# VALUE AT RISK AND EXTREME VALUE THEORY: APPLICATION TO THE JOHANNESBURG SECURITIES EXCHANGE

**R. Williams[a], J.D. van Heerden[b]\* and W.J. Conradie[c]**

## Abstract

Value at Risk (VaR) has been established as one of the most important and commonly used financial risk management tools. Nevertheless, the attractive features and wide-spread use of VaR could not help to avoid a number of financial crises and its severe impact on economies globally, the latest being the 2008 financial crisis. In isolation, VaR has, in the past, mostly focused on events that occur with a 1% or 5% probability. This is a popular reason offered for its failure of 'predicting' the financial crises, as the latter are viewed as 'extreme' events and can therefore not be classified as events with a 1% or 5% probability of happening. The use of Extreme Value Theory (EVT) in calculating VaR is a relatively new approach and attempts to expand on the traditional VaR-only approach to include potential extreme events. This approach has provided good results in developed markets and in this article we investigate if the same holds true in the more volatile South African equity space. We examine and compare the application of seven VaR and VaR-EVT models on the FTSE/JSE Total Return All Share Index. Our results suggest that the Filtered Historical Simulation VaR method is the best all-round model. It is, however, worthwhile to employ EVT in the form of the conditional Generalized Pareto Distribution (GPD) model when calculating very extreme quantiles such as the 0.1% quantile. Our results further highlight the importance of filtering the data in order to account for the conditional heteroskedasticity of the financial time series.

---

[\*a]Postgraduate Student, Department of Statistics and Actuarial Science, Stellenbosch University
[b]Faculty of Agribusiness and Commerce, PO Box 85084, Lincoln University, Lincoln 7647, Christchurch, New Zealand
[c]Department of Statistics and Actuarial Science, Stellenbosch University
Email: jd.vanheerden@lincoln.ac.nz

# 1    Introduction

Value at Risk (VaR) is a common statistical risk measure that summarises the maximum potential loss over a specific time horizon at a given confidence level. It is a particularly pertinent risk measure in today's high-risk financial climate and has become increasingly popular due to its ability to state a financial risk situation in a single figure. However, a number of financial crises that had severe adverse effects on financial markets, the latest being the financial crisis of 2008, have brought the application of VaR into question. VaR has, in the past, mostly focused on events that occur with a 1% or 5% probability which could, in theory, explain why the use of VaR alone is insufficient to predict such crises as these are regarded as extreme events rather than events associated with a 5% or even 1% probability of happening. In keeping with the devastating impact these crises had on economies around the world, it can be argued that for risk management and regulatory purposes, it has become even more important to also accurately predict the probability of an extreme event. Within the context of financial equity markets, these extreme events are reflected in extreme returns. The latter are found in the tails of the underlying return distribution and to be able to accurately predict it, the tails need to be accurately modeled. Although most financial returns are found to be fat-tailed (Jansen & De Vries, 1991), common VaR measures rely on the simplifying assumption that returns have a specific parametric distribution. These distributions do not display the required fat-tails to accurately model the financial returns and subsequently tend to underestimate the VaR at extreme quantiles (Dicks, Conradie & De Wet, 2014).

The use of Extreme Value Theory (EVT) in calculating VaR is a relatively new addition to the tool kit of the financial risk manager. It produced good results in an Engineering sphere where it has been used to design flood walls and dykes (Danielsson, 2011), before it was applied to finance. EVT could also be appropriate for financial risk management because it fits extreme quantiles better than conventional approaches for heavy-tailed data (Gençay & Selçuk, 2004). It does not make a prior assumption about the underlying return distribution but instead focuses only on the modeling of the extreme returns found in the tail of the distribution. While more traditional VaR methods may be able to adequately estimate the 5% and 1% quantile, EVT may be better suited to the goal of estimating very small quantiles such as 0.5% and 0.1% (Diebold, Schuermann & Stroughair, 2000). Additionally, EVT is able to model the left and the right tails independently which is important because risk and reward are not equally likely, especially in emerging markets (Gençay & Selçuk, 2004).

The aim of this study is to determine whether the use of EVT in calculating VaR for the South African equity market provides similarly good results as those associated

with more developed markets. As part of this objective, we examine which models are best suited to calculating VaR for the FTSE/JSE Total Return All Share Index (ALSI) across a range of quantiles. To reach this objective, seven candidate VaR models found to have commonly been used in prior research are examined for the ALSI. The candidate VaR models examined in this study are grouped into parametric, non-parametric and semi-parametric categories. Specifically, models that fall into the parametric category include the location-scale Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, Historical Simulation (HS) is employed as a non-parametric method and Filtered Historical Simulation (FHS), the Generalised Pareto Distribution (GPD) EVT method and the conditional GPD EVT methods are included as semi-parametric VaR methods.

This study adds to the existing body of related literature in the following ways: a) It is the first study to test and compare so many distinct VaR models on the broad South African equity market, making it the most comprehensive study of its type to date; b) we apply the independence test when back testing the models (the importance of applying the independence test is discussed in Section 3) which is novel for the South African market; c) we use a window period of 1000 days (rather than 250 days used in prior studies) to bring the results more in line with international literature; d) the accuracy of the models are examined at more quantiles than prior studies, namely at the 0.95, 0.99, 0.995 and 0.999 quantiles.

The remainder of article is structured as follows: Section 2 provides a theoretical overview of EVT, VaR, the models employed and VaR back-testing methods. Relevant literature is discussed in Section 3, followed by a discussion on the data used and method applied in Section 4. The findings are discussed in Section 5 and the article is concluded in Section 6.

## 2 Theoretical background

From the literature (discussed in Section 3) we identified seven VaR models that we apply and compare. The models include three GARCH models with different innovations, the HS model, the unconditional GPD model, the FHS model and the conditional GPD model. In this section we provide a brief summary of the theory that underlies these models. Before the theoretical properties of each model is discussed, a brief description of EVT and VaR is in order. The interested reader is referred to Coles, Bawa, Trenner and Dorazio (2001) and Alexander (2009) for a more detailed theoretical discussion on EVT and VaR respectively.

## 2.1    Extreme value theory

Extreme value theory studies the statistical behaviour of the maximum, denoted $M_n$, of a dataset over $n$ time units of observation. Let $X_1, X_2, \ldots, X_n$ denote the return series. The variables $X_i$, $i = 1,2, \ldots, n$ have a common distribution function $F(x) = P(X_t \leq x)$ with mean $\mu$ and standard deviation $\sigma$. The distribution function is unknown and it is assumed that the variables are independent and identically distributed (i.i.d).

The maximum can be written,
$$M_n = \max(X_1, X_2, \ldots, X_n) \tag{1}$$

The Fisher-Tippett theorem states that if there exists a sequence of constants $\{a_n > 0\}$ and $\{b_n\}$ such that:

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G(z) \ as \ n \rightarrow \infty \tag{2}$$

where $G$ is a non-degenerate distribution function, then $G$ belongs to one of three families of distributions namely the Gumbel, Fréchet or Weibull. The Fisher-Tippet theorem suggests that, regardless of the original distribution of the observed data, the asymptotic distribution of the maxima belongs to one of the above three distributions. By taking the reparameterisation $\gamma = 1/\alpha$, due to Von Mises (1936) and Jenkinson (1955), the Gumbel, Fréchet and Weibull distribution can be written as a single model with just one parameter:

$$G_\gamma(x) = \begin{cases} \exp\{-(1 + \gamma x)^{-\frac{1}{\gamma}}\}, & if \ \gamma \neq 0, 1 + \gamma x > 0 \\ \exp\{-\exp(-x)\}, & if \ \gamma = 0 \end{cases} \tag{3}$$

This representation is known as the Generalised Extreme Value Distribution (GEV), the parameter $\gamma$ is called the extreme value index (EVI) and $\alpha$ is the tail index.

According to Marimoutou, Raggad and Trabelsi (2009) an efficient approach to modelling extreme events in practice is to attempt to focus not only the maximum events, but on all events greater than some large pre-set threshold. An exceedance of a threshold $u$ occurs when $X_t > u$ for any $t$ in $t = 1, 2, \ldots, n$. An excess over $u$ is defined by $y = X_i - u$.

The conditional distribution of X, given that X exceeds some threshold $u$ is given by:

$$F_u(x) = P(X - u \leq y | X > u) \tag{4}$$

This represents the probability that X exceeds the threshold $u$ by at most an amount $y$, given that X exceeds the threshold $u$. This can also be written as:

$$F_u(x) = \frac{F(u+y) - F(u)}{1 - F(u)}, \quad y \geq 0 \tag{5}$$

Since $x = y + u$ for $X > u$, $F(x)$ can be written as

$$F(x) = \big(1 - F(u)\big)F_u(y) + F(u) \tag{6}$$

A theorem by Balkema and De Haan (1974) and Pickands (1975) states that for large enough $u$, the distribution of $X - u$, given that $X > u$, may be approximated by the GPD, which is defined as:

$$G_{\gamma,\sigma,v}(x) = \begin{cases} 1 - (1 + \gamma\frac{x-v}{\beta})^{-\frac{1}{\gamma}}, & if\ \gamma \neq 0 \\ 1 - e^{-\left(\frac{x-v}{\beta}\right)}, & if\ \gamma = 0 \end{cases} \tag{7}$$

where $\quad x \in \begin{cases} [v, \infty], & if\ \gamma \geq 0 \\ \left[v, v - \frac{\beta}{\gamma}\right], & if\ \gamma < 0 \end{cases}$

$\gamma = 1/\alpha$ is the shape parameter
$\alpha$ is the tail index
$\beta$ is the scale parameter
$v$ is the location parameter

when $v = 0$ and $\beta = 1$ then the representation is known as the standard GPD.

One can either specify the number of upper order statistics in the tail used to model the GPD, or $u$, the threshold above which to model the GPD. In this study we follow the latter approach. There are a number of ways in which one can select a threshold, but there is no widely accepted method for determining $u$. Graphical methods, the approach we follow, include inspecting the mean excess function of the GPD. By detecting an area on the graph with a linear shape it is possible to choose an appropriate threshold. The choice of threshold is of importance when calculating the tail estimator (and consequently the VaR as will be seen in Section 2.5.2).

Following from equation (6), since $F_u(y)$ converges to the GPD for sufficiently large $u$ and since $x = y + u$ for $X > u$, we have

$$F(x) \approx \left(1 - F(u)\right) G_{\gamma,\beta,v}(x - u) + F(u) \tag{8}$$

after determining a high threshold $u$, $F(u)$ can be estimated by $\frac{N-N_u}{N}$ where $N_u$ is the number of exceedences and $N$ is the sample size.

Subsequently it can be shown that the tail estimator becomes:

$$\hat{F}(x) = 1 - \frac{N_u}{N}\left(1 + \hat{\gamma}\frac{x-u}{\hat{\beta}}\right)^{-\frac{1}{\hat{\gamma}}} \tag{9}$$

given that

$$G_{\gamma,\beta,u}(x) = 1 - \left(1 + \gamma\frac{x-u}{\beta}\right)^{-\frac{1}{\gamma}} \tag{10}$$

where $\hat{\gamma}$ and $\hat{\beta}$ are the maximum likelihood estimators of $\gamma$ and $\beta$ respectively, and $u$ is the threshold.

## 2.2 Value at risk

Let $r_t = \log\left(\frac{p_t}{p_{t-1}}\right)$ be the return at time t where $p_t$ is the price of an asset at time t and let $r_1, r_2, \ldots, r_n$ be independent and identically distributed (i.i.d.) random variables.

Adapted from Abad, Benito and López (2014), let $F(r)$ denote the cumulative distribution function $F(r) = P(r < r|\Omega_{t-1})$ conditionally on the information set $\Omega_{t-1}$ that is available at time $t$-1.

Assume that $\{r_t\}$ follows the stochastic process:

$$r_t = \mu + \varepsilon_t \tag{11}$$

$$\varepsilon_t = z_t\sigma_t \text{ and } z_t \sim \text{i.i.d}(0,1) \tag{12}$$

where $\sigma_t^2 = E(z_t^2|\Omega_{t-1})$ and $z_t$ has the conditional distribution function $G(z)$ where

$$G(z) = P(z_t < z|\Omega_{t-1}) \tag{13}$$

The VaR with a given probability $\alpha \in (0,1)$, is defined as the $\alpha$ quantile of the probability distribution of financial returns (for ease of exposition the conditionality is not shown explicitly):

$$F\big(VaR(\alpha)\big) = P\big(r_t < VaR(\alpha)\big) = \alpha \tag{14}$$

or

$$VaR(\alpha) = \inf\{v | P(r_t \le v) = \alpha\} \tag{15}$$

This VaR quantile can be estimated in one of two ways; either inverting the distribution function of the financial returns $F(r)$ or inverting the distribution function of the innovations $G(z)$. In the latter case it is also necessary to estimate $\sigma_t^2$. Hence VaR can also be written as

$$VaR(\alpha) = F^{-1}(\alpha) = \mu + \sigma_t G^{-1}(\alpha) \tag{16}$$

There are three types of VaR methodology, via which $F(r)$ or $G(z)$ can be estimated, namely parametric methods, non-parametric methods and semi-parametric methods. In this study we examine three parametric, one non-parametric and three semi-parametric methods.

## 2.3    Parametric methods

GARCH models explicitly model the conditional volatility as a function of past conditional volatilities and returns. We assume that returns belong to a location-scale family of probability distributions of the form:

$$r_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t z_t \tag{17}$$

where    $\mu_t$ is the location parameter and
$\sigma_t$ is the scale parameter

$\mu_t$ and $\sigma_t$ are determined by the data available at time *t-1*.

$z_t \sim iid\ f_Z(.)$ where $f_Z$ is a zero-location, unit-scale probability density function that can have additional shape parameters. The original GARCH models took $z_t$ to be Gaussian, although this assumption is often not appropriate for financial returns data. A fat-tailed and possibly asymmetric distribution could be found to be a better alternative.

The VaR forecast based on information up to time t can be written as:

$$\widehat{VaR}_{t+1,\alpha} = -(\hat{\mu}_{t+1} + \hat{\sigma}_{t+1}Q_\alpha(z)) \tag{18}$$

where $Q_\alpha(z)$ is the $\alpha$-quantile implied by $f_Z$.

When estimating VaR with a GARCH type model the innovation distribution can follow various distributions, such as a normal distribution, Student's $t$ distribution and skew Student's $t$ distribution. In this article we examine these three innovation distributions.

## 2.4    Non-parametric methods

HS is an example of a non-parametric VaR method. In HS the empirical quantile estimator is estimated from a sample of historical data. In other words, the empirical distribution of the financial returns is used as an approximation for $F(r)$.

Mathematically, HS VaR can be defined as:

$$VaR_{t+1,\alpha} = Quantile\{\{r_t\}_{t=1}^n\} \tag{19}$$

where $r_t$ is the return on day $t$

## 2.5    Semi-parametric methods

Semi-parametric VaR methods combine both the parametric and the non-parametric approach. They are designed to be able to take into account the time varying structures evident in financial time series by means of a parametric GARCH-type model, without placing the restriction of an assumption like that of normality when estimating the residual distribution.

In contrast to non-parametric methods, semi-parametric methods deal with i.i.d data, instead of relying on resampling procedures for non-i.i.d. data and the associated assumptions for those to hold (Mancini & Trojani, 2005)

### 2.5.1    Filtered historical simulation

FHS combines a GARCH model with HS. This model can accommodate volatility clustering and the skewness inherent in the empirical distribution (Ghorbel & Trabelsi, 2009). First, a GARCH model is fitted to the return data and the

standardized residuals are extracted. If the model fits well then the standardized residuals should be i.i.d and HS can be applied to determine the VaR:

$$VaR_{t+1,\alpha} = \mu_{t+1} + \sigma_{t+1} Quantile\{\{z_t\}_{t=1}^n\} \tag{20}$$

where $Quantile\{\{z_t\}_{t=1}^n\}$ is the left quantile at $\alpha\%$ of the standardized residuals.

### 2.5.2   Unconditional GPD

Following our discussion in Sections 2.1 and 2.2, an EVT estimate of VaR can be obtained by applying the following steps (Rocco, 2014):

1) Assume that the data are in the maximum domain of attraction of a GEV distribution.
2) Fix a high threshold $u$ and fit the GPD to the exceedances over $u$.
3) Obtain estimates for $\hat{\gamma}$ and $\hat{\beta}$.
4) Estimate the tail probability using equation (9).
5) Invert the formula to obtain an estimate of the $\alpha$ quantile:

$$VaR_\alpha(X) = u + \frac{\hat{\beta}}{\hat{\gamma}}\left[\left(\frac{N}{N_u}(1-\alpha)\right)^{-\hat{\gamma}} - 1\right] \tag{21}$$

where all variables are as previously defined.

### 2.5.3   Conditional GPD

The unconditional GPD method assumes i.i.d data. To deal with the non-i.i.d. nature of financial returns a two-step procedure is used that first models the correlation structure of the observations and then performs the estimation of the GPD distribution on the resulting residuals which can be considered to be roughly i.i.d.

This two-step procedure was first suggested by Diebold *et al.* (2000) and implemented by McNeil and Frey (2000). It can be summarized as follows (McNeil & Frey, 2000):

1) Fit a GARCH-type model to the return data making no assumptions about $F(z)$ and using pseudo-maximum-likelihood estimation (PML). Calculate the estimates of the conditional mean and variance for day *t+1* and extract the residuals.

2) Consider the residuals to be a realisation of a strict white noise process and use EVT to model the tail of $F_Z(x)$. Use this EVT model to estimate the quantile of interest. The VaR is then calculated as:

$$VaR_{\alpha,t+1} = \mu_{t+1} + \sigma_{t+1}VaR_t(Z) \tag{22}$$

where all variables are as previously defined.

## 2.6 Back-testing

Back-testing is an important tool that can be used to check the adequacy of a particular VaR model and to compare various VaR models. It takes ex ante VaR forecasts from a particular model and compares them with the ex post realized return. When the realized loss exceeds the VaR, a violation is said to have occurred (Danielsson, 2011). An accurate VaR model should correctly measure the frequency of VaR exceedances as well as determine whether exceedances occur independently of each other (Campbell, 2006). Kupiec's unconditional coverage test (1995) checks that the exceedance rate is in line with the expected number of violations, while the independence test checks that violations occur independently of each other.

### 2.6.1 Violation ratio

The violation ratio is defined as the total number of violations divided by the total number of one-day VaR forecasts (Danielsson, 2011):

$$VR = \frac{observed\ number\ of\ violations}{expected\ number\ of\ violations} = \frac{E}{\alpha \times N} \tag{23}$$

where   $E$ is the number of exceedances
$\alpha$ is the confidence level at which the VaR was calculated
$N$ is the number VaR forecasts made

A violation ratio of 1 is expected. A violation ratio greater than one means that the VaR model has under forecasted the risk and if it is smaller than one then the model has over forecasted the risk (Danielsson, 2011). The Kupiec Test (1995) can be used to determine whether any value other than one is statistically significant.

### 2.6.2 Kupiec's unconditional coverage test

Kupiec (1995) proposed a proportion of failures (POF) test that examines how many times a financial institution's VaR is violated over a given time frame. If the number of violations is significantly different from the expected number of failures, $\alpha \times$

100% of the sample, then the accuracy of the underlying VaR model is called into question (Campbell, 2006).

Let $N = \sum_{t=1}^{T} I_{t+1}$ be the number of days over a $T$ period that the portfolio loss was larger than the VaR estimate, where $I_{t+1}$ is a sequence of violations that can be defined as:

For the left tail:

$$I_{t+1} = \begin{cases} 1, if \ X_{t+1} < VaR_{t+1} \\ 0, if \ X_{t+1} \geq VaR_{t+1} \end{cases} \tag{24}$$

For the right tail:

$$I_{t+1} = \begin{cases} 1, if \ X_{t+1} > VaR_{t+1} \\ 0, if \ X_{t+1} \leq VaR_{t+1} \end{cases} \tag{25}$$

Let $p$ be the expected failure rate. If the total number of trials is $T$, then the number of failures $F$ can be modelled with a binomial distribution with probability of occurrence $\alpha$.

The null and alternate hypothesis can be written as:

$$H_0: \frac{F}{T} = \alpha$$
$$H_1: \frac{F}{T} \neq \alpha$$

We want to determine whether the observed failure rate is significantly different from the expected failure rate. Kupiec's POF test (1995) is conducted via a likelihood-ratio (LR) test.

The likelihood ratio statistic is:

$$LR_{UC} = 2 \left[ \log((\tfrac{F}{T})^F \left(1 - \tfrac{F}{T}\right)^{T-F}) - \log(\alpha^F (1 - \alpha)^{T-F}) \right] \tag{26}$$

Under $H_0$, $LR_{UC} \rightarrow \chi^2(1)$ i.e. the likelihood ratio statistic is asymptotically chi-squared distributed with one degree of freedom. If $LR_{UC}$ exceeds the critical value of the $\chi^2(1)$ distribution then the null hypothesis will be rejected and the model is said to not accurately model the number of VaR exceedances (Nieppola, 2009).

### 2.6.3  Independence testing

Define an indicator variable *I* that is assigned a value of 1 if the VaR is exceeded and a value of 0 if it is not exceeded. Next define $n_{ij}$ as the number of days when condition $j$ occurred assuming that condition $i$ occurred on the previous day. Therefore $n_{10}$ means that a day with no VaR violation followed a day that experienced a VaR violation.

The possible outcomes can then be displayed as follows:

|            | $I_{t-1} = 0$     | $I_{t-1} = 1$     |                   |
|------------|-------------------|-------------------|-------------------|
| $I_t = 0$  | $n_{00}$          | $n_{10}$          | $n_{00} + n_{10}$ |
| $I_t = 1$  | $n_{01}$          | $n_{11}$          | $n_{01} + n_{11}$ |
|            | $n_{00} + n_{01}$ | $n_{10} + n_{11}$ | $N$               |

Let $\pi_i$ be the sample probability of observing an exceedence conditional on state $i$ on the previous day:

$$(27)$$

$$\pi_0 = \frac{n_{01}}{n_{00} + n_{01}}$$

$$(28)$$

$\pi$ represents the violation rate:

$$\pi = \frac{n_{01} + n_{11}}{n_{00} + n_{01} + n_{10} + n_{11}} \tag{29}$$

under the null hypothesis: $H_0: \pi_1 = \pi_0$

In other words a VaR exceedance does not depend on whether or not an exceedance occurred the previous day (Nieppola, 2009). The test statistic of independences of exceptions is a likelihood ratio:

$$LR_{ind} = 2\ln\left(\frac{(1-\pi)^{n_{00}+n_{10}}\pi^{n_{01}+n_{11}}}{(1-\pi_0)^{n_{00}}\pi_0^{n_{01}}(1-\pi_1)^{n_{10}}\pi_1^{n_{11}}}\right) \tag{30}$$

under $H_0$, $LR_{ind} \rightarrow \chi^2(1)$

If the test statistic is above the critical value then the null hypothesis is rejected and the model does not generate VaR exceedances that are independent. This test does not depend on the true value of the expected failure rate. It only tests for

independence of violations and thus should be assessed in conjunction with Kupiec's unconditional coverage test.

# 3      Literature review

We start off our discussion on related research by focusing on literature associated with developed markets, followed by emerging markets and in the last section we turn our attention to South African related research.

## 3.1      Developed markets

EVT was introduced to a financial setting by Koedijk, Schafgans and De Vries (1990) and Jansen and De Vries (1991). Since then there has been much work done combining EVT and VaR to better model extreme quantiles that are of interest to financial risk managers.

Danielsson and De Vries (2000) compare the J.P Morgan RiskMetrics VaR technique with HS and their own semi-parametric method. This method uses the empirical distribution for smaller risks and extreme value theory for the largest risks. A window period of 1500 days of return data is used and they find that at low probability RiskMetrics under predicts the VaR while HS over predicts the VaR. They conclude that their semi-parametric method is more accurate than the other two methods.

McNeil and Frey (2000) combine the fitting of a GARCH-type model, to estimate the current volatility, with EVT, to estimate the tail of the innovation distribution of the GARCH model. They develop a two-step method (discussed in Section 2.5.3) for calculating a conditional EVT-VaR measure which they test on the Standard and Poor's and DAX index. An AR(1)-GARCH(1,1) model with normal innovations is used to model the volatility and then a GPD is fitted to the tails of the extracted standardized residuals. A moving window period of 1000 days is used and they test at the 0.95, 0.99 and 0.995 quantiles. A simulation study is conducted to determine the threshold choice for use in their two-step method. It is determined that the choice of a threshold equal to100 (or 10% of the window size) is optimal. They find that their procedure gives better results than those methods which ignore the heavy tails of the innovations or the stochastic nature of the volatility.

Gençay, Selçuk and Ulugülyağci (2003) compare the Variance-Covariance, HS, GARCH(1,1) with both normal and Student t innovations, adaptive and nonadaptive GPD models. Three different rolling window sizes of 500, 1000 and 2000 are used to calculate the high quantiles for both the Istanbul Stock Exchange Index (ISE-100) and the S&P-500. They find that the quantile forecasts of the GARCH models are

very volatile in comparison to the GPD quantile forecast. The GPD model is found to be a robust quantile forecasting tool which is practical to implement and regulate for VaR purposes.

Marimoutou *et al.* (2009) apply EVT to the oil market. They compare an unconditional Normal VaR model, HS, FHS, an AR(1)-GARCH(1,1) model with both normal and Student *t* innovations, a GPD and a conditional GPD model. Both the FHS and conditional GPD are filtered using an AR(1)-GARCH(1,1) model with normal innovations. The VaR methods are all calculated using a rolling window of 1000 days and at quantiles of 0.95, 0.99, 0.995 and 0.999. A sensitivity analysis is done for the conditional GPD approach to determine the optimal threshold value which is set at 10% of the window size. This is in concurrence with McNeil and Frey's (2000) suggestion. They conclude that the conditional GPD and the FHS VaR methods provide improved results over the more conventional methods and that the filtering process is important for the success of these two methods.

## 3.2 Emerging markets

Gençay and Selçuk (2004) compare the Variance-Covariance method with the normal and Student-t distribution, HS and the unconditional GPD VaR method. They test the models on the daily stock market returns of nine different emerging markets, namely Argentina, Brazil, Hong Kong, Indonesia, Korea, Mexico, Singapore, Taiwan and Turkey. Sliding windows of three different sizes are used, specifically 500, 1000 and 1500 days, except for the GPD method where they use all the data up to the point of the VaR estimation. The upper 2.5% of the data points were used for the GPD approach. It was found that risk and reward are not equally likely in the developing markets which they modeled. They conclude that the GPD VaR estimate was the most accurate at higher quantiles.

Pattarathammas, Mokkhavesa, and Nilla-Or (2008) study VaR methods using EVT on ten Asian equity markets. They use Normal VaR, HS and the GPD VaR method. They also filter each method using an exponentially weighted moving average (EWMA), as used in RiskMetrics, and a GARCH(1,1) model resulting in nine different VaR methods. The conditional approaches are based on the two-step method of McNeil and Frey (2000) using a threshold of 100. A rolling window of 1000 data points is used and they test at the 0.95 and 0.99 quantile. They find that unconditional GPD and simple HS perform less accurately when calculating the VaR estimate, especially at higher confidence levels, when compared to FHS. The conditional GPD does not perform much differently from FHS and there is not much difference between the use of the EWMA and the GARCH-based filter. GARCH

models may reflect more flexible volatility adjustment than EWMA, but the models perform quite similarly.

Angelidis and Benos (2008) evaluate many different VaR methods for Greek stocks. These included the Variance-Covariance method, RiskMetrics with GARCH, EGARCH and TARCH volatility modeling under the normal, Student-t and Skewed Student-t distributions as well as the HS, FHS and GPD methods. They find that FHS performs the best at the 99% confidence level and that the GPD method also performs acceptably well. At the lower confidence level of 97.5% most of the models that they tested gave similar, good results.

### 3.3    South Africa

Seymour and Polakow (2003) use the methods proposed by Danielsson and De Vries (2000) and McNeil and Frey (2000) as well as HS to calculate the VaR at high confidence levels on a portfolio of South African stocks. A threshold of 10% of the window size is used in accordance with McNeil and Frey (2000) and they find that McNeil and Frey's conditional GPD method works the best for the South African market, but that none of the methods worked nearly as well as when tested in developed markets.

McMillan and Thupayagale (2010) compare the RiskMetrics model with GARCH models that include asymmetric and long memory models when calculating VaR for the JSE All Share Index. They find that GARCH models consistently outperform the RiskMetrics model and conclude that the latter may not be of great relevance in the South African equity market. GARCH models that incorporate long memory components or asymmetric effects, or both, are found to perform best.

Dicks, Conradie and de Wet (2014) use McNeil and Frey's (2000) two-step process combining both symmetric and non-symmetric GARCH models with EVT to the JSE Financial Index. They use a window period of 250 days which results in their GARCH models not converging and they propose a method to overcome this. They calculate VaR at a 99% confidence level, taking a threshold equal to 20% of the window size and also look at various VaR scaling methods. They conclude that none of their models is universally optimal.

## 4    Data and Methodology

Daily data for the FTSE/JSE Total Return All Share Index was obtained from I-NET Bridge and covers the period from 30 June 1995 to 17 November 2014, resulting in a total of 4 844 observations in the dataset. The daily log-returns of the index are

presented in Figure 1. All data modeling was performed using the statistical programming language R.

The summary statistics and related statistical tests[1] confirm that the data follows a non-normal distribution, conditional heteroskedasticity is present and the time series is stationary. We note that the data is not i.i.d. which is a necessary condition for the application of EVT. Hence, it is necessary to first filter the returns with a GARCH model in order to get approximately i.i.d. data to which EVT can then be applied.

## 4.1    HS

We apply equation (19) using a rolling sample of 1000 observations in order to calculate the one-day ahead VaR forecast for $\alpha \in \{0.95, 0.99, 0.995, 0.999\}$.

## 4.2    GARCH approach

Similar to McNeil and Frey (2000) we use the GARCH(1,1) process for the volatility and an AR(1) model for the dynamics of the conditional mean. We investigate an AR(1)-GARCH(1,1) model with normal, Student's t and skewed Student-t distributed innovations. All the parameter estimates are significant indicating that our models fit the data well[2]. The three different AR(1)-GARCH(1,1) models are used directly to calculate VaR, as described in section 2.3. The AR(1)-GARCH(1,1) specification is estimated using a rolling window of 1000 days. For each rolling window one one-day-ahead VaR forecast is calculated.

## 4.3    GPD modeling

Following the approach discussed in Section 2.5 we fit the GPD to the right hand tail of the first 1000 data points. EVT is designed to work with maximums so when modeling the left tail of the distribution the returns are multiplied by -1. The parameters are extracted from the modeling and the predicted VaR is calculated for the 1001$^{st}$ day using the calculation method as described in Section 2.5. The window is then moved forward by one day and the procedure is repeated until the last day, resulting in a total of 3844 VaR forecasts.

As discussed in Section 2.1, the choice of threshold is critical in the fitting of a GPD to data. Following a similar approach to McNeil and Frey (2000) and Marimoutou *et*

---

[1] The table of summary statistics is available from the authors on request.
[2] Tables with results of fitting the GARCH models are available from the authors on request.

*al*. (2009) we conclude that a threshold of 100 is suitable to be used on each rolling window to calculate the relevant VaR values[3].

## 4.4     FHS and conditional GPD

After examining the parameter estimation results and the graphs of the standardized residuals[4], it is seen that there is very little difference to using an AR(1)-GARCH(1,1) model with normal innovations compared to one with Student-t or skew Student-t innovations. As such, we continue forward using only the AR(1)-GARCH(1,1) with normal innovations to filter the return data for the purposes of applying the FHS and conditional GPD models.

The AR(1)-GARCH(1,1) specification is estimated on the entire data set and the standardised residuals are extracted from the estimated model. The standardized residuals are used to investigate the adequacy of the fitted model and also to use to filter the data for the use in the FHS and the conditional GPD models. The residual series is found to have significant excess kurtosis and skewness, is independently distributed and there are no signs of heteroskedasticity in the residuals. This means that the series has been filtered satisfactorily and we are now dealing with i.i.d. data which can be used in the FHS and conditional GPD risk measurement methods.

## 4.5     Evaluation of VaR models

We use a combination of the violation ratio, Kupiec's unconditional coverage test and the independence test as discussed in Section 2.6 to compare and evaluate the different VaR models.

## 5     Results

In Table 1 the violation ratios for the left and right hand tails respectively are reported. In Table 2 the *p*-values for the unconditional coverage test are reported and in Table 3 the *p*-values of the independence test are reported. Note that the following abbreviations are used to refer to the different models: The GARCH VaR model with normal innovations (GARCH~n); the GARCH VaR model with Student t distributed innovations (GARCH~t); the GARCH VaR model with skew Student t distributed innovations (GARCH~st); Historical Simulation VaR (HS), GARCH filtered Historical Simulation VaR (FHS~n); unconditional Generalised Pareto Distribution VaR (GPD) and the conditional GPD EVT VaR (GPD~n).

---

[3] A detailed discussion on the approach followed to determine the threshold level used in our study is available on request.
[4] Results and graphs are available from the authors on request.

**Table 1: Violation ratios**

This table reports the violation ratios of the return distribution of the ALSI as calculated by the different VaR models. The expected value of the violation ratio is the corresponding tail size i.e. the expected VaR violation ratio for the 5% quantile is 5% (Marimoutou *et al.,* 2009). A violation ratio greater than the expected value at that confidence level indicates that the model has under forecasted the risk and if it is less than the expected value then the model has over forecasted the risk. The ranking of the model for each quantile, $\alpha \in \{0.95, 0.99, 0.995, 0.999\}$, is shown in parenthesis.

| VaR model | *Left tail violation ratios* | | | | *Right tail violation ratios* | | | |
| | $\alpha$ | | | | $\alpha$ | | | |
| | 5% | 1% | 0.5% | 0.1% | 5% | 1% | 0.5% | 0.1% |
|---|---|---|---|---|---|---|---|---|
| **GARCH~n** | 5,489 (5) | 1,639 (7) | 1,119 (7) | 0,494 (7) | 3,460 (7) | 0,780 (6) | 0,442 (3) | 0,156 (5) |
| **GARCH~t** | 5,775 (6) | 1,138 (5) | 0,806 (6) | 0,182 (3) | 3,720 (6) | 0,546 (7) | 0,234 (7) | 0,052 (3) |
| **GARCH~st** | 5,281 (2) | 1,093 (4) | 0,572 (3) | 0,104 (1) | 4,604 (4) | 0,911 (3) | 0,416 (4) | 0,052 (3) |
| **HS** | 4,683 (3) | 0,989 (1) | 0,572 (3) | 0,260 (6) | 4,657 (3) | 0,911 (3) | 0,598 (5) | 0,260 (7) |
| **FHS~n** | 5,151 (1) | 0,911 (3) | 0,468 (2) | 0,182 (3) | 4,917 (1) | 1,119 (5) | 0,546 (2) | 0,156 (5) |
| **GPD** | 4,630 (4) | 0,937 (2) | 0,520 (1) | 0,182 (3) | 4,527 (5) | 0,989 (1) | 0,624 (6) | 0,078 (1) |
| **GPD~n** | 3,824 (7) | 0,702 (6) | 0,390 (5) | 0,052 (2) | 4,865 (2) | 1,067 (2) | 0,468 (1) | 0,078 (1) |

**Table 2: p-Values of unconditional coverage test**

This table reports the p-values of the unconditional coverage test. Under $H_0$ the exceedances are correct. A p-value greater than 5% indicates that the number of exceedances is correct.

| | $\alpha$ | | | | | | | |
| | 5% | | 1% | | 0.5% | | 0.1% | |
| VaR model | Left | Right | Left | Right | Left | Right | Left | Right |
|---|---|---|---|---|---|---|---|---|
| **GARCH~n** | 0,170 | 0,000 | 0,000 | 0,155 | 0,000 | 0,605 | 0,000 | 0,310 |
| **GARCH~t** | 0,031 | 0,000 | 0,026 | 0,002 | 0,013 | 0,009 | 0,149 | 0,230 |
| **GARCH~st** | 0,428 | 0,255 | 0,570 | 0,571 | 0,534 | 0,448 | 0,937 | 0,230 |
| **HS** | 0,362 | 0,323 | 0,944 | 0,571 | 0,534 | 0,402 | 0,009 | 0,009 |
| **FHS~n** | 0,669 | 0,812 | 0,571 | 0,468 | 0,778 | 0,688 | 0,149 | 0,310 |
| **GPD** | 0,288 | 0,171 | 0,689 | 0,943 | 0,859 | 0,293 | 0,149 | 0,654 |
| **GPD~n** | 0,000 | 0,699 | 0,050 | 0,681 | 0,315 | 0,778 | 0,300 | 0,654 |

**Table 3: p-Values of independence test**

This table reports the p-values of the independence test. Under $H_0$ the exceedances are not dependent on whether or not an exceedance was recorded the day before. A p-value greater than 5% indicates that the exceedances are independent.

| $\alpha$ | 5% | | 1% | | 0.5% | | 0.1% | |
|---|---|---|---|---|---|---|---|---|
| | Left | Right | Left | Right | Left | Right | Left | Right |
| **GARCH~n** | - | 0,1596 | 0,1473 | 0,4920 | 0,3712 | 0,6473 | 0,6639 | 0,8730 |
| **GARCH~t** | - | 0,0693 | 0,2509 | 0,5828 | 0,4777 | 0,8371 | 0,9091 | 0,9636 |
| **GARCH~st** | - | 0,1196 | 0,3239 | 0,4093 | 0,5828 | 0,6975 | 0,9091 | 0,9636 |
| **HS** | 0,0000 | 0,0000 | 0,3836 | 0,0416 | 0,6147 | 0,0069 | 0,8193 | 0,8193 |
| **FHS~n** | - | - | 0,4225 | 0,5080 | 0,6806 | 0,6310 | 0,8730 | 0,8910 |
| **GPD** | 0,0000 | 0,0000 | 0,4093 | 0,0584 | 0,6473 | 0,0083 | 0,8730 | 0,9454 |
| **GPS~n** | 0,0140 | - | 0,5365 | 0,4606 | 0,7317 | 0,6806 | 0,9636 | 0,9454 |

The results reported in Table 1 through Table 3 are discussed for each respective model below.

## 5.1 GARCH~n

The AR(1)-GARCH(1,1) VaR model with normally distributed innovations does not perform well for neither the left hand nor right hand tail at any of the calculated quantiles. In seven of the eight confidence levels that we tested it ranked $5^{th}$ or worse. The poor performance is confirmed by the unconditional coverage test where we reject the null hypothesis of the correct number of exceedances in four of the eight cases at the 5% testing level. However, it is found that the exceedances are independent of each other.

The model takes the heteroskedastic nature of the volatility of the returns into account, but due to the assumption of normality for the innovations it is the model which performs the worst. For the left hand tail, the GARCH~n model underestimates risk, while for the right hand tail the model overestimates risk. This is intuitive due to the negative skewness of the returns as well as the excess kurtosis of the returns over that of a normal distribution. Since a risk manager is primarily interested in the left hand tail's extreme returns the use of the normal assumption with this particular GARCH model is not recommended.

Figure 2 displays the back-testing results of the model. The 5% and 0.1% forecasted VaR is plotted against the observed returns for the period between 1999 and 2014.

## 5.2 GARCH~t

The AR(1)-GARCH(1,1) VaR model with Student-t distributed innovations performs ever so slightly better than the same GARCH model with normal innovations. The left tail is still underestimated while the right tail is overestimated. The reason for this is that the skewness of the return distribution is not taken into account.

In six of the eight cases the GARCH~t model performs 5th or worse and the unconditional coverage test is rejected in six of the eight cases. The model performs marginally better at the 0.1% quantile. The hypothesis of independent exceedances is accepted at all confidence levels. Figure 3 shows the back-testing results of the model.

## 5.3 GARCH~st

The AR(1)-GARCH(1,1) VaR model with skew Student-t distributed innovations performs the best of the three GARCH VaR models, attaining rankings between 1 and 4 for all quantiles. Specifically, the model produces good results for the left hand tail indicating that the skew Student's t distribution for the innovations is able to take the skewness of the return distribution into account. The unconditional coverage and independence exceedance hypothesis is accepted at every confidence level. Figure 4 shows the back-testing results of the model.

## 5.4 HS

The HS VaR model performs relatively well for the 5% and 1% confidence levels, but the performance of the model decreases as the confidence level increases. This is due to there being very few observations in the tails at the 0.5% and 0.1%. The poor performance of the HS VaR model is confirmed with the unconditional coverage test rejecting the hypothesis of correct exceedances at the 0.1% level for both the left and the right tail. The model fails the independence test for the right tail, except at the 0.1% confidence level. This is because the heteroskedastic nature of the volatility is not taken into account.

In Figure 5 one can see how extreme negative and positive returns affect the predicted VaR. Extreme returns will increase the VaR and will affect the VaR until they fall out of the 1000 day rolling window period used to calculate the VaR. This is particularly noticeable at the 0.1% level.

## 5.5    FHS~n

The FHS VaR model performs well at all confidence levels, achieving high rankings of 1, 2 and 3 for the left tail. It performs slightly worse for the right tail. The model satisfies the hypothesis of correct exceedances and independence of exceedances at all confidence levels. The FHS model offers an improvement on the HS method by taking into account the heteroskedastic nature of the volatility of the returns, by filtering the data with an AR(1)-GARCH(1,1) model with normally distributed innovations in order to produce i.i.d data. Figure 6 shows the back-testing results of the model.

## 5.6    GPD

The EVT-VaR method using the GPD only takes the tails of the return distribution into account as detailed in Section 2.1. Since the return data is not i.i.d it is not expected that the unconditional GPD method will perform very well. The method ranked in the top 4 for 6 of the 8 confidence levels. The hypothesis of correct exceedances is supported at every confidence level although the hypothesis of independence of exceedances is rejected at the 5% level as well as for the right hand tail at the 1% and 0.5% level. As can be seen in Figure 7, the VaR estimates are not quick to adjust following large positive or negative returns and as with HS, VaR estimates only return to normal levels once the extreme values have fallen out of the rolling window 1000 days later.

## 5.7    GPD~n

The conditional GPD VaR model does not perform well for the 5% and 1% quantiles, but the model's results improve as the quantile size decreases. The model also appears to model the right hand tail better than it does the left. For the right tail the model ranks $1^{st}$ or $2^{nd}$ for all 4 confidence levels and for the left tail it achieves a rank of $2^{nd}$ at the 0.1% confidence level. The model fails the unconditional coverage test for the left hand tail at the 5% and 1% confidence level and it fails the independence test at the 5% level. This model therefore appears to perform well only at the 0.1% quantile.

Benefits of using this model are that the heteroskedastic nature of the volatility is taken into account as well as the fact that the observations of importance, i.e. the extreme returns above a certain high threshold, are taken into account in the modelling. Figure 8 shows the back-testing results of the model.

# 6    Conclusion

In this study we examine seven different VaR models in order to determine which model is best to use when calculating VaR for the South African equity market. The models are applied to the ALSI for the period 30 June 1995 to 17 November 2014. VaR forecasts are made based on each model and these are back-tested against the observed returns. The violation ratio, unconditional coverage test and independence test are used to rank the models and analyse the results statistically.

Of the three different parametric GARCH VaR models the AR(1)-GARCH(1,1) model with skew Student t distributed innovation performed the best over all the quantiles tested. The semi-parametric Filtered Historical Simulation VaR model performed the best overall for quantiles 5%, 1% and 0.5%, while the conditional GPD VaR model performed very well when calculating quantiles 0.1% and smaller.

Our findings suggest that the use of EVT has a place in calculating VaR, but it must be used for the correct purpose, which is that of calculating very extreme quantiles. EVT becomes increasingly inaccurate as we move further away from the very extreme quantiles. Our findings further suggest that the Filtered Historical Simulation VaR method is best to apply when calculating VaR for the South African equity market for quantiles 5%, 1% and 0.5%, while the conditional GPD method is superior when calculating VaR at the 0.1% quantile.
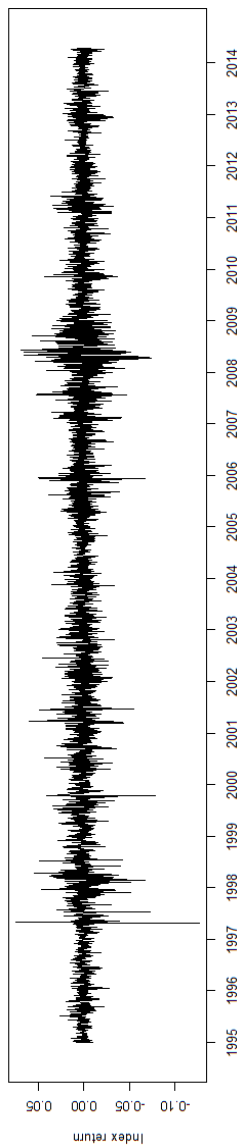
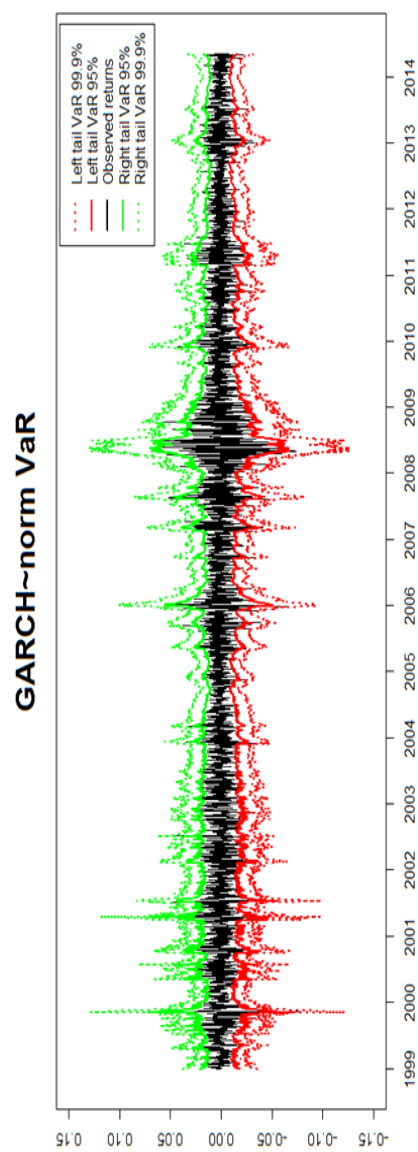**Figure 1: Daily returns form 30 June 1995 to 17 November 2014**



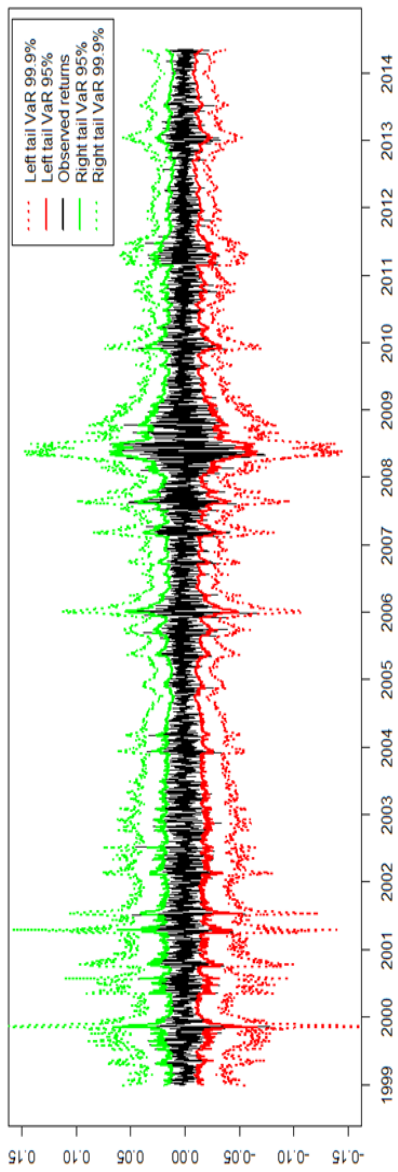**Figure 2: Back-testing the AR(1)-GARCH(1,1) VaR model with normal innovations at the 5% and 0.1% quantiles**

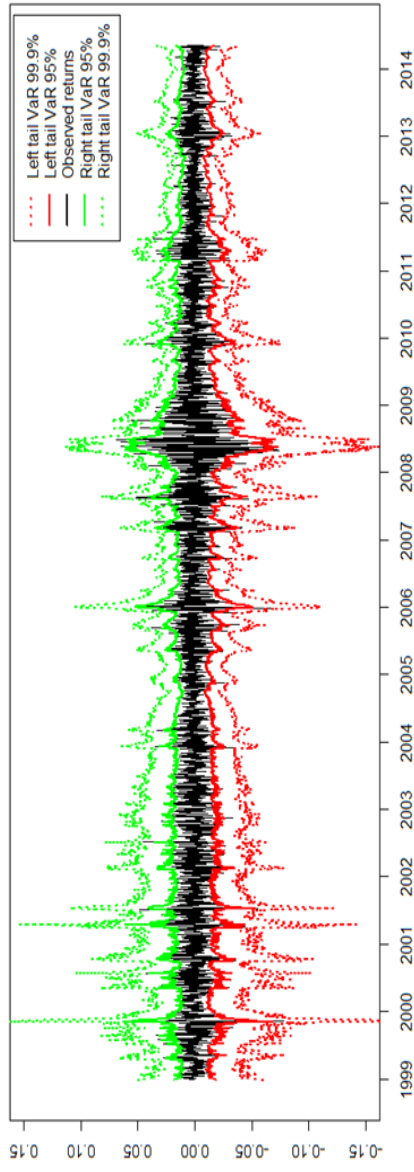**Figure 3: Back-testing the AR(1)-GARCH(1,1) VaR model with Student t distributed innovations at the 5% and 0.1% quantile**

**Figure 4: Back-testing the AR(1)-GARCH(1,1) VaR model with skew Student t distributed innovations at the 5% and 0.1% quantiles**
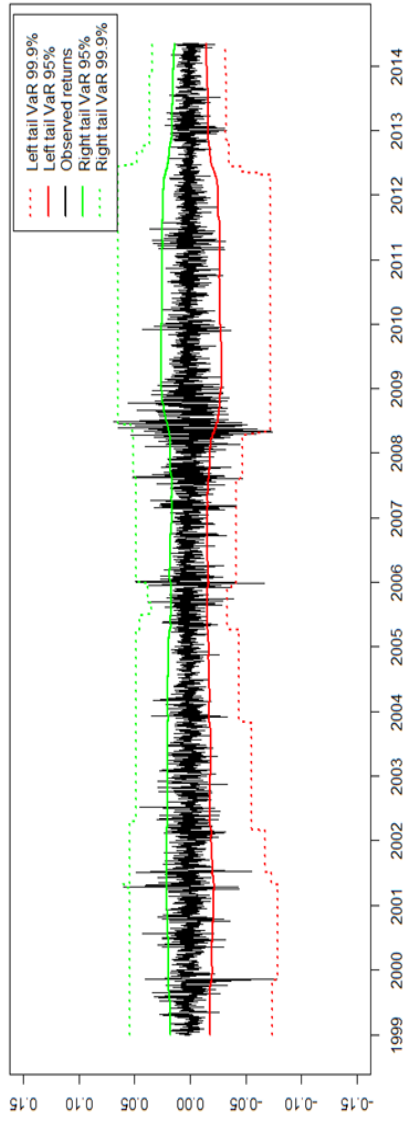
**Figure 5: Back-testing the Historical Simulation VaR model at the 5% and 0.1% quantiles**
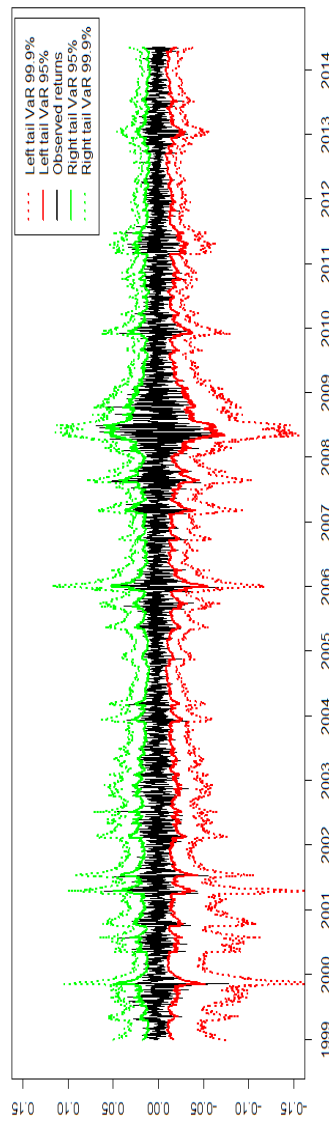


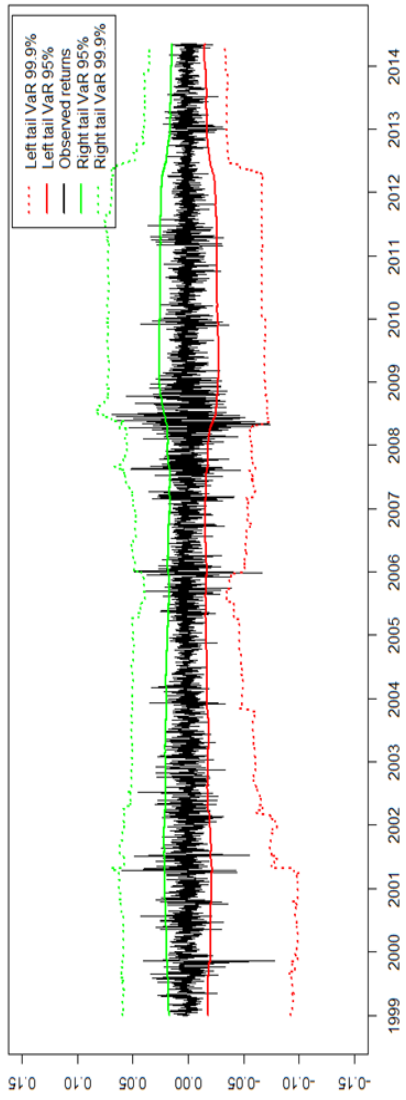**Figure 6: Back-testing the Filtered Historical Simulation VaR model at the 5% and 0.1% quantiles**

**Figure 7: Back-testing the Extreme Value Theory GPD VaR model at the 5% and 0.1% quantiles**
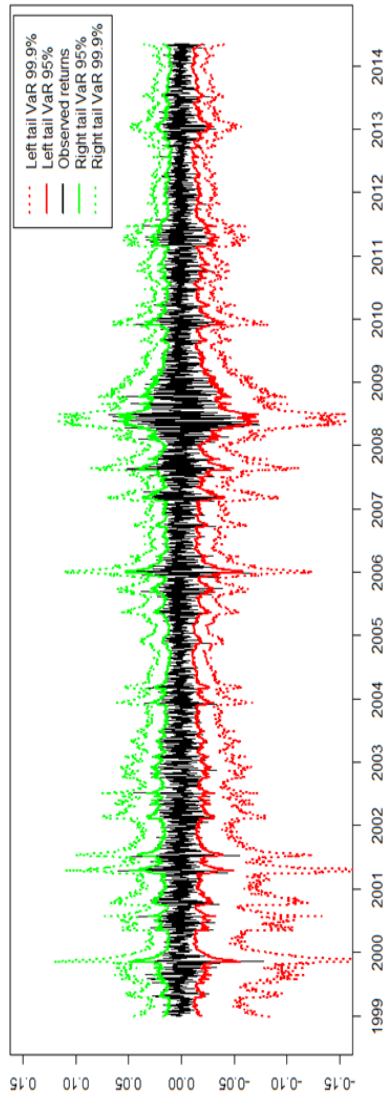
**Figure 8: Back-testing the conditional GPD VaR model at the 5% and 0.1% quantiles**

# References

Abad, P., Benito, S. & López, C. 2014. 'A comprehensive review of value at risk methodologies', *The Spanish Review of Financial Economics*, **12**(1): 15–32.

Alexander, C. 2009. *Market risk analysis, value at risk models* (Vol. 4). England: John Wiley & Sons.

Angelidis, T. & Benos, A. 2008. 'Value-at-risk for Greek stocks', *Multinational Finance Journal*, **12**(1): 67–104.

Balkema, A.A. & De Haan, L. 1974. 'Residual life time at great age', *The Annals of Probability*, **2**(5): 792 - 804.

Campbell, S.D. 2006. 'A review of backtesting and backtesting procedures', *The Journal of Risk*, **9**(2): 1–17.

Coles, S., Bawa, J., Trenner, L. & Dorazio, P. 2001. *An introduction to statistical modeling of extreme values* (Vol. 208). London: Springer.

Daníelsson, J. 2011. *Financial risk forecasting: The theory and practice of forecasting market risk with implementation in R and Matlab* (Vol. 588). United Kingdom: John Wiley & Sons.

Danielsson, J. & De Vries, C.G. 2000. 'Value-at-risk and extreme returns', *Annales d'Economie et de Statistique*, **60**: 239-270.

Dicks, A., Conradie, W. J. & De Wet, T. 2014. 'Value at risk using GARCH volatility models augmented with extreme value theory', *Journal of Studies in Economics and Econometrics*, **38**(3): 1 - 18

Diebold, F.X., Schuermann, T. & Stroughair, J.D. 2000. 'Pitfalls and opportunities in the use of extreme value theory in risk management', *The Journal of Risk Finance*, **1**(2): 30-35.

Gençay, R. & Selçuk, F. 2004. 'Extreme value theory and value-at-risk : Relative performance in emerging markets', *International Journal of Forecasting*, **20**(2): 287–303.

Gençay, R., Selçuk, F. & Ulugülyağci, A. 2003. 'High volatility, thick tails and extreme value theory in value-at-risk estimation', *Insurance: Mathematics and Economics*, **33**(2): 337-356.

Ghorbel, A. &Trabelsi, A. 2009. 'Measure of financial risk using conditional extreme value copulas with EVT margins', *Journal of Risk*, **11**(4): 51–85.

Jansen, D.W. & De Vries, C.G. 1991. 'On the frequency of large stock returns: Putting booms and busts into perspective', *The Review of Economics and Statistics*, **73**(1): 18-24.

Koedijk, K.G., Schafgans, M.M. & De Vries, C.G. 1990. 'The tail index of exchange rate returns', *Journal of International Economics*, **29**(1-2): 93-108.

Kupiec, P.H. 1995. 'Techniques for verifying the accuracy of risk measurement models', *The Journal of Derivatives*, **3**(2): 79-84.

Mancini, L. & Trojani, F. 2005. *Robust semiparametric bootstrap methods for value at risk prediction under GARCH-type volatility processes.* SSRN Working Paper. University of St. Gallen.

Marimoutou, V., Raggad, B. & Trabelsi, A. 2009. 'Extreme value theory and value at risk: Application to oil market', *Energy Economics*, **31**(4): 519–530.

McMillan, D. & Thupayagale, P. 2010. 'Evaluating stock index return value-at-risk estimates in South Africa: Comparative evidence for symmetric, asymmetric and long memory GARCH models', *Journal of Emerging Market Finance*, **9**(3): 325–345.

McNeil, A.J. & Frey, R. 2000. 'Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach', *Journal of Empirical Finance*, **7**(3-4): 271–300.

Nieppola, O. 2009. *Backtesting value-at-risk models*. Helsinki School of Economics. [Online] Available: https://aaltodoc.aalto.fi/handle/123456789/181

Pattarathammas, S., Mokkhavesa, S. & Nilla-Or, P. 2008. Value-at-risk and expected shortfall under extreme value theory framework: An empirical study on Asian markets. In 2nd European Risk Conference, Milano.

Pickands, J. 1975. 'Statisticak inference using extreme order statistics', *The Annals of Statistics*, **3**(1): 119 - 131.

Rocco, M. 2014. 'Extreme value theory for finance: A survey', *Journal of Economic Surveys*, **28**(1): 82–108.

Seymour, A.J. & Polakow, D.A. 2003. 'A coupling of extreme-value theory and volatility estimation in emerging markets: A South African test', *Multinational Finance Journal*, **7**(1/2): 3–23.