arXiv:1305.5238v1 [q-fin.RM] 22 May 2013

RISK MEASURE ESTIMATION ON FIEGARCH PROCESSES

T.S. Prass* and S.R.C. Lopes†

Instituto de Matemática - UFRGS
Porto Alegre - RS - Brazil

This version: November, 2011

Abstract

We consider the Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroskedasticity process, denoted by FIEGARCH(p, d, q), introduced by Bollerslev and Mikkelsen (1996). We present a simulated study regarding the estimation of the risk measure VaR_p on FIEGARCH processes. We consider the distribution function of the portfolio log-returns (univariate case) and the multivariate distribution function of the risk-factor changes (multivariate case). We also compare the performance of the risk measures VaR_p, ES_p and MaxLoss for a portfolio composed by stocks of four Brazilian companies.

Mathematics Subject Classification (2000). Primary 60G10, 62G05, 62G35, 62M10, 62M15; Secondary 62M20.

Keywords. Long Memory Models, Volatility, Risk Measure Estimation, FIEGARCH Processes.

1 Introduction

In financial terms, risk is the possibility that an investment will have a return different from the expected, including the possibility of losing part or even all the original investment. A portfolio is a collection of investments maintained by an institution or a person. In this paper portfolio will be used to indicate a collection of stocks. The selection of an efficient portfolio is an important issue and it is discussed in Prass and Lopes (2009). The authors consider an approach based in the mean-variance (MV) method introduced by Markowitz (1952).

In finance one of the most important problems is risk management, which involves risk measures estimation. Among the applications of a risk measure we can say that it can be

*E-mail: taianeprass@gmail.com
†E-mail: silviarc.lopes@gmail.com

used to determine the capital in risk, that is, we can measure the exposure to the risks of a financial institution, in order to determine the necessary amount to honour possible unexpected losses. A more detailed study on quantitative risk management, including theoretical concepts and practical examples, can be found in MacNeil et al. (2005).

Financial time series present an important characteristic known as volatility, which can be defined in different ways but it is not directly observable. In financial terms, the volatility is associated to the risk of an asset. The volatility can be seen as the statistic measure of the possibility that the value of an asset will significantly increase or decrease, several times, in a given period of time. As a risk measure, the volatility can be calculated by different approaches. The most common one, but not unique, is to use the variance (or the standard deviation) of the historical rentability of a given investment.

In order to model time series, in the presence of volatility clusters, heteroskedastic models need to be considered. This class of models consider the variance as a function not only of the time but also of the past observations. The fitted models are then used to estimate the (conditional) mean and variance of the process and these values in turn are used in several risk measures estimation.

Among the most used non-linear models we find the ARCH-type model (Autoregressive Conditional Heteroskedasticity) and its extension. Such models are used to describe the conditional variance of a time series. The ARCH(p) models were introduced by Engle (1982). The main assumption of the model is that the random variables $\{X_t\}_{t\in\mathbb{Z}}$ are not correlated, but the conditional variance depends on the square of the past p values of the process. This model was generalized by Bollerslev (1986) with the introduction of the GARCH(p, q) models (Generalized ARCH). In this model the conditional variance depends not only on the past values of the process but also on the past values of the conditional variance.

ARCH and GARCH models appear frequently in the literature due to their easy implementation. However, this class of models present a drawback. These models do not take into account the signal of the process $\{X_t\}_{t\in\mathbb{Z}}$, since the conditional variance is a quadratic function of those values. In order to deal with this problem, Nelson (1991) introduced the EGARCH(p, q) models (Exponential GARCH). As in the linear case, where the ARFIMA models are presented as a generalization of ARMA models, in the non-linear case the FIGARCH models (Fractionally Integrated GARCH), introduced by Baillie et al. (1996) and FIEGARCH models (Fractionally Integrated EGARCH), introduced by Bollerslev and Mikkelsen (1996), appear to generalize the GARCH and EGARCH models, respectively. A study on ARFIMA and FIEGARCH models, among other non-linear processes, can be found in Lopes (2008) and Lopes and Mendes (2006).

Lopes and Prass (2009) present a theoretical study on FIEGARH process and data analysis. The authors present several properties of these processes, including results on their stationarity and their ergodicity. It is shown that the process $\{g(Z_t)\}_{t\in\mathbb{Z}}$, in the definition of FIEGARCH processes, is a white noise and from this result the authors prove that if $\{X_t\}_{t\in\mathbb{Z}}$ is a FIEGARCH(p, d, q) process then, $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ is an ARFIMA(q, d, p) process. Also, under mild conditions, $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is an ARFIMA($q, d, 0$) process with non-Gaussian innovations. Lopes and Prass (2009) also analyze the autocorrelation and spectral density functions decay of the $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ process and the convergence order for the polynomial coefficients that describes the volatility.

The most used methodology to calculate risk measures is based on the assumption that the distribution of the data is Gaussian. A drawback present in this method is that the distribution function of financial series usually present tails heavier than the normal distribution. A very well known risk measure is the Value-at-Risk (VaR_p). Nowadays it has being changed but the most used methods to estimate the VaR_p consider the Gaussian assumption. Khindanova and Atakhanov (2002) present a comparison study on VaR_p estimation. The authors consider the Gaussian, empirical and stable distributions. The study demonstrates that stable modeling captures asymmetry and heavy-tails of returns, and, therefore, provides more accurate estimates of VaR_p. It is also known that, even considering heavier tail distributions, the risk measures are underestimated for events with small occurrence probability (extreme events). Embrechts et al. (1997) present ideas on the modeling of extremal events with special emphasis on applications to insurance and finance.

Another very common approach is to consider the scenario analysis. Usually one does not make any assumption on the data distribution. Since this analysis does not provide information on an event probability, it should be used as a complementary tool to other risk measure estimation procedures. Even though the scenarios need to be chosen in such a way that they are plausible and for that it is necessary to have an idea of the occurrence probability for each scenario. The maximum loss, denoted by MaxLoss, introduced by Studer (1997), can be viewed either as risk measure or as a systematic way to perform a stress test (scenario analysis). In a portfolio analysis, this risk measure can be interpreted as the worst possible loss in the portfolio value.

This paper is organized as follows. Section 2 gives some definitions and some properties of FIEGARCH(p, d, q) processes. Section 3 defines some risk measures considered here and their relationship. Section 4 gives a simulation study. Section 5 presents a portfolio analysis. Section 6 concludes the paper.

2 FIEGARCH Process

Financial time series present characteristic common to another time series such as, trend, seasonality, outliers and heteroskedasticity. However, empirical studies show that return (or log-return) series present some stylized facts. We can say that the return series are not i.i.d. although they show little serial correlation, while the series of absolute or squared returns show profound serial correlation, the conditional expected returns are close to zero, volatility appears to vary over time, return series are leptokurtic (or heavy-tailed) and extreme returns appear in clusters. Due to these characteristics, modeling these time series require considering a class of non-linear heteroskedastic models. In this section we present the *Fractionally Integrated Exponential Generalized Autoregressive Conditional Heteroskedasticity* process, denoted by FIEGARCH(p, d, q). This class of processes, introduced by Bollerslev and Mikkelsen (1996), describes the volatility varying on time, volatility clusters (known as ARCH and GARCH effects), volatility long-range dependence and also asymmetry.

Definition 2.1. Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stochasric process. Then, $\{X_t\}_{t\in\mathbb{Z}}$ is a *Fractionally Integrated* EGARCH *process*, denoted by FIEGARCH(p, d, q), if and only if,

$$X_t = \sigma_t Z_t, \tag{2.1}$$

$$\ln(\sigma_t^2) = \omega + \frac{\alpha(\mathcal{B})}{\beta(\mathcal{B})(1-\mathcal{B})^d} g(Z_{t-1}), \quad \text{for all } t \in \mathbb{Z}, \tag{2.2}$$

where $\omega \in \mathbb{R}$, $\{Z_t\}_{t\in\mathbb{Z}}$ is a process of i.i.d. random variables with zero mean and variance equal to one, $\alpha(\cdot)$ and $\beta(\cdot)$ are the polynomials defined by $\alpha(\mathcal{B}) \equiv \sum_{i=0}^{p}(-\alpha_i)\mathcal{B}^i$ and $\beta(\mathcal{B}) \equiv \sum_{j=0}^{q}(-\beta_j)\mathcal{B}^j$, with $\alpha_0 = -1 = \beta_0$, $\beta(\mathcal{B}) \neq 0$, for all $\mathcal{B}$ such that $|\mathcal{B}| \leq 1$. The function $g(\cdot)$ is defined by

$$g(Z_t) = \theta Z_t + \gamma\big[|Z_t| - \mathbb{E}\big(|Z_t|\big)\big], \quad \text{for all } t \in \mathbb{Z}, \text{ and } \theta, \gamma \in \mathbb{R}, \tag{2.3}$$

and the operator $(1-\mathcal{B})^d$ is defined as $(1-\mathcal{B})^d = \sum_{k=0}^{\infty}(-1)\,\delta_{d,k}\,\mathcal{B}^k \equiv \delta_d(\mathcal{B})$, where $\delta_{d,0} = -1$ and

$$\delta_{d,k} = d \times \frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(1-d)} = \delta_{d,k-1}\left(\frac{k-1-d}{k}\right), \quad \text{for all } k \geq 1.$$

REMARK 2.1. If $d = 0$, in expression (2.2), we have an EGARCH(p, q) model.

Some properties of the $\{g(Z_t)\}_{t\in\mathbb{Z}}$ process can be found in Lopes and Prass (2009). The authors present a theoretical study on the FIEGARCH process properties, including results on their stationarity and their ergodicity. The authors also show that the process $\{g(Z_t)\}_{t\in\mathbb{Z}}$ is a white noise and use this result to prove that if $\{X_t\}_{t\in\mathbb{Z}}$ is a FIEGARCH(p, d, q) process then, $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ is an ARFIMA(q, d, p) process. Moreover, under mild conditions, $\{\ln(X_t^2)\}_{t\in\mathbb{Z}}$ is an ARFIMA$(q, d, 0)$ process with non-Gaussian innovations. The autocorrelation and spectral density functions decay of the $\{\ln(\sigma_t^2)\}_{t\in\mathbb{Z}}$ process and the convergence order for the polynomial coefficients that describes the volatility are also analyzed in Lopes and Prass (2009).

In the literature one can find different definitions of FIEGARCH processes. Definition 2.1 is the same as in Bollerslev and Mikkelsen (1996). In Zivot and Wang (2005), expression (2.2) is replaced by expression (2.4) in the definition of a FIEGARCH process. The following proposition shows that, under the restrictions given in (2.5), expressions (2.4) and (2.2) are equivalent. This result is crucial for a Monte Carlo simulation study (see Section 4).

PROPOSITION 2.1. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a* FIEGARCH(p, d, q) *process, given in Definition 2.1. Then, the expression* (2.2) *can be rewritten as*

$$\beta(\mathcal{B})(1-\mathcal{B})^d \ln(\sigma_t^2) = a + \sum_{i=0}^{p}\Big(\psi_i|Z_{t-1-i}| + \gamma_i Z_{t-1-i}\Big), \tag{2.4}$$

where $a \equiv (-\gamma)\,\alpha(1)\mathbb{E}\big(|Z_t|\big)$, $\psi_i = -\gamma\alpha_i$ *and* $\gamma_i = -\theta\alpha_i$, *for all* $0 \leq i \leq p$. (2.5)

Proof: See Lopes and Prass (2009). □

Clearly the definition given by Zivot and Wang (2005) is more general than the one presented by Bollerslev and Mikkelsen (1996). In fact, in the definition given by Zivot and Wang (2005), the coefficients ψ_j and γ_j, for $j = 0, 1, \cdots, p$, do not necessarily satisfy the restrictions given in (2.5).

3 Risk Measures

In this section we present the concept of risk factor, loss distribution and the definition of some risk measures and some approaches for estimating them. We also show that different approaches can lead to equivalent results depending on the assumptions made.

Risk measures are directly related to risk management. McNeil et al. (2005) classify the existing approaches to measuring the risk of a financial position in four different groups: *Notional-amount approach, Factor-sensitivity measures, Risk measures based on loss distribution and Scenario-based risk measures* (see McNeil et al., 2005, page 34). A frequent concept associated to risk is the volatility, which can have different definitions. The most common approach is to define the volatility as the variance (or the conditional variance) of the processes. In this paper we focus our attention to the variance, the Value-at-Risk (VaR_p) and Expected Shortfall (ES_p), which are risk measures based on loss distributions, and the Maximum-loss, a scenario-based risk measure.

Risk Factors

Consider a given portfolio and denote its value at time t by $V(t)$ (we assume that $V(t)$ is known at time t). For a given time horizon h we denote by L_{t+h} the loss of the portfolio in the period $[t, t+h]$, that is, $L_{t+h} := -(V(t+h) - V(t))$. The distribution of the random variables L_{t+h} is termed *loss distribution*. In risk management, the main concern is to analyze the probability of large losses, that is, the right tail of the loss distribution.

The usual approach is to assume that the random variable V_t, for all t, is a function of time and an m-dimensional random vector $\boldsymbol{Z}_t = (Z_{1,t}, \cdots, Z_{m,t})'$ of *risk factors*, that is, $V_t = f(t, \boldsymbol{Z}_t)$, for some measurable function $f : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$. Since we assume that the risk factors are observable, $\boldsymbol{Z}_t$ is known at time t. The choice of the risk factors and of the function $f(\cdot)$ is a modeling issue and depends on the portfolio and on the desired level of accuracy.

Remark 3.1. The random vector $\boldsymbol{Z} = (Z_1, \cdots, Z_m)'$ is also called a *scenario* which describes the situation of the market and, consequently, $f(\boldsymbol{Z})$ is referred to as the value of the portfolio under the scenario $\boldsymbol{Z}$.

In some cases it is more convenient to consider instead the time series of *risk-factors change* $\{\boldsymbol{X}_t\}_{t\in\mathbb{N}}$. This time series is defined by $\boldsymbol{X}_t \equiv \boldsymbol{Z}_t - \boldsymbol{Z}_{t-1}$, for all $t \in \mathbb{N}$. For instance, if we consider a portfolio with m stocks and $Z_{i,t} = \ln(P_{i,t})$, where $P_{i,t}$ is the price of the i-th asset at time t, then $\boldsymbol{X}_t = (X_{1,t}, \cdots, X_{m,t})'$ is the vector of log-returns. In this case, for $h = 1$, the value of the portfolio can be written as $L_{t+1} = -\big(f(t+1, \boldsymbol{Z}_t + \boldsymbol{X}_{t+1}) - f(t, \boldsymbol{Z}_t)\big)$, where $\boldsymbol{Z}_t$ is known at time t. The loss distribution is then determined by $\boldsymbol{X}_{t+1}$ risk-factor change distribution.

3.1 Risk Measures Based on Loss Distributions

The variance of the loss distribution is one of the most used risk measures. However, as a risk measure, the variance presents two problems. First of all one needs to assume that the loss distribution has finite second moment. Moreover, this measure does not

distinguish between positive and negative deviations of the mean. Therefore, the variance is a good risk measure only for distributions that are (approximately) symmetric, such as the Gaussian or t-Student (with finite variance) distributions.

Another common approach is the quantile analysis of the loss distribution. Consider a portfolio $\mathcal{P}$ with some risky assets and a fixed time horizon h. Let $F_L(\ell) \equiv \mathbb{P}(L \leq \ell)$ be the distribution of the associated loss. The main idea is to define a statistic based on $F_L(\cdot)$ capable of measuring the risk of the portfolio over a period h. Since for several models the support of $F_L(\cdot)$ is unbounded, the natural candidate (which is the maximum possible loss), given by $\inf\{\ell \in \mathbb{R} : F_L(\ell) = 1\}$, is not the best choice. The idea is to consider instead the "maximum loss which is not exceeded with a high probability". This probability is called *confidence level*.

DEFINITION 3.1. Let $\mathcal{P}$ be a fixed portfolio. Given a confidence level $\mathrm{p} \in (0,1)$, the *Value-at-Risk* of the portfolio, denoted by VaR_p, is defined as

$$\mathrm{VaR}_\mathrm{p} \equiv \inf\{\ell \in \mathbb{R} : \mathbb{P}(L \geq \ell) \leq 1 - \mathrm{p}\} = \inf\{\ell \in \mathbb{R} : F_L(\ell) \geq \mathrm{p}\}. \tag{3.1}$$

In a probabilistic sense, VaR_p is the p-quantile of the loss distribution function. For practical purposes, the most commonly used values are $\mathrm{p} \in \{0.95, 0.99\}$ and $h \in \{1, 10\}$ days.

As mentioned before, the risk analysis by considering the variance presents some drawbacks and so does the VaR_p. Artzner *et al.* (1999) define coherent risk measures and show that the VaR_p does not satisfy the subadditivity axiom. That is, given a fixed number of portfolios, the VaR_p of the sum of the portfolios may not be bounded by the sum of the VaR_p of the individual portfolios. This result contradicts the idea that the risks can be decreased by diversification, that is, by buying or selling financial assets.

In the following definition we present a coherent risk measure in the sense of Artzner *et al.* (1999) definition.

DEFINITION 3.2. Let L be a loss with distribution function $F_L(\cdot)$, such that $\mathbb{E}(|L|) < \infty$. The *Expected Shortfall*, denoted by ES_p, at confidence level $\mathrm{p} \in (0,1)$, is defined as

$$\mathrm{ES}_\mathrm{p} \equiv \frac{1}{1-\mathrm{p}} \int_\mathrm{p}^1 q_u(F_L) du,$$

where $q_u(\cdot)$ is the quantile function defined as $q_u(F_L) \equiv \inf\{\ell \in \mathbb{R} : F_L(\ell) \geq u\}$.

The risk measures ES_p and VaR_p are related by the expression

$$\mathrm{ES}_\mathrm{p} \equiv \frac{1}{1-\mathrm{p}} \int_\mathrm{p}^1 \mathrm{VaR}_u du$$

and it can be shown that, if L is integrable, with continuous distribution function $F_L(\cdot)$, then $\mathrm{ES}_\mathrm{p} = \mathbb{E}(L | L \geq \mathrm{VaR}_\mathrm{p})$ (see McNeil et al., 2005).

REMARK 3.2. In the literature one can find variations for the risk measure ES_p, given in the Definition 3.2, such as *tail conditional expectation* (TCE), *worst conditional expectation* (WCE) and *conditional* VaR (CVaR). Besides having slightly different definitions, all these risk measures are equivalent when the distribution function is continuous.

In practice, in order to calculate the VaR_p and ES_p values, one needs to estimate the loss distribution function $F_L(\cdot)$. Obviously, the use of different methods to estimate $F_L(\cdot)$ will lead to different values of those measures. The most common approaches to calculate the VaR_p are:

1. **Empirical VaR**. This is a non-parametric approach. The empirical VaR_p is the p-quantile of the empirical distribution function of the data. Under this approach, the VaR_p of h periods is the same as the VaR_p of 1 period.

2. **Normal VaR or Variance-Covariance Method**. Under this approach we assume that the data is normally distributed with mean and variance constants. The VaR_p is then the p-quantile of the Gaussian distribution. The mean and the variance (or the covariance matrix if the data is multidimensional) are estimated by their sample counter parts. For this approach it is also very common to assume that the conditional distribution function of the data is Gaussian instead of the distribution function itself.

3. **RiskMetrics Approach**. This methodology was developed by J.P. Morgan and it considers the conditional distribution function of the data. Consider first the case in which the portfolio has only one asset. Let r_t be the return (or log-return) of the portfolio at time t (the loss is then $-r_t$). The methodology assumes that

$$r_t|\mathcal{F}_{t-1} \sim \mathcal{N}(\mu_t, \sigma_t^2),$$

   where the conditional mean and variance are such that

$$\mu_t = 0 \quad \text{e} \quad \sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2, \quad 0 < \lambda < 1, \tag{3.2}$$

   that is, $\{\sigma_t^2\}_{t\in\mathbb{Z}}$ follows an exponentially weighted moving average (EWMA) model (see Roberts, 1959). Then, the VaR_p at time $t+1$, is the p-quantile of the Gaussian distribution with mean μ_t and variance σ_t^2.

   It can be shown that, using this method, the VaR_p for a period h is given by $\text{VaR}_{\text{p},t}[h] = \Phi^{-1}(\text{p})\sqrt{h}\sigma_{t+1}$, where $\Phi(\cdot)$ is the standard Gaussian distribution function and σ_{t+1}^2 is defined by the expression (3.2). It is easy to see that $\text{VaR}_{\text{p},t}[h] = \sqrt{h}\text{VaR}_{\text{p},t+1}$. However, if $\mu_t \neq 0$, in expression (3.2), this equality no longer holds.

   The multivariate case assumes that the conditional distribution function of the data is a multivariate normal one and $\text{Cov}(r_{i,t+1}, r_{j,t+1}) = \gamma_{ij,t+1}$ is estimated by the expression

$$\gamma_{ij,t} = \lambda\gamma_{ij,t-1} + (1-\lambda)r_{i,t-1}r_{j,t-1}, \quad \text{for } 0 < \lambda < 1.$$

   For more details see Zangari (1996).

4. **Econometric Approach**. This approach is similar to the RiskMetrics one. However, in this case, a more general class of models is considered. Generally, the time series mean is modeled by a linear model, such as the ARMA model, and the volatility is estimated by using a heteroskedastic model such as the FIEGARCH model defined in Section 2.

In the following proposition we present an expression for ES_p under the normality assumption.

PROPOSITION 3.1. *Let L be the random variable which represents the portfolio loss. If L has Gaussian distribution function with mean μ and variance σ^2 then,*

$$\mathrm{ES_p} = \mu + \sigma \frac{\phi\big(\Phi^{-1}(\mathrm{p})\big)}{1-\mathrm{p}}, \quad \textit{for all} \quad \mathrm{p} \in (0,1),$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the density and the distribution function of a standard normal random variable.

Proof: By setting $Z = \dfrac{L-\mu}{\sigma}$ and noticing that $\lim\limits_{z\to\infty} \dfrac{1}{\sqrt{2\pi}} e^{-z^2/2} = 0$, the proof follows directly from the fact that $\mathbb{P}(L \geq \mathrm{VaR_p}) = 1 - \mathrm{p}$ and $\mathrm{ES_p} = \mathbb{E}(L | L \geq \mathrm{VaR_p})$. □

3.2 A Scenario Based Risk Measure

The maximum loss, denoted by MaxLoss, introduced by Studer (1997), can be viewed either as a risk measure or as a systematic way of performing a stress test. This risk measure can be viewed as the worst possible loss. In many cases, the worst scenario may not exist since the function to calculate the value of a portfolio may be unbounded from below. It is known that the probability of scenarios occurrence which are very far away from the present market state is very low. Therefore, the idea is to restrict attention to scenarios under a certain admissibility domain, also denominated by *confidence region*, that is, a certain set of scenarios with high probability of occurrence. For example, if we assume that the data has an elliptic distribution function, such as the t-Student or the Gaussian distribution, then the admissibility domain is an ellipsoid (see Studer, 1997).

DEFINITION 3.3. Given an admissibility domain A, the *maximum loss* of a portfolio contained in A is given by

$$\mathrm{MaxLoss}_A(f) \equiv f(\boldsymbol{Z}_{AM}) - \min_{\boldsymbol{Z}\in A} \big\{ f(\boldsymbol{Z}) \big\},$$

where $f(\cdot)$ is the function that determines the portfolio's price and the vector $\boldsymbol{Z}_{AM} = (Z_{AM,1}, \cdots Z_{AM,m})'$ represents the m risk factors for the current market situation.

The maximum loss is a coherent risk measure in the same sense as defined by Artzner et al. (1999). Furthermore, the maximum loss gives not only the loss dimension but also the scenario in which this loss occurs.

REMARK 3.3.

(i) Note that, in order to compute the maximum loss, we need to set a closed confidence region A, with a certain probability p of occurrence. Then, an equivalent definition of maximum loss is the following

$$\mathrm{MaxLoss}_A(f) = \max \Big\{ f(\boldsymbol{Z}_{AM}) - f(\boldsymbol{Z}) : \boldsymbol{Z} \in A \text{ and } \mathbb{P}(A) = \mathrm{p} \Big\}.$$

(ii) Since $f(\cdot)$ gives the portfolio value, the expression $f(\boldsymbol{Z}_{AM}) - f(\boldsymbol{Z})$ represents the loss L (or $-L$ if $\boldsymbol{Z}$ is measured previously to AM) in the portfolio.

A portfolio is called *linear* if the loss L is a linear function with respect to each one of the risk-factors change. The following theorem gives the expression of the MaxLoss for a linear portfolio $\mathcal{P}$ with risk-factors change normally distributed.

THEOREM 3.1. *Let $\mathcal{P}$ be a linear portfolio and $f(\cdot)$ be the function that determines the portfolio value. Then, $f(\boldsymbol{X}) = \boldsymbol{a}'\boldsymbol{X}$, where $\boldsymbol{a} \in \mathbb{R}^m$ is a vector of real constants and $\boldsymbol{X} \in \mathbb{R}^m$, is the vector of risk-factors change. It follows that, given a confidence level* p*, the maximum loss of the portfolio is given by*

$$\text{MaxLoss} = -\sqrt{c_{\text{p}}}\sqrt{\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}}, \tag{3.3}$$

where $\boldsymbol{\Sigma}$ is the covariance matrix of the risk-factors change and c_{p} is the p*-quantile $\mathcal{X}^2_m$ distribution function with m degrees of freedom. Moreover, the worst scenario is given by*

$$\boldsymbol{Z}^* = -\frac{\sqrt{c_{\text{p}}}}{\sqrt{\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}}}\boldsymbol{\Sigma}\boldsymbol{a}.$$

Proof: See Theorem 3.15 in Studer (1997). □

REMARK 3.4. Studer (1997) considers the *Profit and Loss distribution* (P&L), which is the distribution of the random variable $-L$, instead of the loss distribution. However, both analysis lead to similar results. The only difference is which tail of the distribution is being analyzed. Considering this fact, notice that expression (3.3) is very similar to the expression for the Normal VaR_{p}, which is $\sqrt{z_{\text{p}}}\sqrt{\boldsymbol{a}'\boldsymbol{\Sigma}\boldsymbol{a}}$; the only difference lies in the scaling factor: c_{p} is the p-quantile of a $\mathcal{X}^2_m$ distribution with m degrees of freedom, whereas z_{p} is the p-quantile of a standard normal distribution. Contrary to the VaR_{p}, MaxLoss depends on the number of risk factors used in the model.

4 Simulation

In this section we present a simulation study related to the estimation of the volatility of the risk measures VaR_{p} on FIEGARCH(p, d, q) processes. A theoretical study related to the generation and the estimation of FIEGARCH(p, d, q) processes, considering the same set of parameters used here, can be found in Lopes and Prass (2009).

The simulation study considers five different models and the generated time series are the same ones used in Lopes and Prass (2009). The representation of the FIEGARCH process is the one proposed by Bollerslev and Mikkelsen (1996), given in Definition 2.1, where $Z_{i,t} \sim \mathcal{N}(0, 1)$, for $i \in \{1, \cdots, 5\}$. For each model we consider 1000 replications, with sample size $n \in \{2000, 5000\}$. The value $n = 2000$ was chosen since this is the approximated size of the observed time series considered in Section 5 of this paper. The value $n = 5000$ was chosen to analyze the asymptotic properties for the estimators. In the following, Mi, for $i \in \{1, \cdots, 5\}$, denotes the simulated FIEGARCH(p, d, q) model, that is,

M1: FIEGARCH(0,0.45,1) M2: FIEGARCH(1,0.45,1) M3: FIEGARCH(0,0.26,4)
M4: FIEGARCH(0,0.42,1) M5: FIEGARCH(0,0.34,1)

The parameters of the models used in the simulation study, are given in Table 4.1. These values are similar to those found in the analysis of the observed time series (see Section 5).

Table 4.1: Parameter Values for the Generated Models.

Model	ω	β_1	β_2	β_3	β_4	α_1	θ	γ	d
M1	0.00	0.45	-	-	-	-	-0.14	0.38	0.45
M2	0.00	0.90	-	-	-	0.80	0.04	0.38	0.45
M3	0.00	0.22	0.18	0.47	-0.45	-	-0.04	0.40	0.26
M4	0.00	0.58	-	-	-	-	-0.11	0.33	0.42
M5	0.00	0.71	-	-	-	-	-0.17	0.28	0.34

4.1 Volatility Estimation

In order to estimate the volatility Lopes and Prass (2009) used the *fgarch* function (from the S-Plus) to fit the FIEGARCH(p, d, q) models to the generated time series. The S-Plus code consider the expression (2.4) instead. For each time series only the $n - 10$ first values are considered, where n is the sample size. The remaining final 10 values are used to estimate the forecast error values. The mean square error (mse) is defined by $mse = \frac{1}{re}\sum_{t=1}^{re} e_t^2$, where $re = 1000$ is the number of replications and $e_t = \theta - \hat{\theta}$ represents the estimation error for the parameter θ, where θ is any parameter given in Table 4.1. For each model, the final value was obtained from the expression $\hat{\theta} = \frac{1}{re}\sum_{k=1}^{re}\hat{\theta}_k$, where $\hat{\theta}_k$ is the k-th estimator for θ in the k-th replication, for $k \in \{1, \cdots, re\}$.

Lopes and Prass (2009) compare the mean of 1000 generated values of $\sigma_{i,t+h}$ (a known parameter) and $X^2_{i,t+h}$, with the mean of the h-step ahead forecast values $\hat{\sigma}_{i,t+h}$ and $\hat{X}^2_{i,t+h}$, by calculating the mean square error (mse) values, for $t = n - 10$, $h \in \{1, \cdots, 10\}$, $n \in \{2000, 5000\}$ and $i = 1, \cdots, 5$. The authors observed that the higher the sample size, the higher the mean square forecast error. For σ_t (the square root of the volatility) the mean square forecast error values vary from 0.0037 to 0.3227 and, for X_t^2, they vary from 2.0725 to 12.1379.

4.2 $\mathrm{VaR_p}$ Estimation

In the following we present estimation results of the risk measure $\mathrm{VaR_p}$ for the generated time series.

In order to calculate the conditional mean and the conditional variance we use the $n - 10$ first values of the generated time series, where n is the sample size. For each one of the 1000 replications we used the approaches described in Subsection 3.1 to obtain the estimated $\mathrm{VaR_p}$. For the Econometric approach we consider EGARCH(p, q) models, with $p = 1 = q$ and the FIEGARCH(p, d, q) fitted to the time series in Lopes and Prass (2009), considering the same values of p and q used to generate the time series.

Remark 4.1. Since $Z_{i,t} \sim \mathcal{N}(0, 1)$, for $i \in \{1, \cdots, 5\}$, the true value of the $\mathrm{VaR}_{\mathrm{p},i,t+1}$, where i stands for the model and $t + 1$ for the period, is given by

$$\mathrm{VaR}_{\mathrm{p},i,t+1} = \Phi^{-1}(\mathrm{p}) \times \sigma_{i,t+1}, \quad i = 1, \cdots, 5, \tag{4.1}$$

where $t = n - 10$, $\Phi(\cdot)$ is the standard normal distribution function and $\sigma^2_{i,t+1}$ is the value of the conditional variance (volatility) generated by the model Mi, for $i = 1, \cdots, 5$.

Table 4.2 presents the true value, given by expression (4.1), and the estimated values of the risk measure $\mathrm{VaR_p}$, for $\mathrm{p} \in \{0.95; 0.99\}$. The values in this table are the mean taken over 1000 replications and $n \in \{2000, 5000\}$. The values for $n = 5000$ appear in parenthesis and mse represents the mean square error value. By comparing the different approaches, we observe that the mean of the $\mathrm{VaR_p}$ estimated values are very close from each other. In all cases, the mean of the estimated values is higher than the mean for the true value of this risk measure, either for $\mathrm{p} = 0.95$ or $\mathrm{p} = 0.99$. Also, there is little difference between the mean square error values when the sample size vary from $n = 2000$ to $n = 5000$. In most cases, the empirical approach leads to estimators with higher mean square error values. However, there is little difference among the values from the Econometric and Normal

approaches. Generally, those two approaches present better results than the RiskMetrics approach.

Lopes and Prass (2009) reported that the parameter estimation for the FIEGARCH models show coefficients with high mean square error values. The results presented in subsection 4.1, show a high mean square error value for the volatility estimation, which was expected since the volatility estimation depends on the parameter estimation value. As a consequence, although the underlying process is a FIEGARCH process, the Econometric approach using this model do not present better results.

Table 4.2: Mean Estimated Values of the Risk Measure $\mathrm{VaR_p}$ under Different Approaches for Sample Sizes $n = 2000$ and, in Parenthesis, $n = 5000$, with $\mathrm{p} \in \{0.95; 0.99\}$.

Approach	$\widehat{\mathrm{VaR}}_{0.95}$	mse	$\widehat{\mathrm{VaR}}_{0.99}$	mse
M1; $n = 2000$ ($n = 5000$); $\mathrm{VaR}_{0.95} = 1.6726$ (1.6673); $\mathrm{VaR}_{0.99} = 2.3656$ (2.3581)				
Empirical	1.7996 (1.7967)	**0.1378** (**0.1287**)	2.8355 (2.8577)	0.5257 (0.5206)
Normal	1.8458 (1.8433)	0.1534 (0.1453)	2.6108 (2.6072)	**0.3051** (**0.2895**)
Risk Metrics	1.7159 (1.6912)	0.2386 (0.2270)	2.4370 (2.4029)	0.4750 (0.4548)
EGARCH	1.7707 (1.7533)	0.1722 (0.1521)	2.5044 (2.4797)	0.3445 (0.3043)
FIEGARCH	1.7759 (1.7567)	0.1936 (0.1621)	2.5117 (2.4845)	0.3872 (0.3243)
M2; $n = 2000$ ($n = 5000$); $\mathrm{VaR}_{0.95} = 1.6765$ (1.6691); $\mathrm{VaR}_{0.99} = 2.3710$ (2.3607)				
Empirical	1.7773 (1.7689)	0.1014 (0.1094)	2.7020 (2.7229)	0.3285 (0.3625)
Normal	1.8004 (1.7938)	0.1077 (0.1170)	2.5462 (2.5369)	0.2152 (0.2337)
Risk Metrics	1.7287 (1.7192)	0.1353 (0.1352)	2.4425 (2.4291)	0.2694 (0.2697)
EGARCH	1.7494 (1.7342)	**0.0925** (**0.0902**)	2.4743 (2.4527)	**0.1850** (**0.1805**)
FIEGARCH	1.7499 (1.7439)	0.1035 (0.1114)	2.4749 (2.4664)	0.2071 (0.2228)
M3; $n = 2000$ ($n = 5000$); $\mathrm{VaR}_{0.95} = 1.6517$ (1.6594); $\mathrm{VaR}_{0.99} = 2.3360$ (2.3469)				
Empirical	1.6852 (1.6968)	**0.0326** (**0.0326**)	2.5002 (2.5208)	0.0965 (0.0967)
Normal	1.7028 (1.7129)	0.0330 (0.0339)	2.4086 (2.4224)	**0.0654** (**0.0674**)
Risk Metrics	1.6864 (1.6805)	0.0683 (0.0598)	2.3858 (2.3778)	0.1349 (0.1190)
EGARCH	1.6886 (1.6976)	0.0414 (0.0359)	2.3882 (2.4009)	0.0828 (0.0717)
FIEGARCH	1.6895 (1.6950)	0.0466 (0.0382)	2.3894 (2.3973)	0.0932 (0.0765)
M4; $n = 2000$ ($n = 5000$); $\mathrm{VaR}_{0.95} = 1.6637$ (1.6691); $\mathrm{VaR}_{0.99} = 2.3360$ (2.3607)				
Empirical	1.7643 (1.7967)	**0.1194** (**0.1197**)	2.7593 (2.8391)	0.4399 (0.4714)
Normal	1.8064 (1.8383)	0.1330 (0.1340)	2.5549 (2.5998)	0.4016 (0.2667)
Risk Metrics	1.6987 (1.7078)	0.1958 (0.1901)	2.4115 (2.4250)	0.3912 (0.3805)
EGARCH	1.7478 (1.7610)	0.1327 (0.1314)	2.4719 (2.4906)	**0.2654** (**0.2629**)
FIEGARCH	1.7546 (1.7656)	0.1559 (0.1512)	2.4816 (2.4971)	0.3117 (0.3025)
M5; $n = 2000$ ($n = 5000$); $\mathrm{VaR}_{0.95} = 1.6763$ (1.6484); $\mathrm{VaR}_{0.99} = 2.3708$ (2.3314)				
Empirical	1.8038 (1.8001)	**0.1473** (**0.1237**)	2.9232 (2.8391)	0.6476 (0.5871)
Normal	1.8672 (1.8593)	0.1708 (0.1464)	2.6411 (2.6298)	**0.3384** (0.2917)
Risk Metrics	1.7402 (1.7078)	0.2926 (0.2636)	2.4790 (2.3665)	0.5869 (0.5277)
EGARCH	1.7863 (1.7200)	0.1780 (0.1314)	2.5264 (2.4906)	0.3560 (**0.2726**)
FIEGARCH	1.7959 (1.7242)	0.2088 (0.1557)	2.5400 (2.4386)	0.4176 (0.3114)

In the following we assume that the simulated time series represent log-returns and we denote by $\{r_{i,t}\}_{t=1}^{n}$ the time series generated from the model Mi, for $i = 1, \cdots, 5$, where n is the sample size.

In practice, the volatility is not observable. Therefore, it is not possible to calculate the true value of the risk measure $\mathrm{VaR_p}$ and the mean square error value for its estimation. Recall that $\mathrm{VaR}_{\mathrm{p},i,t+1}$, $i = 1, \cdots, 5$, represents the maximum loss that can occur with a given probability p in the instant $t+1$. Also, $-r_{i,t+1}$, $i = 1, \cdots, 5$, can be understood as the loss at time $t+1$. The usual approach is then to compare $\mathrm{VaR_p}$ estimated values with the value of the observed log-returns (in our case, the simulated time series).

Table 4.3 presents the mean value of $-r_{i,t+1}$ (which is known), for $i = 1, \cdots, 5$, and the mean values of the estimated $\mathrm{VaR}_{\mathrm{p},i,t+1}$, for $t = n-10$, where n is the sample size and $\mathrm{p} \in \{0.95; 0.99\}$. We observe that, for each model and each p value, different approaches lead to similar results. In most cases, the value of the risk measure $\mathrm{VaR_p}$ estimated under the RiskMetrics approach presents the smallest mean square error value. Also, the values estimated under this approach are closer to the observed values (the log-returns) than the ones estimated by the Empirical and Normal approaches. By comparing both, the RiskMetrics and Econometric approaches we observe almost no difference between the estimated values. As in the previous case, we need to take into account that the estimation of the FIEGARCH models has strong influence in the results.

Table 4.3: Mean Square Error Values for the Risk Measure $\mathrm{VaR_p}$ under Different Approaches for Sample Sizes $n = 2000$ and, in Parenthesis, $n = 5000$, with $\mathrm{p} \in \{0.95; 0.99\}$.

Approach		Empirical	Normal	RiskMetrics	EGARCH	FIEGARCH
$r_{1.t+1}$ = 0.0649 (-0.0185)	p = 0.95	4.5099 (4.7085)	4.6781 (4.8815)	4.4996 (**4.6719**)	**3.5713** (4.8073)	3.6421 (4.8570)
	p = 0.99	4.5099 (4.7085)	4.6781 (4.8815)	4.4996 (**4.6719**)	**3.5713** (4.8073)	3.6421 (4.8570)
$r_{2.t+1}$ = -0.0015 (0.0424)	p = 0.95	4.4530 (**4.2751**)	4.5403 (4.3669)	4.5726 (4.3423)	**3.3268** (4.3238)	3.3688 (4.4094)
	p = 0.99	4.5099 (4.7085)	4.6781 (4.8815)	4.4996 (**4.6719**)	**3.5713** (4.8073)	3.6421 (4.8570)
$r_{3.t+1}$ = -0.0380 (0.0353)	p = 0.95	4.0809 (**3.8915**)	4.1427 (3.9452)	4.1591 (3.9260)	**2.9299** (3.9708)	2.9587 (3.9892)
	p = 0.99	4.5099 (4.7085)	4.6781 (4.8815)	4.4996 (**4.6719**)	**3.5713** (4.8073)	3.6421 (4.8570)
$r_{4.t+1}$ = -0.0369 (-0.0103)	p = 0.95	4.6457 (**4.6180**)	4.8021 (4.7760)	4.7666 (4.7099)	**3.4045** (4.8067)	3.4854 (4.8756)
	p = 0.99	4.5099 (4.7085)	4.6781 (4.8815)	4.4996 (**4.6719**)	**3.5713** (4.8073)	3.6421 (4.8570)
$r_{5.t+1}$ = -0.0589 (0.0525)	p = 0.95	4.9339 (**4.3492**)	5.1815 (4.5625)	5.2978 (4.3588)	**3.6491** (4.3689)	3.7537 (4.4497)
	p = 0.99	4.5099 (4.7085)	4.6781 (4.8815)	4.4996 (**4.6719**)	**3.5713** (4.8073)	3.6421 (4.8570)

Note: The mean square error value is measured with respect to the log-returns instead of the true $\mathrm{VaR_p}$.

5 Analysis of Observed Time Series

In this section we present the estimation and the analysis of risk measures for a portfolio $\mathcal{P}$ of stocks. This portfolio is composed by stocks of four Brazilian companies. These assets are denoted by A_i, $i = 1, \cdots, 4$, where:

A_1: represents the Bradesco stocks. A_2: represents the Brasil Telecom stocks.

A_3: represents the Gerdau stocks. A_4: represents the Petrobrás stocks.

These stocks are negotiated in the Brazilian stock market, that is, in the São Paulo Stock, Mercantile & Futures Exchange (Bovespa). The notation A_M (or, equivalently, A_5) is used to denote the financial market. The market portfolio values are represented by the São Paulo Stock Exchange Index (Bovespa Index or IBovespa).

Prass and Lopes (2009) present a comparison study on risk analysis using CAPM model, VaR_p and MaxLoss on FIEGARCH processes. They consider the same portfolio $\mathcal{P}$ considered here and they also calculate the vector $a = (a_1, a_2, a_3, a_4)$ of weights for this portfolio. The same weights found by Prass and Lopes (2009) are considered in this paper, that is, $(a_1, a_2, a_3, a_4) = (0.3381, 0.1813, 0.3087, 0.1719)$.

In the following we fixed:

- c_i is the number of stocks of the asset A_i, $i = 1, \cdots, 4$. It follows that $c_i = \dfrac{a_i V_0}{P_{i,0}}$, where V_0 is the initial capital invested in the portfolio, $P_{i,0}$ is the unitary price of the stock at the initial time, and $\boldsymbol{a} = (a_1, a_2, a_3, a_4)'$ are the weights of the assets. The value of the portfolio $\mathcal{P}$, at time t, is then given by

$$V_t = V_0 \left(\frac{a_1}{P_{1,0}} P_{1,t} + \frac{a_2}{P_{2,0}} P_{2,t} + \frac{a_3}{P_{3,0}} P_{3,t} + \frac{a_4}{P_{4,0}} P_{4,t} \right), \quad t = 1, \cdots, n,$$

  where n is the sample size;

- for this portfolio we assume that the initial time is the day of the first observation;

- the risk-factors for this portfolio $\mathcal{P}$ are the logarithm of the prices of the assets. That is, the risk-factors vector is given by

$$\boldsymbol{Z}_t = (Z_{1,t}, Z_{2,t}, Z_{3,t}, Z_{4,t})' = (\ln(P_{1,t}), \ln(P_{2,t}), \ln(P_{3,t}), \ln(P_{4,t}))',$$

  where $P_{i,t}$ is the price of the asset A_i, $i = 1, \cdots, 4$, at time t. It follows that the risk-factors change is given by

$$\boldsymbol{X}_t = (X_{1,t}, X_{2,t}, X_{3,t}, X_{4,t})' = (r_{1,t}, r_{2,t}, r_{3,t}, r_{4,t})',$$

  where $r_{i,t}$, is the log-return of the asset A_i, $i = 1, \cdots, 4$, at time t;

- the loss of the portfolio $\mathcal{P}$, at time t, is the random variable L_t given by

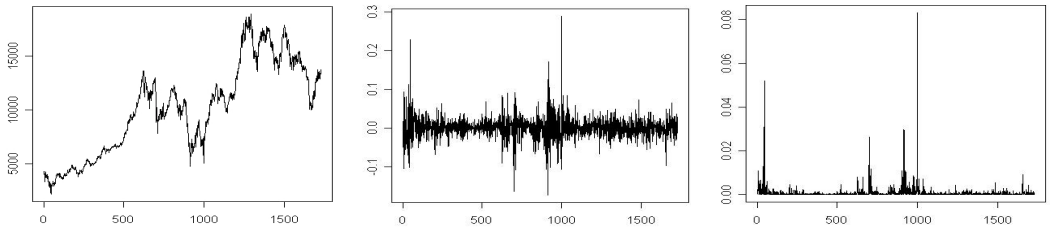
$$L_t = -V_{t-1} \sum_{i=1}^{4} a_i R_{i,t} \simeq -V_{t-1} \sum_{i=1}^{4} a_i r_{i,t} = -V_{t-1} r_{\mathcal{P},t}, \tag{5.1}$$

  where V_{t-1} is the value of $\mathcal{P}$, at time $t-1$, $\boldsymbol{a} = (a_1, a_2, a_3, a_4)'$ is the vector of weights, $R_{i,t} = \dfrac{P_t - P_{t-1}}{P_t}$, $r_{i,t}$ is the log-return of the asset A_i at time t and $r_{\mathcal{P},t}$ is the log-return of the portfolio, at time t;

- for VaR_p and ES_p estimation, in all cases (univariate or multivariate), we assume normality (or conditional normality) of the data.

5.1 Characteristics of the Observed Time Series

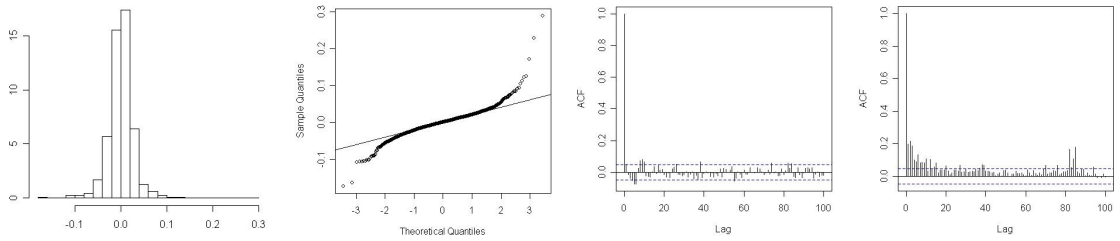
Figure 5.1 presents the time series with $n = 1729$ observations of the São Paulo Stock Exchange Index (Bovespa Index or IBovespa) in the period of January, 1995 to December, 2001, the IBovespa log-return series and the square of the log-return series. Observe that the log-return series presents the stylized facts mentioned in Section 2, such as apparently stationarity, mean around zero and clusters of volatility. Also, in Figure 5.2 we observe that, while the log-return series presents almost no correlation, the sample correlation of the square of the log-return series assumes high values for several lags, pointing to the existence of both heteroskedasticity and long memory characteristics. Regarding the histogram and the QQ-Plot, we observe that the distribution of the log-return series seems approximately symmetric and leptokurtic.



(a) (b) (c)

Figure 5.1: (a) São Paulo Stock Exchange Index (Bovespa Index or IBovespa) in the Period of January, 1995 to December, 2001; (b) IBovespa Log-returns and (c) Square of the IBovespa Log-returns.

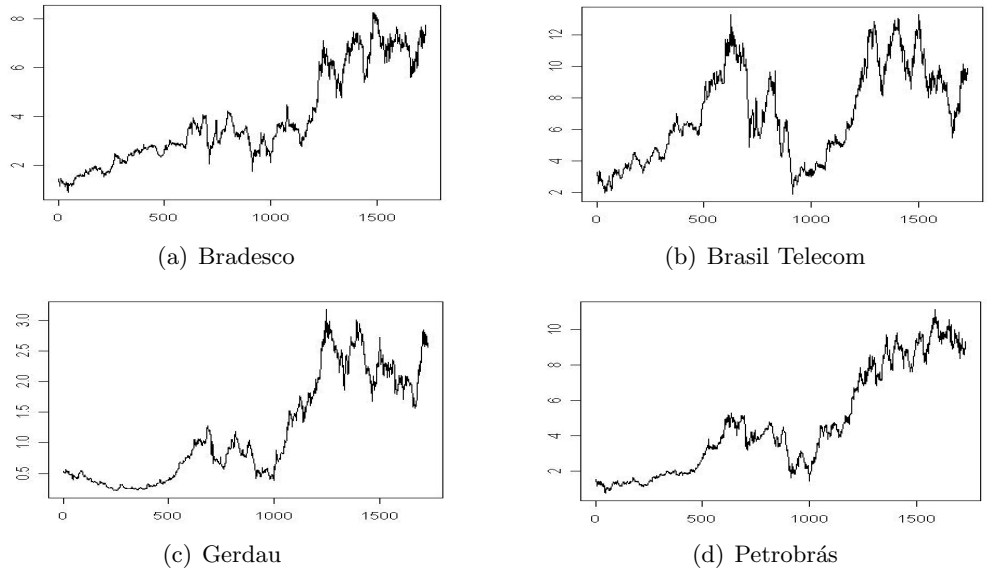
Figure 5.3 presents the stock prices of the companies Bradesco, Brasil Telecom, Gerdau and Petrobrás, in the period of January, 1995 to December, 2001, with $n = 1729$ observations. These time series present the same characteristics observed in the IBovespa time series (and log-return time series). In Figures 5.1 and 5.3 we observe that both, the market index and the stock prices, present a strong decay in their values close to $t = 1000$ (January 15, 1999). This period is characterized by the Real (the Brazilian currency) devaluation.



(a) (b) (c) (d)

Figure 5.2: (a) Histogram; (b) QQ-Plot and (c) Sample Autocorrelation of the IBovespa Log-return series and (d) Sample Autocorrelation of the Square of the IBovespa Log-return series.

Remark 5.1. The presence of long memory was also tested by analyzing the periodogram



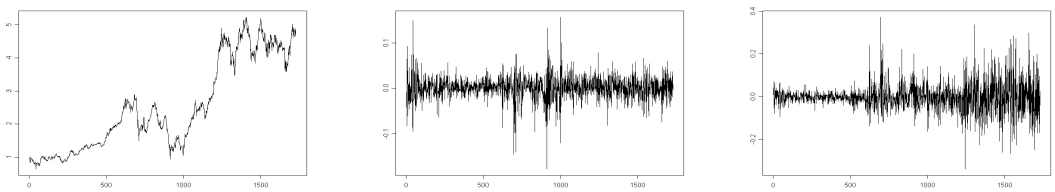
(a) Bradesco (b) Brasil Telecom

(c) Gerdau (d) Petrobrás

Figure 5.3: Time Series of the Stock Prices in the Period of January, 1995 to December, 2001.

of the square of the log-return time series and by using some known hypothesis test such as GPH test, R/S, modified R/S, V/S and KPSS statistics. All these tests confirm the existence of long memory characteristics.

Figure 5.4 presents the time series $\{V_t\}_{t=1}^{n}$ of the values of the portfolio $\mathcal{P}$ and the portfolio log-returns series in the period of January, 1995 to December, 2001. By comparing the portfolio values with the market index time series we observe a similar behaviour.



(a) (b) (c)

Figure 5.4: (a) Time Series of the Portfolio $\mathcal{P}$ Values; (b) Log-returns and (c) Portfolio Loss for the Period of January, 1995 to December, 2001.

The time series of the portfolio loss is shown in Figure 5.4 (c). We observe that the highest loss occurred at $t = 698$,with $L_t = 0.3729$ (this means that the highest earning value is approximately equal to R$ 0.37 per Real invested) and the smallest loss occurred at $t = 1244$, with $L_t = -0.3402$ (that is, approximately R$ 0.34 per each Real invested). The highest loss corresponds to the change in the value of the portfolio $\mathcal{P}$ from 10/24/1997 (Friday) to 10/27/1997 (Monday). In this period the Bovespa index changed

from 11,545.20 to 9,816.80 points, which represents a drop of 14.97%. In this same date we also observed high drops in the Dow Jones (7.18%) and S&P 500 (6.87%) indexes. This period coincides with the crises in Asia. The highest earning in the portfolio $\mathcal{P}$ corresponds to the change in the value of the portfolio from 01/13/2001 to 01/14/2001. The IBovespa showed an increasing of 2.08% in this period. In the date 01/14/2000 the Dow Jones index was 11,722.98 points. This value only was surpassed in 10/03/2006, when the index reached 11,727.34 points.

5.2 Fitting the FIEGARCH Models

In the following we present the models fitted to the observed time series. The selection of the class of models used, the ARMA-FIEGARCH models, was based on the analysis of the sample autocorrelation and on periodogram functions and results from the long memory tests. The fitted models are then used to estimate the volatility and consequently the risk measures for these processes.

In all cases considered, the analysis of the sample autocorrelation function suggests an ARMA(p_1, q_1)-FIEGARCH(p_2, d, q_2) model. The ARMA models are used to model the correlation among the log-returns while the FIEGARCH models take into account the long memory and the heteroskedasticity characteristics of the time series.

In order to estimate the parameters of the model we consider the *fgarch* function from S-PLUS software. We consider $p_1, q_1 \in \{0, 1, 2, 3\}$ e $p_2, q_2 \in \{0, 1\}$. The selection of the final model was based on the values of the log-likelihood and on the AIC and BIC criteria. The residual analysis indicated that none of these models were adequated for the Gerdau log-return time series since the square of the log-returns still presented correlation. The problem was solved by choosing a FIEGARCH model with $p_2 = 4$. After the residual analysis, the following models were selected:

- For the IBovespa log-returns $\{r_{M,t}\}_{t=1}^{1728}$: ARMA$(0, 1)$-FIEGARCH$(0, 0.339, 1)$,

$$\begin{aligned} r_{M,t} &= X_{M,t} - 0.078X_{M,t-1} \\ X_{M,t} &= \sigma_{M,t}Z_{M,t} \\ (1 - 0.706\mathcal{B})(1 - \mathcal{B})^{0.339}\ln(\sigma^2_{M,t}) &= -0.346 + 0.275|Z_{M,t-1}| - 0.166Z_{M,t-1}. \end{aligned}$$

- For the Bradesco log-returns $\{r_{1,t}\}_{t=1}^{1728}$: ARMA$(0, 1)$-FIEGARCH$(0, 0.446, 1)$,

$$\begin{aligned} r_{1,t} &= X_{1,t} - 0.129X_{1,t-1} \\ X_{1,t} &= \sigma_{1,t}Z_{1,t} \\ (1 - 0.453\mathcal{B})(1 - \mathcal{B})^{0.446}\ln(\sigma^2_{1,t}) &= -0.374 + 0.381|Z_{1,t-1}| - 0.135Z_{1,t-1}. \end{aligned}$$

- For the Brasil Telecom log-returns $\{r_{2,t}\}_{t=1}^{1728}$: ARMA$(0, 1)$-FIEGARCH$(1, 0.447, 1)$,

$$\begin{aligned} r_{2,t} &= 0.002 + X_{2,t} - 0.103X_{2,t-1} \\ X_{2,t} &= \sigma_{2,t}Z_{2,t} \\ (1 - 0.905\mathcal{B})(1 - \mathcal{B})^{0.447}\ln(\sigma^2_{2,t}) &= -0.053 + 0.382|Z_{2,t-1}| - 0.331|Z_{2,t-2}| \\ &\quad +0.044Z_{2,t-1} - 0.066Z_{2,t-2}. \end{aligned}$$

- For the Gerdau log-returns $\{r_{3,t}\}_{t=1}^{1728}$: ARMA$(1, 0)$-FIEGARCH$(0, 0.256, 4)$,

$$\begin{aligned} r_{1,t} &= -0.1409r_{3,t-1} + X_{3,t} \\ X_{3,t} &= \sigma_{3,t}Z_{3,t} \\ \beta(\mathcal{B})(1 - \mathcal{B})^{0.256}\ln(\sigma^2_{3,t}) &= -0.769 + 0.395|Z_{3,t-1}| - 0.046Z_{3,t-1}, \end{aligned}$$

  where $\beta(\mathcal{B}) = 1 - 0.216\mathcal{B} - 0.184\mathcal{B}^2 - 0.470\mathcal{B}^3 + 0.450\mathcal{B}^4$.

- For the Petrobrás log-returns $\{r_{4,t}\}_{t=1}^{1728}$: ARMA$(0,1)$-FIEGARCH$(0, 0.416, 1)$,

$$\begin{aligned} r_{4,t} &= 0.001 + X_{4,t} - 0.119X_{4,t-1} \\ X_{4,t} &= \sigma_{4,t}Z_{4,t} \\ (1-0.575\mathcal{B})(1-\mathcal{B})^{0.416}\ln(\sigma_{4,t}^2) &= -0.347 + 0.326|Z_{4,t-1}| - 0.110Z_{4,t-1}. \end{aligned}$$

- For the portfolio log-returns $\{r_{\mathcal{P},t}\}_{t=1}^{1728}$: ARMA$(1,0)$-FIEGARCH$(0, 0.233, 1)$,

$$\begin{aligned} r_{\mathcal{P},t} &= -0.001 - 0.173r_{\mathcal{P},t-1} + X_{\mathcal{P},t} \\ X_{\mathcal{P},t} &= \sigma_{\mathcal{P},t}Z_{\mathcal{P},t} \\ (1-0.754\mathcal{B})(1-\mathcal{B})^{0.233}\ln(\sigma_{\mathcal{P},t}^2) &= -0.498 + 0.285|Z_{\mathcal{P},t-1}| + 0.127Z_{\mathcal{P},t-1}. \end{aligned}$$

5.3 Conditional Mean and Volatility Forecast

In order to estimate the risk measures presented in Section 3 we first consider the estimation of the conditional mean and the volatility, that is, the conditional standard deviation of the log-return time series. Theoretical results regarding forecast on ARMA and FIEGARCH models can be found, respectively, in Brockwell and Davis (1991) and Lopes and Prass (2009).

The forecast for the conditional mean, $\hat{r}_{i,n+h}$, and for the volatility, $\hat{\sigma}_{i,n+h}$, for $i = 1, \cdots, 5$ and $h = 1, \cdots, 10$ are presented in Table 5.1. We observe that, for $h > 1$, the value of $\hat{\sigma}_{M,n+h}$ is constant, which does not occur for the assets A_i, $i = 1, \cdots, 4$. Apparently, for A_1 and A_4 the forecast values increase, while for A_2 they decrease. For A_3 the behaviour seems to be random.

Table 5.1: Forecast Values of the Conditional Mean and Volatility for the Log-return Time Series of IBovespa (A_M), Bradesco (A_1), Brasil Telecom (A_2), Gerdau (A_3) and Petrobrás (A_4), for $h = 1, \cdots, 10$.

	A_M		A_1		A_2		A_3		A_4	
h	$\hat{r}_{M.n+h}$	$\hat{\sigma}_{M.n+h}$	$\hat{r}_{1.n+h}$	$\hat{\sigma}_{1.n+h}$	$\hat{r}_{2.n+h}$	$\hat{\sigma}_{2.n+h}$	$\hat{r}_{3.n+h}$	$\hat{\sigma}_{3.n+h}$	$\hat{r}_{4.n+h}$	$\hat{\sigma}_{4.n+h}$
1	-0.0012	0.0179	0.0031	0.0215	-0.0030	0.0319	-0.0041	0.0256	0.0019	0.0176
2	0.0000	0.0182	0.0000	0.0221	0.0018	0.0219	0.0000	0.0248	0.0014	0.0179
3	0.0000	0.0182	0.0000	0.0226	0.0018	0.0194	0.0000	0.0195	0.0014	0.0185
4	0.0000	0.0182	0.0000	0.0230	0.0018	0.0181	0.0000	0.0293	0.0014	0.0190
5	0.0000	0.0182	0.0000	0.0233	0.0018	0.0175	0.0000	0.0235	0.0014	0.0194
6	0.0000	0.0182	0.0000	0.0236	0.0018	0.0171	0.0000	0.0219	0.0014	0.0197
7	0.0000	0.0182	0.0000	0.0238	0.0018	0.0169	0.0000	0.0280	0.0014	0.0200
8	0.0000	0.0182	0.0000	0.0239	0.0018	0.0168	0.0000	0.0219	0.0014	0.0202
9	0.0000	0.0182	0.0000	0.0241	0.0018	0.0169	0.0000	0.0233	0.0014	0.0204
10	0.0000	0.0182	0.0000	0.0242	0.0018	0.0170	0.0000	0.0261	0.0014	0.0206

The forecast for the conditional mean, $\hat{r}_{\mathcal{P},n+h}$, and for the volatility, $\hat{\sigma}_{\mathcal{P},n+h}$, of the portfolio $\mathcal{P}$ are presented in Table 5.2. Note that, while $\hat{r}_{\mathcal{P},n+h}$ is constant for $h > 2$, $\hat{\sigma}_{\mathcal{P},n+h}$ is slowly increasing.

Table 5.2: Forecast Values of the Conditional Mean and Volatility for the Portfolio Log-returns, for $h = 1, \cdots, 10$.

h	$\hat{r}_{\mathcal{P},t+h}$	$\hat{\sigma}_{\mathcal{P},t+h}$	h	$\hat{r}_{\mathcal{P},t+h}$	$\hat{\sigma}_{\mathcal{P},t+h}$
1	0.0001	0.0149	6	-0.0010	0.0161
2	-0.0008	0.0152	7	-0.0010	0.0162
3	-0.0010	0.0155	8	-0.0010	0.0164
4	-0.0010	0.0157	9	-0.0010	0.0165
5	-0.0010	0.0159	10	-0.0010	0.0166

5.4 $\mathrm{VaR_p}$ and $\mathrm{ES_p}$ Estimation

We considered two different approachs in the analysis of the risk measures. The first one considers the log-return series of the portfolio $\mathcal{P}$ (see Palaro and Hotta, 2006) and its loss distribution. We consider either, the conditional (RiskMetrics approach) and the unconditional (variance-covariance method) distribution of the risk-factors change in order to estimate the risk measures $\mathrm{VaR_p}$ and $\mathrm{ES_p}$. As a second approach we calculate the risk measures for each one of the assets in the portfolio. Since $\mathrm{ES_p}$ is a coherent risk measure and $\mathrm{VaR_p} \leq \mathrm{ES_p}$, by calculating $\mathrm{ES_p}$ we found an upper bound for the $\mathrm{VaR_p}$ of the portfolio $\mathcal{P}$.

Table 5.3 presents the estimated values of $\mathrm{VaR_p}$ and $\mathrm{ES_p}$ for the portfolio log-return time series. The observed values, at time $n+1$, of the assets log-returns are, respectively, -0.0026, 0.0301, 0.0680 and 0.0021. Therefore, the portfolio log-return value, at this time, is 0.0259 (without loss of generality, we assume $V_n = 1$. The loss is then given by a $L_{n+1} = -V_n \times r_{\mathcal{P},n+1} = -0.0259$). By comparing this value with the estimated values given in the Table 5.3 we observe that the Econometric approach, using FIEGARCH model, presents the best performance.

Table 5.3: Estimated Values of the Risk Measures $\mathrm{VaR_p}$ and $\mathrm{ES_p}$ for the Portfolio Log-return Time Series, at Confidence Level p = 95% and, in Parenthesis, p = 99%, for $h = 1$ day (Univariate Case).

Approach	$\mathrm{VaR}_{\mathrm{p}_{\mathrm{p},n+1}}$	$\mathrm{ES}_{\mathrm{p}_{\mathrm{p},n+1}}$
Empirical	0.0369 (0.0703)	0.0588 (0.0966)
Normal	0.0398 (0.0566)	0.0712 (0.0923)
RiskMetrics	0.0321 (0.0461)	0.0583 (0.0759)
EGARCH	0.0353 (0.0499)	0.0625 (0.0808)
FIEGARCH	0.0247 (0.0349)	0.0437 (0.0564)

Table 5.4 presents the results obtained by considering the multivariate distribution function of the risk-factor changes $\boldsymbol{X}_t = (r_{1,t}, r_{2,t}, r_{3,t}, r_{4,t})'$. We consider the unconditional (Normal approach) and the conditional distribution (RiskMetrics approach). By comparing the results in Tables 5.3 and 5.4 we observe that while for the Normal approach the values are the same, for the RiskMetrics approach the estimated value obtained using the univariate distribution was closer to the negative value of the observed log-return.

Tables 5.5 and 5.6 present the estimated values of the risk measures $\mathrm{VaR_p}$ and $\mathrm{ES_p}$ obtained by considering the univariate distribution function of each one of the risk-factor

Table 5.4: Estimated Values of the Risk Measures $\mathrm{VaR}_{\mathrm{p}}$ and ES_{p} for the Portfolio Log-return Time Series, at Confidence Level p = 95% and, in Parenthesis, p = 99%, for $h = 1$ day (Multivariate Case).

Approach	$\mathrm{VaR}_{\mathrm{p}_{\mathrm{p},n+1}}$	$\mathrm{ES}_{\mathrm{p}_{\mathrm{p},n+1}}$
Normal	0.0398 (0.0566)	0.0712 (0.0923)
RiskMetrics	0.2131 (0.3018)	0.3788 (0.4898)

changes $r_{1,t}, r_{2,t}, r_{3,t}, r_{4,t}$. Upon comparison of the values in Tables 5.5 and 5.6 we observe that, for this portfolio $\mathcal{P}$, both inequalities are satisfied:

$$\mathrm{VaR}_{\mathrm{p},\mathcal{P},n+1} \leq \sum_{i=1}^{4} a_i \mathrm{VaR}_{\mathrm{p},i,n+1} \quad \text{and} \quad \mathrm{ES}_{\mathrm{p},\mathcal{P},n+1} \leq \sum_{i=1}^{4} a_i \mathrm{ES}_{\mathrm{p},i,n+1}.$$

Also, the $\mathrm{VaR}_{\mathrm{p}}$ estimated by Econometric approach using FIEGARCH processes were the ones closer to the observed loss $-r_{i,1729}$, $i = 1, \cdots, 4$, given by 0.0026, −0.0301, −0.0680 and −0.0021. It is easy to see that, in all cases, the loss was superestimated. This fact is well known and discussed in the literature. This occurs because of the normality assumption in the risk measure estimation. Khindanova and Atakhanov (2002) presents a comparison study which demonstrate that stable modeling captures asymmetry and heavy-tails of returns, and, therefore, provides more accurate estimates of $\mathrm{VaR}_{\mathrm{p}}$.

Table 5.5: $\mathrm{VaR}_{\mathrm{p}}$ Estimated Values for the Assets Log-return Time Series, at Confidence Level p = 95% and, in Parenthesis, p = 99%, for $h = 1$ day.

Approach	$\mathrm{VaR}_{\mathrm{p},1,n+1}$	$\mathrm{VaR}_{\mathrm{p},2,n+1}$	$\mathrm{VaR}_{\mathrm{p},3,n+1}$	$\mathrm{VaR}_{\mathrm{p},4,n+1}$	$\sum_{i=1}^{4} a_i \mathrm{VaR}_{\mathrm{p},i,n+1}$
Empirical	0.0427 (0.0785)	0.0508 (0.1023)	0.0494 (0.0870)	0.0483 (0.0905)	0.0472 (0.0875)
Normal	0.0487 (0.0693)	0.0582 (0.0825)	0.0535 (0.0761)	0.0539 (0.0767)	0.0528 (0.0750)
RiskMetrics	0.2648 (0.3751)	0.3056 (0.4327)	0.3024 (0.4281)	0.2511 (0.3557)	0.2814 (0.3985)
EGARCH	0.0370 (0.0536)	0.0633 (0.0875)	0.0565 (0.0783)	0.0342 (0.0489)	0.0473 (0.0666)
FIEGARCH	0.0385 (0.0531)	0.0495 (0.0712)	0.0380 (0.0555)	0.0308 (0.0428)	0.0390 (0.0554)

5.5 MaxLoss Estimation

Since the considered portfolio $\mathcal{P}$ is linear, from Theorem 3.1, given a confidence level p, the MaxLoss is estimated by the expression (3.3).

Table 5.7 presents the MaxLoss values under different values of p and the scenarios under which this loss occurs. By definition, $r_{i,ML}$, $i = 1, \cdots, 4$ is the log-return of the asset A_i under the MaxLoss scenario. By comparing the values in Table 5.7 with those ones found in the previous analysis, we observe that the loss estimated under the scenario

Table 5.6: ES_p Estimated Values for the Assets Log-return Time Series, at Confidence Level p = 95% and, in Parenthesis, p = 99%, for $h = 1$ day.

Approach	$ES_{p,1,n+1}$	$ES_{p,2,n+1}$	$ES_{p,3,n+1}$	$ES_{p,4,n+1}$	$\sum_{i=1}^{4} a_i ES_{p,i,n+1}$
Empirical	0.0672 (0.1133)	0.0731 (0.1163)	0.0623 (0.1049)	0.0679 (0.1202)	0.0669 (0.1124)
Normal	0.0871 (0.1128)	0.1037 (0.1341)	0.0956 (0.1238)	0.0964 (0.1249)	0.0943 (0.1221)
RiskMetrics	0.4707 (0.6086)	0.5428 (0.7017)	0.5370 (0.6942)	0.4464 (0.5771)	0.5001 (0.6465)
EGARCH	0.0680 (0.0887)	0.1085 (0.1388)	0.0973 (0.1246)	0.0616 (0.0799)	0.0833 (0.1074)
FIEGARCH	0.0658 (0.0841)	0.0901 (0.1172)	0.0706 (0.0924)	0.0532 (0.0682)	0.0695 (0.0900)

analysis approach is higher than the loss estimated by VaR_p and ES_p (see Tables 5.5 and 5.6). For all values of p, the MaxLoss value is higher (in absolute value) than the observed loss.

Table 5.7: Portfolio Maximum Loss Values for Different Values of p and Their Respective Scenario.

		Scenario			
p	$MaxLoss$	$r_{1,ML}$	$r_{2,ML}$	$r_{3,ML}$	$r_{4,ML}$
0.50	-0.0453	-0.0451	-0.0460	-0.0453	-0.0449
0.55	-0.0475	-0.0473	-0.0482	-0.0475	-0.0471
0.65	-0.0521	-0.0519	-0.0529	-0.0521	-0.0517
0.75	-0.0574	-0.0572	-0.0583	-0.0574	-0.0569
0.85	-0.0642	-0.0640	-0.0652	-0.0642	-0.0637
0.95	-0.0762	-0.0759	-0.0773	-0.0762	-0.0756
0.99	-0.0901	-0.0897	-0.0915	-0.0901	-0.0894

6 Conclusion

Here we consider the same time series generated and analyzed in Lopes and Prass (2009) to estimate the risk measures VaR_p, ES_p and MaxLoss. For those time series we fit FIEGARCH models to estimate the conditional variances of the time series. We observe that the higher the sample size, the higher the mean square error value. We use the estimated variances to estimate the risk measure VaR_p for the simulated time series under different approaches. Since the estimated values for this risk measure, under different approaches, are very close from one another, we cannot say that one method is better than

the others. For this simulated study, the Econometric approach, considering FIEGARCH models, does not perform as well as one would expected. However, the results obtained by using these models have strong influence from the model parameter estimation, which is based on the quasi-likelihood method. Asymptotic properties of the quasi-likelihood estimator are still an open issue and it could explain the unexpected results.

Regarding the observed time series, we consider two different approaches for analyzing the portfolio risk. We consider the distribution function of the portfolio log-returns (univariate case) and the multivariate distribution function of the risk-factor changes (multivariate case). Also, we consider either, the conditional and the unconditional distribution functions in all cases.

In the VaR_p and ES_p calculation, all approaches present similar results. By comparing the observed loss, the values estimated using the econometric approach (and FIEGARCH models) were the closest to the observed values. In all cases, the estimated loss was higher than the observed one. We also observe that the values estimated by considering the univariate distribution of the portfolio log-returns were smaller than the values estimated by considering the multivariate distribution of the risk-factor changes. By comparing the estimated values of VaR_p, ES_p and MaxLoss we observe that the loss estimated under the scenario analysis approach is higher than the loss estimated by VaR_p and ES_p. For all values of p, the MaxLoss value is higher (in absolute value) than the observed loss.

Acknowledgments

T.S. Prass was supported by CNPq-Brazil. S.R.C. Lopes research was partially supported by CNPq-Brazil, by CAPES-Brazil, by *INTC in Mathematics* and also by Pronex *Probabilidade e Processos Estocásticos* - E-26/170.008/2008 -APQ1.

References

Artzner, P., F. Delbaen, J. Eber and D. Heath (1999). "Coherent Measures of Risk". *Math. Finance*, Vol. **9**, 203-228.

Baillie, R., T. Bollerslev e H. Mikkelsen (1996). "Fractionally Integrated Generalised Autoregressive Conditional Heteroscedasticity". *Journal of Econometrics*, Vol. **74**, 3-30.

Beran, J. (1994). *Statistics for Long-Memory Processes*. New York: Chapman & Hall.

Billingsley, P. (1995). *Probability and Measure*. New York: John Wiley.

Bollerslev, T. and H.O. Mikkelsen (1996). "Modeling and Pricing Long Memory in Stock Market Volatility". *Journal of Econometrics*, Vol. **73**, 151-184.

Brockwell, P.J. and R.A. Davis (1991). *Time Series: Theory and Methods*. New York: Springer-Verlag.

Embrechts, P., C. Klppelberg and T. Mikosh (1997). *Modeling Extremal Events*. New York: Springer-Verlag.

Engle, R.F. (1982). "Autoregressive Conditional Heteroskedasticity with Estimates of Variance of U.K. Inflation". *Econometrica*, Vol. **50**, 987-1008.

Khindanova, I. and Z. Atakhanova (2002). "Stable Modeling in Energy Risk Management". *Mathematical Methods of Operations Research*, Vol. **55**(2), 225-245.

Lopes, S.R.C. and T.S. Prass (2009). "Theoretical Results on FIEGARCH Processes". Submitted.

Lopes, S.R.C. (2008). "Long-range Dependence in Mean and Volatility: Models, Estimation and Forecasting", In *In and Out of Equilibrium 2 (Progress in Probability)*, V. Sidoravicius and M.E. Vares (eds.), Vol. **60**, pp. 497-525, Birkhäuser.

Lopes, S.R.C and B.V.M. Mendes (2006). "Bandwidth Selection in Classical and Robust Estimation of Long Memory". *International Journal of Statistics and Systems*, Vol. **1**(2), 177-200.

Markowitz, H. (1952). "Portfolio Selection". *Journal of Finance*, Vol. **7**(1), 77-91.

McNeil, A.J., R. Frey and P. Embrechts (2005). *Quantitative Risk Manangement*. New Jersey: Princeton University Press.

Nelson, D.B. (1991). "Conditional Heteroskedasticity in Asset Returns: A New Approach". *Econometrica*, Vol. **59**, 347-370.

Prass, T.S. and S.R.C. Lopes (2009). "Stress Tests, Maximum Loss and Value-at-Risk on FIEGARCH Process". Submitted.

Ruiz, E. and H. Veiga (2008). "Modelling long-memory volatilities with leverage effect: A-LMSV versus FIEGARCH". *Computational Statistics and Data Analysis*, Vol. **52**(6), 2846-2862.

Roberts, S.W. (1959). "Control Charts Based on Geometric Moving Averages". *Technometrics*, Vol. **1**, 239-250.

Zangari, P. (1996). "RiskMetrics Technical Document". The URL address is http://www.riskmetrics.com/rmcovv.html.