(-\frac{x_n^2}{2\sigma_n^2}) I_{\{\omega>0\}} I_{\{\beta\geqslant 0\}} I_{\{\phi\geqslant 0\}} I_{\{-1<\gamma<1\}} I_{\{\delta\geqslant 0\}} I_{\{0<d<0.5\}}. \end{aligned}$$

Hence, the probability of catastrophe conditional on D_t and $\mathbf{x}_2 = \{x_t\}$, takes the form

$$P(C_{t,2}|x_t, D_t) = \int_\Theta P(C_{t,2}|x_t, D_t, \boldsymbol{\theta}) h(\boldsymbol{\theta}|D_t) d\boldsymbol{\theta}, \tag{10}$$

with Θ being the parameter space. On the other hand, the probability of catastrophe conditional on D_t, will be given by

$$P(C_{t,2}|D_t) = \int_\Theta P(C_{t,2}|D_t, \boldsymbol{\theta}) h(\boldsymbol{\theta}|D_t) d\boldsymbol{\theta}, \tag{11}$$

where $P(C_{t,2}|x_t, D_t, \boldsymbol{\theta})$ and $P(C_{t,2}|D_t, \boldsymbol{\theta})$ are calculated through (8) and (9), respectively. However, due to the complexity of expressions (8) and (9) analytical calculations are not possible. Nonetheless, since by definition

$$P(C_{t,2}|x_t, D_t) = E_{\boldsymbol{\theta}|D_t}[P(C_{t,2}|x_t, D_t, \boldsymbol{\theta})] \quad \text{and} \quad P(C_{t,2}|D_t) = E_{\boldsymbol{\theta}|D_t}[P(C_{t,2}|D_t, \boldsymbol{\theta})],$$

[2]In fact, the consistency and asymptotic normality of the QMLE estimator had been formally established for the IGARCH(1,1) process. Baillie et al. (1996) followed a dominance-type argument to extend this result to the FIGARCH$(1, d, 0)$ case and refer the need for a formal proof of consistency and asymptotic normality for the general IGARCH(p, q) and FIAGARCH(p, d, q) cases.

their respective Monte Carlo approximations can be used, that is

$$\widehat{P}(C_{t,2}|x_t, D_t) = \frac{1}{m}\sum_{i=1}^{m} P(C_{t,2}|x_t, D_t, \boldsymbol{\theta}_i) \quad \text{and} \quad \widehat{P}(C_{t,2}|D_t) = \frac{1}{m}\sum_{i=1}^{m} P(C_{t,2}|D_t, \boldsymbol{\theta}_i),$$

where the observations $\boldsymbol{\theta}_i = (\omega_i, \beta_i, \phi_i, \gamma_i, \delta_i, d_i)$ with $i = 1, 2, \ldots, m$ constitute a sample of the posterior distribution $h(\boldsymbol{\theta}|D_t)$. A similar procedure is applied to approximate the operating characteristics.

4 Simulation results

In this section we present a simulation study to illustrate the performance of the alarm system constructed for the FIAPARCH$(1, d, 1)$ model. In particular we consider the set of parameters $\boldsymbol{\theta} = (0.40, 0.28, 0.10, 0.68, 1.27, 0.30)$. Figure 1 below shows a simulated sample path for this specific FIAPARCH model.

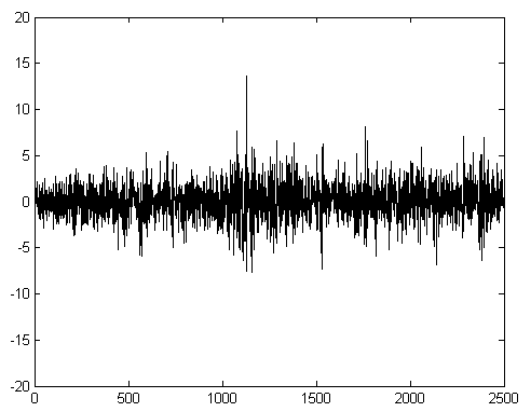

Figure 1: FIAPARCH$(1, d, 1)$ process with $\boldsymbol{\theta} = (0.40, 0.28, 0.10, 0.68, 1.27, 0.30)$

Parameter estimates, $\hat{\boldsymbol{\theta}}$, and their corresponding standard errors were obtained for this sample, following the QMLE procedure of Bollerslev and Wooldrige (1992). Robust standard errors are estimated from the product $A(\hat{\boldsymbol{\theta}})^{-1}B(\hat{\boldsymbol{\theta}})A(\hat{\boldsymbol{\theta}})^{-1}$, where $A(\hat{\boldsymbol{\theta}})$ and $B(\hat{\boldsymbol{\theta}})$ denote the Hessian and the outer product of the gradients evaluated at $\hat{\boldsymbol{\theta}}$, respectively.

Moreover, Bayesian estimates were also obtained for this single sample. Since the standard Gibbs methodology is difficult to implement to FIAPARCH models partially due to the non-standard forms of the full conditional densities, the Metropolis-Hastings algorithm was implemented in the software Matlab. In addition, a multivariate t-distribution was used as the proponent one. The sampler algorithm ran 100000 iterations including a *burn-in* period of 40000 observations which are discarded for the posterior analysis, as suggested by Dellaportas (2000). Furthermore, only every twentieth iteration is stored in order to obtain an, approximately, independent and identically distributed sample. The estimates were taken as the means of the posterior distribution. The convergence of the Markov chain was analyzed through the R criterium of Gelman and Rubin (1992), the Z-score test of Geweke (1992) and by graphical methods.

The analysis of the alarm system is carried out at $t = 2000$, i.e., $\mathbf{x}_2 = \{x_{2000}\}$. The event of interest is the two step ahead catastrophe defined by the upcrossing of the fixed level u, at time $t+2$: $C_{2000,2} = \{(x_{2001}, x_{2002}) \in \mathbb{R}^2 : x_{2001} \leq u < x_{2002}\}$. In a first stage, two values of u were chosen, accordingly to the sample quantiles, namely the 90th percentile ($Q_{0.90}$), and the 95th percentile ($Q_{0.95}$). The choice of these values is justified by the fact that we are interested in relatively rare events. For both fixed levels of u, the probabilities $P(C_{t,2}|x_t, D_t, \theta)$ and $P(C_{t,2}|D_t, \theta)$ were numerically approximated as described in the previous section. In order to compute the optimal alarm region for each case, one has to obtain the region for several values of k, accordingly to expression (7) and then, for each value of k, compute the operating characteristics of the alarm system, i.e., the size of the region, $\alpha_{t,2}$, the probability of correct alarm, $P(C_{t,2}|A_{t,2}, D_t)$ and the probability of detection, $P(A_{t,2}|C_{t,2}, D_t)$. For every fixed value of k the region has to be obtained through a systematic search in a three dimensional region for (x_t, x_{t+1}, x_{t+2}). We considered a thin grid of values of x_t in $[-100, 100]$ and determined, for each value of x_t, whether $P(C_{t,2}|x_t, D_t)$ exceeds k. This procedure is repeated for k ranging from $P(C_{t,2}|D_t)$ to $P(C_{t,2}|D_t) + n \times 0.005$, with $n \in \mathbb{R}^+$. This procedure is repeated for both the classical (using the true values of the parameters and their QMLE estimates) and the Bayesian approach. The results are shown in Table 1 below.

Considering the true values of the parameters, the probability of the alarm being correct, does not exceed 5.6% in the $u = Q_{0.95}$ case, or 9.7% in the $u = Q_{0.90}$ case. The probability of detection for this sample, ranges from 2.5% to 49.0% for $u = Q_{0.95}$, or from 1.7% to 53.4% for $u = Q_{0.90}$. The results obtained with the QMLE estimates do not differ considerably, in particular in what concerns the probability of correct alarm. Regarding the probability of detecting the event, we can say the alarm system behaves better since the detection probability reaches 54.0% for $u = Q_{0.95}$ and 60.6% for $u = Q_{0.90}$. Considering now the Bayesian approach, the probability of detection is the lowest obtained. It does not even reach 22%. On the other hand, the estimation procedure involved in the Bayesian approach seems to be able to produce higher probabilities of correct alarm, depending on an accurate choice of k. The probability of correct alarm ranges from lower values than in the classical approach to more than the double of this values, with increasing k, reaching 24.7% in the $u = Q_{0.90}$ case. Furthermore, note that as the probability of correct alarm increases, the probability of detecting the event decreases, as expected. This can be justified by the fact that as k increases, the size of the alarm region decreases, which implies that the number of alarms should decrease, so as the probability of detection, $P(A_{t,2}|C_{t,2})$. However, as the number of alarms decreases, the probability of false alarms also decreases and therefore the probability of the alarm being correct, $P(C_{t,2}|A_{t,2})$, increases.

As already discussed, it is not possible, in general, to maximize both probabilities, $P(C_{t,2}|A_{t,2})$ and $P(A_{t,2}|C_{t,2})$, simultaneously. Hence, a compromise should be reached by the proper choice of k. In doing so, several criteria have been already proposed. Svensson et al. (1996), for example, suggested that k should be chosen so that the probability of correct alarm and the probability of detecting the event are approximately equal, $P(C_{t,2}|A_{t,2}) \simeq P(A_{t,2}|C_{t,2})$. On the other hand, Antunes et al. (2003) suggested that k should be chosen so that the alarm size is about twice the probability of having a catastrophe given the past values of the process,

Table 1: Operating Characteristics at time point $t = 2000$.

				$u = Q_{0.95} = 3.136$					
	True Parameters $P(C_{t,2}) = 0.0746$			QML Estimates $P(C_{t,2}) = 0.0893$			Bayesian Estimates $P(C_{t,2}) = 0.0817$		
k	α_2	$P(C_{t,2}\|A_{t,2})$	$P(A_{t,2}\|C_{t,2})$	α_2	$P(C_{t,2}\|A_{t,2})$	$P(A_{t,2}\|C_{t,2})$	α_2	$P(C_{t,2}\|A_{t,2})$	$P(A_{t,2}\|C_{t,2})$
0.0350	0.4789	0.0335	0.4903	0.5353	0.0346	0.5446	0.1904	0.0267	0.2155
0.0400	0.2998	0.0345	0.3155	0.3255	0.0355	0.3400	0.1902	0.0257	0.2074
0.0450	0.2072	0.0349	0.2209	0.2971	0.0359	0.3133	0.1211	0.0264	0.1354
0.0500	0.2067	0.0344	0.2173	0.2102	0.0363	0.2247	0.0718	0.0283	0.0862
0.0600	0.1377	0.0347	0.1458	0.1413	0.0360	0.1496	0.0397	0.0318	0.0535
0.0700	0.0864	0.0363	0.0957	0.0896	0.0373	0.0983	0.0203	0.0391	0.0337
0.0800	0.0509	0.0390	0.0605	0.0535	0.0398	0.0625	0.0097	0.0555	0.0227
0.0900	0.0282	0.0439	0.0377	0.0300	0.0454	0.0401	0.0042	0.0982	0.0177
0.1000	0.0146	0.0558	0.0248	0.0158	0.0563	0.0262	0.0017	0.2061	0.0151
				$u = Q_{0.90} = 2.293$					
	True Parameters $P(C_{t,2}) = 0.1267$			QML Estimates $P(C_{t,2}) = 0.1439$			Bayesian Estimates $P(C_{t,2}) = 0.1353$		
k	α_2	$P(C_{t,2}\|A_{t,2})$	$P(A_{t,2}\|C_{t,2})$	α_2	$P(C_{t,2}\|A_{t,2})$	$P(A_{t,2}\|C_{t,2})$	α_2	$P(C_{t,2}\|A_{t,2})$	$P(A_{t,2}\|C_{t,2})$
0.0850	0.5303	0.0832	0.5339	0.6042	0.0846	0.6055	0.1904	0.0722	0.1984
0.0900	0.3209	0.0837	0.3250	0.3490	0.0853	0.3528	0.1902	0.0719	0.1974
0.0950	0.2960	0.0844	0.3021	0.3033	0.0849	0.3050	0.1211	0.0717	0.1252
0.1000	0.2069	0.0843	0.2109	0.2117	0.0864	0.2167	0.1211	0.0713	0.1245
0.1100	0.1377	0.0852	0.1420	0.2101	0.0859	0.2137	0.0718	0.0730	0.0757
0.1200	0.0864	0.0862	0.0901	0.1413	0.0864	0.1446	0.0397	0.0773	0.0442
0.1300	0.0509	0.0887	0.0546	0.0535	0.0905	0.0573	0.0203	0.0825	0.0242
0.1400	0.0282	0.0888	0.0302	0.0535	0.0904	0.0572	0.0042	0.1123	0.0069
0.1500	0.0146	0.0965	0.0170	0.0158	0.1054	0.0197	0.0017	0.2474	0.0062

$P(C_{t,2}|D_t) \simeq \frac{1}{2}P(A_{t,2}|D_t)$, stating that in this situation the system will be spending twice the time in the alarm state than in the catastrophe region. We analyzed both criteria in this work and from hereafter, the former criterion will be designated by Criterion 2 and the last by Criterion 1.

In order to test the alarm system, three extra values of the series were simulated: $(\mathbf{x}_2, \mathbf{x}_3) = (x_t, x_{t+1}, x_{t+2})$. This procedure was repeated 10000 times with the same informative experience, D_t. With the alarm regions calculated before for $u = Q_{0.90} = 2.293$ and for the two criteria already mentioned, we observed, for each of the 10000 samples, whether an alarm was given or not and wether a catastrophe occurred or not. Results are given in Table 2.

Table 2: Results at time point $t = 2000$. Percentages in parenthesis.

	Criterion	Alarms		Catastrophes	
		False	Total	Detected	Total
True	1	1112 (0.8330)	1335	223 (0.2059)	1083
Parameters	2	651 (0.8314)	783	132 (0.1273)	1037
QMLE	1	1163 (0.8526)	1364	201 (0.1963)	1024
Approach	2	380 (0.8261)	460	80 (0.0771)	1037
Bayesian	1	1161 (0.8401)	1382	221 (0.2103)	1051
Approach	2	668 (0.8477)	788	120 (0.1204)	997

Finally, we illustrate how the online prediction performs in practice. The event to predict is $C_{t,2} = \{(x_{t+1}, x_{t+2}) \in \mathbb{R}^2 : x_{t+1} \leq u < x_{t+2}\}$, $t = 2000, \ldots, 2010$, again with $u = Q_{0.90} = 2.293$. Alarm regions and respective operating characteristics are presented in Table 3 for Criterion 1 and in Table 4 for Criterion 2.

Overall, Criterion 1 provides better estimates for the operating characteristics. The probability of detection, for instance, reaches values around 0.22 in some cases for the classical approach

Table 3: Operating characteristics at different time points, with Criterion 1.

Approach	t	$P(C_{t,2}\|D_t)$	k	Alarm Region	α_2	$P(C_{t,2}\|A_{t,2})$	$P(A_{t,2}\|C_{t,2})$
True Parameters	2000	0.0827	0.1100	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1377	0.0852	0.1420
	2001	0.1047	0.1047	$[-\infty, -1.5] \cup [5.5, +\infty]$	0.1848	0.1093	0.1929
	2002	0.0936	0.0936	$[-\infty, -2.0] \cup [9.5, +\infty]$	0.1209	0.0980	0.1265
	2003	0.0923	0.1073	$[-\infty, -1.5] \cup [7.5, +\infty]$	0.2167	0.0947	0.2224
	2004	0.0897	0.0977	$[-\infty, -1.5] \cup [8.0, +\infty]$	0.2076	0.0914	0.2116
	2005	0.0879	0.0979	$[-\infty, -1.5] \cup [7.5, +\infty]$	0.2036	0.0893	0.2069
	2006	0.0803	0.0953	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1311	0.0831	0.1356
	2007	0.0687	0.0887	$[-\infty, -2.0] \cup [8.5, +\infty]$	0.1286	0.0716	0.1340
	2008	0.0573	0.0873	$[-\infty, -2.0] \cup [9.5, +\infty]$	0.1194	0.0614	0.1279
	2009	0.0508	0.0758	$[-\infty, -2.0] \cup [8.5, +\infty]$	0.1045	0.0522	0.1075
	2010	0.0545	0.0845	$[-\infty, -2.0] \cup [8.5, +\infty]$	0.0924	0.0566	0.0960
QMLE	2000	0.0844	0.1200	$[-\infty, -2.0] \cup [10.5, +\infty]$	0.1413	0.0864	0.1446
	2001	0.1097	0.1047	$[-\infty, -1.5] \cup [6.0, +\infty]$	0.1867	0.1123	0.2002
	2002	0.0969	0.0969	$[-\infty, -2.0] \cup [9.5, +\infty]$	0.1230	0.1005	0.1276
	2003	0.0946	0.1096	$[-\infty, -1.5] \cup [7.5, +\infty]$	0.2202	0.0972	0.2262
	2004	0.0919	0.1019	$[-\infty, -1.5] \cup [7.5, +\infty]$	0.2110	0.0943	0.2165
	2005	0.0900	0.1000	$[-\infty, -1.5] \cup [7.5, +\infty]$	0.2066	0.0917	0.2104
	2006	0.0821	0.0971	$[-\infty, -2.0] \cup [8.5, +\infty]$	0.1340	0.0843	0.1376
	2007	0.0697	0.0897	$[-\infty, -2.0] \cup [8.5, +\infty]$	0.1314	0.0723	0.1363
	2008	0.0594	0.0894	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1217	0.0619	0.1269
	2009	0.0506	0.0756	$[-\infty, -2.0] \cup [8.0, +\infty]$	0.1059	0.0528	0.1104
	2010	0.0544	0.0844	$[-\infty, -2.0] \cup [8.5, +\infty]$	0.0930	0.0566	0.0966
Bayesian	2000	0.0693	0.0950	$[-\infty, -2.0] \cup [8.5, +\infty]$	0.1211	0.0717	0.1252
	2001	0.0911	0.0911	$[-\infty, -1.5] \cup [6.0, +\infty]$	0.1685	0.0939	0.1736
	2002	0.0820	0.0820	$[-\infty, -2.0] \cup [9.5, +\infty]$	0.1047	0.0845	0.1078
	2003	0.0794	0.0994	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1297	0.0820	0.1340
	2004	0.0764	0.0914	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1218	0.0797	0.1271
	2005	0.0715	0.0915	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1176	0.0779	0.1282
	2006	0.0680	0.0830	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1144	0.0711	0.1196
	2007	0.0576	0.0776	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1121	0.0598	0.1165
	2008	0.0498	0.0748	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.1038	0.0513	0.1068
	2009	0.0419	0.0669	$[-\infty, -2.0] \cup [9.0, +\infty]$	0.0902	0.0441	0.0948
	2010	0.0447	0.0747	$[-\infty, -2.0] \cup [9.5, +\infty]$	0.0790	0.0467	0.0825

whereas with Criterion 2 this probability is nearly only half the former.

5 Exploring the IBOVESPA returns data set

In this section, we model the data set IBOVESPA which contains daily returns of the S. Paulo Stock Market during the period 04/07/1994 to 02/10/2008 (www.ipeadata.gov.br). Data consists on the closing rates of stocks, I_t, being the log-returns calculated as $y_t = \ln(I_t/I_{t-1})$, $t =$

Table 4: Operating characteristics at different time points, with Criterion 2.

Approach	t	$P(C_{t,2}\|D_t)$	k	Alarm Region	α_2	$P(C_{t,2}\|A_{t,2})$	$P(A_{t,2}\|C_{t,2})$
True Parameters	2000	0.0827	0.1200	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0864	0.0862	0.0901
	2001	0.1047	0.1247	$[-\infty,-2.0]\cup[10.5,+\infty]$	0.1153	0.1088	0.1198
	2002	0.0936	0.1036	$[-\infty,-2.5]\cup[12.0,+\infty]$	0.0717	0.1001	0.0767
	2003	0.0923	0.1223	$[-\infty,-2.5]\cup[12.0,+\infty]$	0.0958	0.0949	0.0985
	2004	0.0897	0.1147	$[-\infty,-2.5]\cup[12.0,+\infty]$	0.0872	0.0924	0.0899
	2005	0.0879	0.1129	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0835	0.0906	0.0862
	2006	0.0803	0.1053	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0805	0.0831	0.0832
	2007	0.0687	0.0987	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0783	0.0726	0.0827
	2008	0.0573	0.1023	$[-\infty,-2.5]\cup[13.0,+\infty]$	0.0705	0.0630	0.0774
	2009	0.0508	0.0908	$[-\infty,-2.5]\cup[12.0,+\infty]$	0.0582	0.0531	0.0608
	2010	0.0545	0.0945	$[-\infty,-2.5]\cup[11.0,+\infty]$	0.0487	0.0593	0.0530
QMLE	2000	0.0844	0.1300	$[-\infty,-3.0]\cup[13.5,+\infty]$	0.0535	0.0905	0.0573
	2001	0.1047	0.1297	$[-\infty,-2.0]\cup[10.5,+\infty]$	0.1174	0.1104	0.1238
	2002	0.0969	0.1069	$[-\infty,-2.5]\cup[12.0,+\infty]$	0.0735	0.1027	0.0780
	2003	0.0946	0.1246	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0992	0.0974	0.1021
	2004	0.0919	0.1169	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0904	0.0947	0.0932
	2005	0.0900	0.1150	$[-\infty,-2.5]\cup[11.0,+\infty]$	0.0863	0.0929	0.0891
	2006	0.0821	0.1121	$[-\infty,-2.5]\cup[12.5,+\infty]$	0.0831	0.0850	0.0860
	2007	0.0697	0.0997	$[-\infty,-2.5]\cup[11.0,+\infty]$	0.0808	0.0731	0.0847
	2008	0.0594	0.0994	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0723	0.0637	0.0776
	2009	0.0506	0.0956	$[-\infty,-2.5]\cup[13.0,+\infty]$	0.0593	0.0529	0.0619
	2010	0.0544	0.0994	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0491	0.0590	0.0533
Bayesian	2000	0.0693	0.1100	$[-\infty,-2.5]\cup[12.5,+\infty]$	0.0718	0.0730	0.0757
	2001	0.0911	0.1011	$[-\infty,-2.0]\cup[8.5,+\infty]$	0.1002	0.0943	0.1037
	2002	0.0820	0.0820	$[-\infty,-2.0]\cup[9.5,+\infty]$	0.1047	0.0845	0.1078
	2003	0.0794	0.1094	$[-\infty,-2.5]\cup[12.0,+\infty]$	0.0793	0.0835	0.0835
	2004	0.0764	0.1014	$[-\infty,-2.5]\cup[12.0,+\infty]$	0.0724	0.0813	0.0771
	2005	0.0715	0.1065	$[-\infty,-2.5]\cup[13.5,+\infty]$	0.0689	0.0794	0.0766
	2006	0.0680	0.0930	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0663	0.0726	0.0707
	2007	0.0576	0.0876	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0643	0.0619	0.0692
	2008	0.0498	0.0848	$[-\infty,-2.5]\cup[12.0,+\infty]$	0.0576	0.0536	0.0619
	2009	0.0419	0.0769	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0470	0.0461	0.0517
	2010	0.0447	0.0847	$[-\infty,-2.5]\cup[11.5,+\infty]$	0.0388	0.0476	0.0413

$1,\ldots,n$. The results obtained from this procedure were then multiplied by 100 just to ensure the stability of posterior calculations. Sáfadi and Pereira (2008) proved that the FIAPARCH$(1,d,1)$ provides a good fit for this kind of data sets. To fit a FIAPARCH$(1,d,1)$ model for the log-returns we proceeded as follows: first, the AR(10) model $y_t = 0.0689 + 0.0645 y_{t-10} + x_t$, is fitted, using the least squares method, in order to eliminate serial dependence. The time series plot of both the IBOVESPA daily returns and the residuals (x_t), hereafter designated by x-returns, are exhibited in Figure 2 below. This is, indeed, the set of data reported to show the common

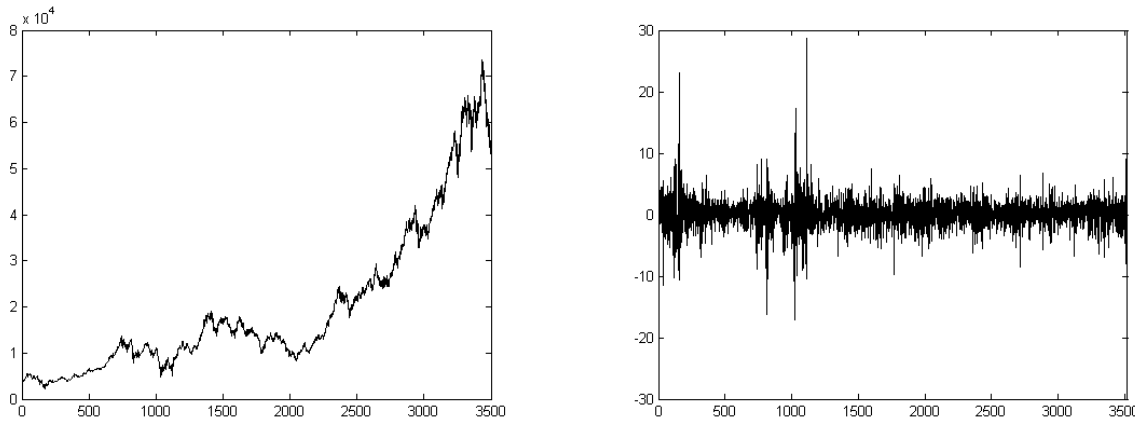

Figure 2: Plot of the IBOVESPA daily returns (left) and the x-returns (right) from 04/07/1994 to 02/10/2008

features of financial time series mentioned in Section 1, that is weak dependence without any evident pattern on the series level and significative dependence on squared and absolute returns.

The FIAPARCH$(1, d, 1)$ model was fitted to the series of x-returns by means of the QMLE procedure and the Bayesian approach described in Section 3. In both cases the adequacy of the fit was checked through the analysis of the standardized residuals. Table 5 presents the estimates obtained for both procedures. Since the IBOVESPA x-returns are related to the daily changes of the stock indexes of S. Paulo Stock Market, we considered that the event of interest is given by

$$C_{t,2} = \{(x_{t+1}, x_{t+2}) \in \mathbb{R}^2 : x_{t+1} \geq u > x_{t+2}\},$$

with $t = 3450, \ldots, 3516$, corresponding to July, August and September of 2008, and $u = Q_{0.25} = -1.219$. Note that, the downcrossing event $C_{t,2}$ can be view as related with a stock market crash. Moreover, the choice of k was done according only to Criterion 1: $P(C_{t,2}|D_t) \simeq \frac{1}{2}P(A_{t,2}|D_t)$. Two reasons justify this choice. First, Criterion 2 is difficult to implement since $P(C_{t,2}|A_{t,2}, D_t)$ may never get so close to $P(A_{t,2}|C_{t,2}, D_t)$ or when it does, some operating characteristics may

Table 5: Parameter estimates. Standard deviations in parenthesis

	QMLE	Bayesian Estimates
ω	0.3903 (0.1092)	0.4227 (0.0576)
ϕ	0.0957 (0.1334)	0.1289 (0.0397)
γ	0.6782 (0.1363)	0.7813 (0.1108)
β	0.2794 (0.1693)	0.3246 (0.0568)
δ	1.2744 (0.1274)	1.2218 (0.1008)
d	0.2952 (0.0642)	0.3020 (0.0258)

show not so good results (at least as compared with those obtained with Criterion 1). Secondly, Criterion 1 results in better estimates of the operating characteristics. For the time period considered, the total number of alarms, the total number of catastrophes, the number of false alarms and the number of detected events was counted. Results are presented in Table 6.

Table 6: Results of the alarm system with $u = -1.219$. Percentages in parenthesis.

Month	Alarms		Catastrophes	
	False	Total	Detected	Total
July	1 (0.5000)	2	1 (0.1667)	6
August	1 (0.5000)	2	1 (0.2000)	5
September	0 (0.0000)	3	3 (0.2727)	11
Trimester	2 (0.2857)	7	5 (0.2273)	22

A closer look to Table 6 reveals that the probability of the alarm being correct is 50% in July and August and raises to 100% in September. In addition, the probability of detecting a catastrophe remains around 20% during the time period considered. We noticed that this online prediction system exhibits an adaptive behavior, that is, as long as the available information is integrated within the informative experience, the system adapts itself in order to produce the minimum number of false alarms. This fact explains on one hand the high probabilities of the alarm given being correct and on the other hand that the system produces few alarms, so the probability of detection can not be very high.

References

[1] Amaral-Turkman, M. A. and Turkman, K. F. (1990). Optimal alarm systems for autoregressive processes; a Bayesian approach. *Comput. Statist. Data Anal.* **10**, 307–314.

[2] Antunes, M.; Amaral-Turkman, M. A. and Turkman, K. F. (2003). A Bayesian approach to event prediction intervals. *J. Time Ser. Anal.* **24**, 631–646.

[3] Baillie, R. T.; Bollerslev, T. and Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heterokedasticity. *J. Econom.* **74**, 3–30.

[4] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedascity. *J. Econom.* **31**, 307–327.

[5] Bollerslev, T. and Mikkelsen, H. O. (1996). Modelling and pricing long memory in stock market volatility. *J. Econom.* **73**, 151–184.

[6] Bollerslev, T. and Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Rev.* **11**, 143-172.

[7] de Maré, J. (1980). Optimal prediction of catastrophes with application to Gaussian process. *Ann. Probab.* **8**, 841–850.

[8] Conrad, C.; Karanasos, M. and Zeng, N. (2008). Multivariate fractionally integrated APARCH modeling of stock market volatility: a multi-country study. (Submitted)

[9] Dellaportas, P.; Politis, D. N. and Vrontos, I. D. (2000). Full bayesian inference for GARCH e EGARCH models. *J. Bus. Econom. Statist.* **18**, 187–196.

[10] Ding, Z.; Engle, R. F. and Granger, C. W. J. (1993). A long memory property of stock market returns and a new model. *J. Empir. Finance* **1**, 83–106.

[11] Engle, R. F. (1982). Autoregressive conditional heteroskedascity with estimates of the United Kingdom inflation. *Econometrica* **50**, 987–1008.

[12] Engle, R. F. and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Rev.* **5**, 1–50.

[13] Engle, R. F. and Gonzalez-Rivera, G. (1991). Semi-parametric ARCH model. *J. Bus. Econom. Statist.* **9**, 345–360.

[14] Fuentes, A-M.; Izzeldin, I. and Kalotychou, E. (2009). On forecasting daily stock volatility: The role of intraday information and market conditions. *Int. J. Forecast.* **25**, 259–281.

[15] Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7**, 457–511.

[16] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. *Bayesian Statistics 4* (J. M. Bernardo, J. O. Berger, A. P. Dawin and A. F. M. Smith, eds.). Oxford: University Press, 169–193, (with discussion).

[17] Koop, G. and Tale, L. (2004). Measuring the health effects of air pollution: to what extent can we really say that people are dying from bad air? *J. Environ. Econ. Manage.* **47**, 30–54.

[18] Lindgren, G. (1980). Model processes in non-linear prediction, with application to detection and alarm. *Ann. Probab.* **8**, 775–792.

[19] Lindgren, G. (1985). Optimal prediction of level crossings in Gaussian processes and sequences. *Ann. Probab.* **13**, 804–824.

[20] Mikosch, T. (2003). Modeling dependence and tails of financial time series. *Extreme Values in Finance, Telecommunications, and the Environment* (B. Finkenstadt and H. Rootzén, eds.). Boca Raton: Chapman & Hall, 185–286.

[21] Monteiro, M.; Pereira, I. and Scotto, M. G. (2008). Optimal alarm systems for count processes. *Commun. Statist. Theory and Methods* **37**, 3054–3076.

[22] Sáfadi, T. and Pereira, I. (2008). Bayesian analysis of FIAPARCH model: an application to So Paulo Stock Market. *Cadernos de Matemática*, CM08; I-4, Departamento de Matemática, Universidade de Aveiro.

[23] Smith, R. L.; Davis, J. M.; Sacks, J.; Speckman, P. and Styer, P. (2000). Regression models for air pollution and daily mortallity: analysis of data from Birmingham, Alabama. *Environmetrics* **11**, 719–743.

[24] Svensson, A.; Lindquist, and R. Lindgren, G. (1996). Optimal prediction of catastrophes in autoregressive moving average processes. *J. Time Ser. Anal.* **17**, 511–531.

[25] Teräsvirta, T. (2009). An introduction to univariate GARCH models. *Handbook of Financial Time Series* (T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch, eds.). New York: Springer, 17–42.

[26] Thomas, L. C. (2000). A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. *Int. J. Forecast.* **16**, 149–172.

[27] Tobias, A. and Scotto, M. G. (2005). Prediction of extreme ozone levels in Barcelona, Spain. *Environ. Monit. Assess.* **100**, 23-32.

[28] Tse, Y. (1998). The conditional heteroskedascity of the Yen-Dollar exchange rate. *J. Appl. Econom.* **13**, 49–55.

[29] Weatherford, L. R. and Kimes, S. E. (2003). A comparison of forecasting methods for hotel revenue management. *Int. J. Forecast.* **19**, 401–415.