STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **10**, March 2022, pp 383–400.
Published online in International Academic Press (www.IAPress.org)

Extreme Value Modelling of the Monthly South African Industrial Index (J520) Returns

Owen Jakata *, Delson Chikobvu

Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, South Africa

Abstract This study uses Extreme Value Theory (EVT), Value-at-Risk (VaR) and Expected Shortfall (ES) analysis as a unified tool for managing extreme financial risk. The study extends the application of the generalised Pareto distribution (GPD) by modelling monthly South African Industrial Index (J520) returns (years: 1995-2018) to quantify the tail-related risk measures. The GPD is used to estimate the tail-related risk measures using the Peak over Threshold (PoT) method. Maximum Likelihood Estimates (MLE) of model parameters were obtained and the models goodness of fit was assessed graphically using Quantile-Quantile (Q-Q) plots, Probability (P-P) plots, scatter plots, residuals, return levels and density plots. The findings are that the GPD provides an adequate fit to the data of excesses (extreme losses or gains). Low frequency but very high or very low returns impact on investment decisions. Calculations of the VaR and ES tail-related risk measures based on the fitted GPD model are given. The results reveal that for an investment in the South African Industrial Index (J520), the prospect of potential extreme losses is less than the prospect of potential extreme gains. There seems to be an upper bound where losses do not seem to exceed easily. The study concluded that EVT, together with VaR and ES analysis are useful tools that can be applied in practice to manage index/stock price risk and help investors improve their investment decisions and trading strategies through better quality information derived from the tools. This study contributes to empirical evidence on EVT methods that help to protect financial systems against unpredictable fluctuations and losses of an extreme nature.

Highlights:

i. the GPD provides an adequate fit to the distribution of extreme financial returns for the South African Industrial Index (J520) returns.
ii. for an investment in the South African Industrial Index (J520), the prospect of potential extreme losses is less than the prospect of potential extreme gains.
iii. there seems to be an upper bound where losses do not seem to exceed easily.

Keywords Extreme Value Theory, Generalised Pareto Distribution, Peak-over-Threshold method, Value-at-Risk, Expected Shortfall, Tail-related risk, monthly South African Industrial Index (J520), South African equity market

AMS 2010 subject classifications 60G70

DOI: 10.19139/soic-2310-5070-906

1. Introduction

The last three decades have been characterised by Global financial instabilities, which include the Asian Financial Crisis (1997-1998), Global Financial Crisis (2007-2008) and the Chinese Stock Market Crash (2015-2016). These periods of economic crisis affected the Global financial markets, including the South African equity market. Traditional methods of predicting risk failed to quantify and predict the financial consequences. Extreme Value

*Correspondence to: Owen Jakata (Email: owenjakata@rocketmail.com). Department of Mathematical Statistics and Actuarial Sciences, University of the Free State, South Africa .

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Theory (EVT) modelling is applied to financial markets to predict risks that are extreme in nature using say, the Value-at-Risk (VaR) and Expected Shortfall (ES). In this study, risk is defined as the transformation of unforeseen future states into numerical values representing possible gains and losses [1].

According to [35], the traditional methods of calculating VaR and ES based on the normal distribution have limitations. ([21]), stated that EVT methods estimate tail-related risk measures with some degree of accuracy and certainty. According to [16], some sub-indices such as the South African Industrial Index (J520) are not always as informationally efficient, unlike the South Africas All Share Index (ALSI) which is said to be weak form efficient, thereby allowing some investors the opportunity of making excess profits/losses (extreme gains/losses) on the South African Industrial Index. The sub-indices of the ALSI are therefore modelled by Extreme Value Distributions (EVD) such as the Generalised Extreme Value Distribution (GEVD) and the Generalised Pareto distribution (GPD) models.

The EVT methods allow for the development of models that help investors, researchers and practitioners to estimate tail-related risk measures. The EVT method is a branch of probability and statistics, dealing with extreme fluctuations found in the tails of a distributions. Extreme equity returns require the application of the EVT methods for modelling tail-related events ([29]). According to [22], extreme events in finance could reveal possible losses leading to bankruptcy. Such events can also bring huge gains as well. The normal distribution is not suitable for modelling low frequency but with very high gains or very large losses since extreme events will occur with a greater frequency than that predicted by the normal distribution.

According to [33], the EVT methods are used for determining the probability of extreme events. According to [36], the two practical methods of the EVT are:

- the Block Maxima (BM) method.
- the Peak over Threshold (PoT) method.

The two parameter GPD uses the PoT method to extract excesses above a certain threshold and has proved to be one of the best ways to apply EVT in practice. The main problem is that of choosing the optimal threshold. According to [13], the GPD is better in modelling insufficient data than the three parameter GEVD which requires the use of large datasets. [12], states that BM method is inferior to the PoT method when using a financial time series. [34] states that there is no unambiguous answer as to which one is better between the BM and PoT methods.

[33] stated that the BM method avoids the dependency problem in the dataset which tends to complicate the use of the threshold method when applying the PoT method. The PoT method is one of the most popular models for extracting and fitting minima/maxima values above a sufficiently high threshold. The estimates of extreme events provided by the BM method may underestimate the extreme events in some cases ([29]). Studies by [31] and [32] show that the BM method is a good approach to financial risk estimation. However studies by [5], revealed that the PoT method is more preferable for tail-related estimation of VaR while BM method is more preferable for return level estimation. According to [30], the PoT method estimates are more accurate and consistent than the BM method. The VaR estimated by the GPD models are, more accurate than those estimated by traditional methods such as the historical simulation ([1]).

The PoT method utilises the data more efficiently to produce reliable findings and is therefore preferred compared to the BM method as all observations above a certain threshold are used to estimate parameters of the tail distribution ([13]). A major drawback of the alternative GEVD or BM method, is that, by extracting only the largest value in each block of returns, it ignores some secondary potential extreme values. According to [14], the GPD is preferred than the GEVD, since it exhibits less variability across all the levels of confidence, which minimises the chances of making incorrect calculations and hence investment decisions and trading strategies. In this study, the GPD model is therefore adopted.

This study models the monthly South African Industrial Index (J520) returns based on the GPD model and estimates the tail-related risk measures.

1.1. Statement of the Problem

The South African equity market was impacted negatively during the international financial crises which include: the Asian Financial Crisis (1997-1998), Global Financial Crisis (2007-2008) and the Chinese Stock Market Crash (2015-2016). The traditional methods failed to predict/forecast these financial crises in terms of their extreme nature. EVT provides a set of ready-made approaches that capture extreme events more accurately than traditional methods in quantifying extreme risk in advance. The estimated tail-related risk measures will be useful for minimising potential extreme losses and maximising on potential extreme gains of financial investments on the South African equity market.

1.2. Justification of Study

The extreme events in the equity market are the main issues of concern to investors, researchers and practitioners. The extreme events are associated with high-unexpected losses that can lead to bankruptcy. Other events can lead to huge and unexpected gains. If only we can understand this behaviour better. The objective and motivation behind this study is to analyse performance of the monthly South African Industrial Index (J520) using the EVT framework in order to understand the behaviour of returns at the two extremes: huge losses and huge gains.

1.3. Objectives of the Study

This study aims at extending the application of the GPD by modelling extremes associated with monthly South African Industrial Index (J520), based on the EVT framework. The specific objectives of this study are to:

- to fit the GPD model to the monthly South African Industrial Index (J520) log returns.
- estimate tail-related risk measures of VaR and ES based on the GPD model.

The research into the behaviour of the extremes will contribute to literature on the South African equity market and its sub-indices. It differs from other studies in that it uses the monthly South African Industrial Index (J520) extremes to model and quantify tail-related risk measures. The study is organised as follows: Section 1 gives the introduction, Section 2 Reviews the Literature, section 3 discusses the research models, Section 4 is on data analysis and discussions of results. Section 5 concludes and gives areas of further study.

2. Review of Literature

The use of EVT is of great importance in economics and finance. The results of some studies by [11] , [4] and [2] showed that EVT methods fit the tails of fat-tailed financial time series better than the traditional distributional approaches. These studies revealed that EVT is the best approach in estimating the tail of a loss distribution. The estimation of the tail of a loss distribution is used to determine and quantify extreme risk. Studies by [17], [19], [22], [20] and [37] similarly apply the GPD model to extreme financial returns.

[17] applied the GPD model to estimate tail-related risk measures of the South African Financial Index (J580) returns. The PoT method was used to extract the excesses above a certain threshold and fitted the GPD model to estimate the parameters which are then used to estimate the tail-related risk measures in the form of VaR and ES. The study reveals that the upside risk is greater than the downside risk for one invested in the index.

[19] fitted the peak-over-threshold GPD model to returns of the daily All Share Index returns (years: 2008 to 2015) on the Nairobi Stock Exchange (NSE).One full stop only. Parameter estimates derived using the MLE method gave a good fit. QQ and density plots were used to assess the goodness of fit of the data. Estimates of risk measures of VaR and ES were calculated. The conclusion was that, the prospect of potential losses was lower than

the prospect of potential gains, if one were to invest on the NSE.

[22] applied the EVT framework to the tails of the All-Shares index returns of the Ghana Stock Exchange (GSE) (years: 2000-2010). The excesses above a certain threshold from the PoT method were fitted to the GPD model. Parameter estimates derived using the MLE method gave a good fit. Quantile-Quantile (QQ), Probability (PP) and density plots were used to test the goodness of fit of the model. Their results indicated that the GPD model provides a good fit to the excesses above a certain threshold. Estimates of VaR and ES at high quantiles were calculated. The conclusion was that, the prospect of potential losses (downside risk) was lower than the prospect of potential gains (upside risk) for an investment on the GSE.

[20] researched on estimation VaR and ES on the NSE using the EVT framework. Their objective was to analyse the tails of the distribution at high quantiles and then estimate VaR and ES using the EVT framework. In their study, Barclays Bank stock returns data were analysed at different confidence levels. The researchers concluded that the GPD model is a preferred method for estimating VaR and ES. The risk metrics of VaR and ES were estimated at high quantiles of 99.9 per cent and 95 per cent, and were used to capture and quantify the associated financial risks.

[37], used the GPD to estimate tail-related risk of the Shanghai stock market data. The researchers estimated the VaR and ES at the 95th and 99th percentiles in an attempt to describe the tail of a loss distribution. The researchers concluded that the GPD is a useful complement to traditional VaR and ES methods in quantifying risk.

Many studies cited in literature have analysed extreme events in finance and economics. These studies also include [15], [36], [23], [8], [6], [10], [24], [13] and [9] who discussed the behaviour of tail-related distributions using the GPD . This study adopts and extends the application of the GPD to model tail-related measures of VaR and ES at high quantiles. This study will contribute to the knowledge and development of EVT models that help to protect financial systems against unpredictable fluctuations and losses of an extreme nature. The models will help investors, researchers, practitioners and policy makers in South Africa to measure risk and optimise the returns of financial investments through better quality information.

3. Research Models

3.1. Extreme value theory

EVT provides us with a set of methods for modelling and predicting extreme risk on the South African equity market.
In this study, the PoT method is adopted. The method uses the GPD with two parameters, and is widely used for modelling excesses above a certain high threshold. The main concept of the method is to extract the excesses above a certain high threshold and then fitting the GPD model. This is done by setting a threshold μ to be some value which exceeds most but not all values defined in the return series. For some sufficiently large threshold μ, the distribution of the values exceeding the threshold will converge to a GPD with some certain shape and scale parameters.

3.1.1. Generalised Pareto Distribution (GPD) Model

The EVT framework by [26] and [3] models the asymptotic behaviour of the exceedances with a two parameter GPD with scale parameter (β) and shape parameter (ξ). The generalized Pareto distribution which models observations exceeding a certain high threshold μ is defined as:

$$G\xi,\beta(x) = \left\{ \begin{array}{ll} 1-\left(1+\frac{\xi(x-\mu)}{\beta}\right)^{\frac{-1}{\xi}}, & \text{if} \quad \xi \neq 0, \\ 1-exp^{-\frac{(x-\mu)}{\beta}} & \text{if} \quad \xi = 0. \end{array} \right\}, \tag{1}$$

where $x > 0$ when $\xi \geq 0$, $0 \leq x \leq \frac{\beta}{\xi}$ when $\xi < 0$ and $\beta > 0$.
ξ is the shape parameter, β is the scale parameter and x represents the return. $(x-\mu)$ is the exceedance above the threshold μ. If $\xi > 0$, then, $G_{(\xi,\beta)}$ is used for very large losses ([25]), if $\xi = 0$ then it is a case of a light tail and $\xi < 0$ then, it is a case of a Pareto type II distribution which is bounded.

3.1.2. *Mean Excess Distribution.*

For a random variable X with a density function F, the excess distribution above a certain threshold μ is defined by:

$$F_u(x) = P(X - \mu \leq x | X > \mu) \tag{2}$$

where $(x - \mu\ )$ represents the size of exceedances over μ. If F is denoted as the distribution function for X, then we may write:

$$F_u(x) = \frac{F(x+\mu) - F(\mu)}{1 - F(\mu)} = \frac{F(x) - F(\mu)}{1 - F(\mu)} \tag{3}$$

F_μ is the conditional excess distribution function.
According to Ren and Giles (2010) the tail estimator for the distribution is given by

$$\hat{F}(x) = 1 - \frac{N_\mu}{n}\left(1 + \frac{\xi}{\beta}(x-\mu)\right)^{\frac{-1}{\xi}} \tag{4}$$

Where n is the total sample size, and N_μ is the number of exceedances above the threshold μ.

3.2. *Measures of Risk*

The most frequently used measures of risk in extreme quantile estimation are VaR, ES and return level. The tail-related risk measures of the right tail and the left tail are important when considering risk.

3.2.1. *Value at risk (VaR)*

VaR is the maximum potential change in value with a given probability over a given period for an equity portfolio. It is an extreme amount, which will be lost over a fixed period of time within some degree of confidence. Alternatively VaR is defined as sufficient capital to cover the potential losses of an equity portfolio over a fixed period.

The inverse of Equation 4 (value of x in Equation 4) with a probability p gives the VAR:

$$\widehat{VaR_p} = \mu + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_\mu}(1-p)\right)^{-\hat{\xi}} - 1\right) \tag{5}$$

VaR is not a coherent risk measure which implies that the risk measures may not behave sensibly under certain conditions.

3.2.2. *Expected shortfall (ES)*

ES is a downside risk measure which is a potential measure of a loss exceeding VaR. According to [7] VaR does not capture the aspect of subadditivity but ES is sub-additive making it a coherent risk measure. ES is also defined as the expected amount of return that exceeds VaR for a certain probability level.

The Expected Shortfall (ES) for the GPD when $\xi < 1$ is:

$$ES_p = \frac{VarR_p}{1-\xi} + \frac{\sigma - \xi\mu}{1-\xi} \tag{6}$$

3.3. Threshold determination

The data was grouped into two sets, negative returns (representing losses) and positive returns (representing gains). In choosing a suitable threshold value for each group, there should be a compromise between choosing a high threshold and a low threshold value. For a high threshold value, bias is reduced as convergence towards the extreme value theory is achieved. There will be fewer data though for estimating the parameters. This increases the variance for the estimators of the parameters of the GPD. A high bias and a low variance and more data for estimation of the parameters will occur as a result of adopting a low threshold value. The three methods that are adopted and used to determine the optimal threshold value are: the Pareto QQ plot, the Hill plot and the Mean excess plot (Mean residual live plot).

3.3.1. Pareto QQ plot

The Pareto QQ plot uses a tangent line to the data points plotted. The value at which the line crosses the y-axis is determined and its exponent is calculated as the threshold required. The ordinary least squares estimation method of an intercept to a regression line is used to estimate this parameter.

3.3.2. The Hill plot

The Hill plot is a graphical tool used to determine the threshold. Let $x_1 > x_2 ... > x_n$ be the ordered statistic of random variables which are independent and identically distributed. Using $k+1$ ordered statistics, the Hill estimator of the tail index is defined by;

$$\hat{\xi} = \hat{\alpha}^{-1} = \frac{1}{k} \sum_{i=1}^{k} lnx_{in} - lnx_{kn} \tag{7}$$

Where $k = k(n) \longrightarrow \infty$ is upper order statistics, n is defined as the sample size and $\alpha = \frac{1}{\xi}$ is defined as the tail index. Using a stable area of the tail of the plot, the threshold μ is estimated.

3.3.3. The mean excess plot/mean residual live plot

According to the findings of [26] and [3]: for a certain threshold u, the excesses above the threshold converge to a GPD. The threshold can be determined by identifying a steady area on the mean excess graph where an approximation by the GPD model is reasonable. In order to adopt the PoT method, the choice of the threshold is critical in order to model the tails of the distribution. The mean excess function for the GPD is:

$$e(\mu) = E(X - \mu X > \mu) = \frac{\sigma + xi\mu}{1-\xi}, \quad \sigma + \xi\mu > 0 \tag{8}$$

σ is a constant. This function is used as a criterion for the selection of μ, $e(\mu)$ is linear and should remain linear in the tail.

From the mean excess function, the sample mean excess function is estimated:

$$\hat{e}_{\mu} = \frac{\sum_{n=1}^{N}(X_n - \mu)I(X_n > \mu)}{\sum_{n=1}^{N} I(X_n > \mu)} \tag{9}$$

where $I(X_n > \mu)$ is an indicator function that is equal to one if $X_n > \mu$ and zero otherwise.

3.4. Parameter estimation

The maximum likelihood method is used to estimate the parameters of the GPD. For an independent and identically distributed sample of size n, the joint density function is given as:

$$f(x_1, x_2, ..., x_n|\theta) = f(x_1|\theta).f(x_2|\theta)...f(x_n|\theta) \tag{10}$$

where $\theta = (\xi, \beta)$ are the parameters of the model $f(x_i\theta)$. Therefore the observed variables are known whereas the parameters given by $\theta = (\xi, \beta)$ for given threshold u are to be estimated. The likelihood function is then given by:

$$lnL(\theta|x_1, ..., x_n|\theta) = \sum_{i=1}^{n} f(x_i|\theta) \tag{11}$$

The estimated parameters are then given by the set $\hat{\theta} = (\hat{\xi}, \hat{\beta})$ which maximises the likelihood function, equation (11) or (12).

3.5. Test for Normality, Stationarity, Heteroscedasticity and Auto-correlation

To test for normality of the monthly South African Industrial Index (J520) return distribution, the QQ graphical plot is used. Stationarity tests are applied to determine whether the data is stationary. The Augmented Dickey Fuller (ADF) test is used to test for stationarity of the monthly South African Industrial Index (J520) returns. The ARCH LM test is used to test for the presence of heteroscedasticity. The ACF and PACF and Box-Ljung test are used to test for auto-correlation of the return distribution

4. Data analysis and Discussion

This study applied the GPD in analysing the monthly South African Industrial Index (J520) returns over the period 1995 to 2018. Data analysis was done in an R programming environment using packages: ismev, ReIns, evd, evir, fExtremes and extRemes.

4.1. Data description

The study employed monthly South African Industrial Index (J520) data, accessed from the website iress expert: https://expert.inetbfa.com (with permission). The data consisted of 272 monthly South African Industrial Index (J520) returns (years: 1995-2018). This index is for values of stocks based on industrial firms listed on the South African equity market. This index measures the overall performance of the South African equity market for this specific industry.

The South African equity market firms are listed under different sectors that take the firm's main business activities into account and groups companies that share general industrial and economic themes together. For every industrial sector on the South African equity market,there is an index, listing all the companies in that sector. The listed companies are in particular industrial sectors. The broad economic sectors are Basic Materials, Oil and Gas, Industrials, Consumer Goods, Consumer Services, Telecommunications, Financials and Exchange Traded Funds.

In this study, the South African Industrial Index (J520) is analysed, and consists of companies in the following sub-categories of indices, Construction and Materials (J235), Aerospace and Defence (J271), General Industrials (J272), Electronic and Electrical Equipment (J273), Industrial Engineering (J275), Commercial Vehicles and Trucks Industrial Transportation (J277) and Support Services (J279). In terms of market capitalisation, the Industrial sector is the most represented sector on the South African equity market.

In parts of this study, losses are positive (right tail) since the loss function in period t for an index log return X is:

$$Y_t = -r_t = -ln(\frac{M_t}{M_{t-1}}) \tag{12}$$

r_t is the monthly log returns in month t, M_t represents the monthly index in month t and ln represents the natural logarithm. The goal is to fit the GPD to the log returns. Notice the left tail of the loss function is the right tail of the return function and represents gains. The right tail of the loss function represents the left tail of the return function and represents losses.

4.2. Descriptive statistics

Table 1: Descriptive statistics for monthly log returns r_t of the South African Industrial Index (J520).

Table 1. Description Values for r_t

Number of Observations	271
Mean	-0.009366
Median	-0.010478
Maximum	0.328471
Minimum	-0.140273
Variance	0.003302
Standard Deviation	0.057467
Skewness	1.016932
Kurtosis	4.420852

The descriptive statistics are given in Table 1 for r_t. The maximum and minimum values are observed to be far from the mean. This confirms the existence of extreme events in the return series. The skewness is positive indicating the presence of extreme values in the right tail. Kurtosis is greater than 3, which suggests that the return series is fat-tailed. These different characteristics observed in this return series allow us to infer that the return series is fat-tailed.

4.3. Graphical plot of the monthly South African Industrial Index (J520) levels.

The time series plot shows a clear upward trend (Figure 1). The log return data r_t is in Figure 2. The Asian Financial Crisis (1997-1998), the Global Financial Crisis (2007-2008) and the Chinese Stock Market Crash. (2015-2016) had a negative effect on the South African equity market which are indicated on the time series plot by sharp down turns of the index levels (Figure 1). According to Onour (2010), as equity markets become increasingly involved in global trade, they become more vulnerable to shocks in global financial markets and become more volatile. Hence the need for extreme risk analysis of the stock and index returns of the South African equity market.

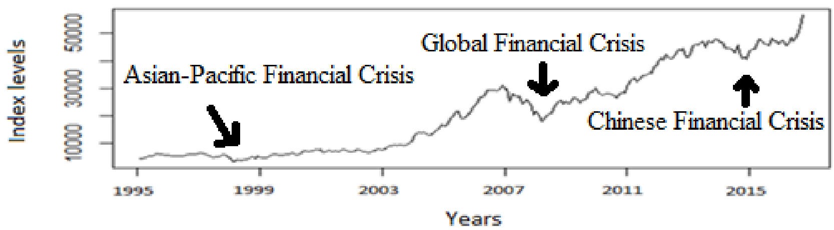

Figure 1. Time series plot of M_t, the monthly South African Industrial Index (J520) levels.

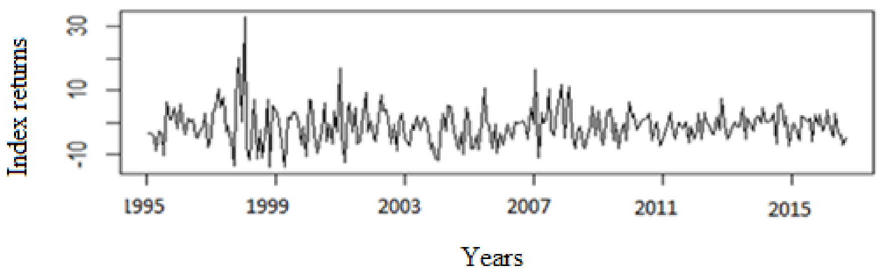

Figure 2. Time series plot of $r_{t,}$ monthly returns for the South African Industrial Index (J520) levels.

4.4. Testing for stationarity, normality, heteroscedasticity and autocorrelation results

4.4.1. Testing for Stationarity

The ADF Test is used to test for stationarity and the results show that a p-value $<$ 0.05 was obtained: Dickey-Fuller = -6.7391, Lag order = 6, p-value = 0.01. It is concluded that the returns data is stationary.

4.4.2. Test for Normality

Two methods are used in this study to test for normality: (i) the Andersen Darling Test and (ii) the Q-Q plot.

(i) Andersen Darling Normality Test

To test if the monthly South African Industrial Index (J520) data is normally distributed, the Andersen Darling Normality Test is used. A p-value $<$ 0.05 is obtained which means that the monthly data series is not normally distributed. This suggests that the returns data is fat-tailed.

(ii) Q-Q plot for the monthly log returns r_t

Some positive returns (gains) points are above the straight line implying that the data is not normally distributed, and suggesting that the data is fat-tailed.

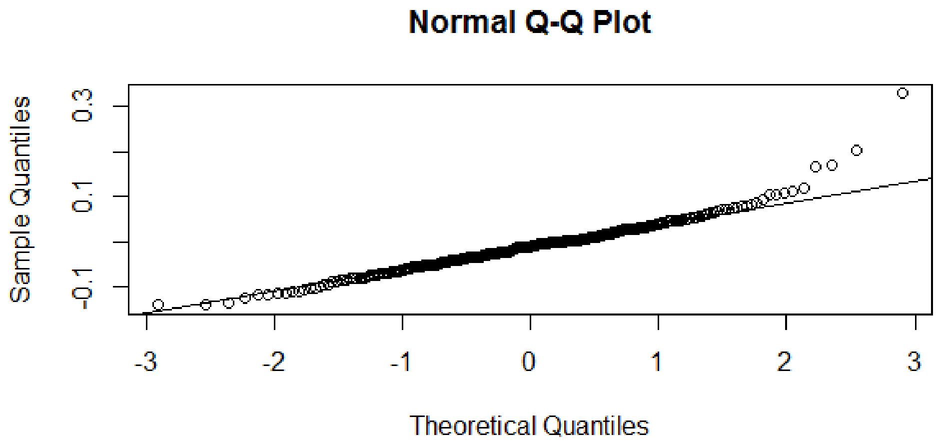

Figure 3. QQ plot for r_t, the South African Industrial Index (J520) monthly log returns

4.4.3. Test for Heteroscedasticity

The Arch LM Test is used to test for the existence of ARCH effects in the monthly South African Industrial Index (J520) returns. The results indicate that there are no significant ARCH effects that exist in the returns data (χ^2 = 8.366974, df = 12, p-value = 0.7558355). Therefore the heteroscedasticity of the monthly returns in this study are insignificant hence the adoption of the unconditional (static) models for modelling using the PoT method .

4.4.4. Test for autocorrelation

For the EVT methods to be applied, the returns data has to be independent and identically distributed. Two tests were carried out:
(i) The ACF and PACF are used to check for autocorrelation.

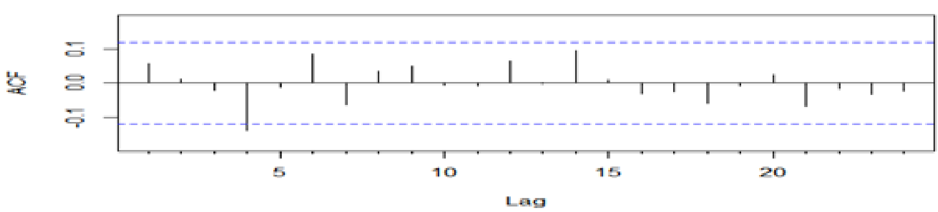

Figure 4. ACF diagram of r_t.

The ACF and PACF indicate that there are no significant autocorrelations in the data.

(ii) The Box-Ljung test for autocorrelation of the monthly South African Industrial Index (J520) return series was performed.

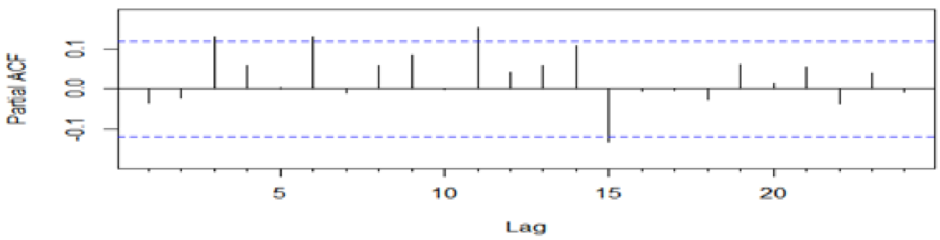

Figure 5. PACF diagram of r_t.

χ^2 = 0.8806, df = 1, p-value = 0.348.

The above test results show that autocorrelation is insignificant since the p-value $>$ 0.05 is obtained. This confirms that the return distribution is independently distributed.

4.5. Histogram of monthly log returns of South African Industrial Index (J520) data.

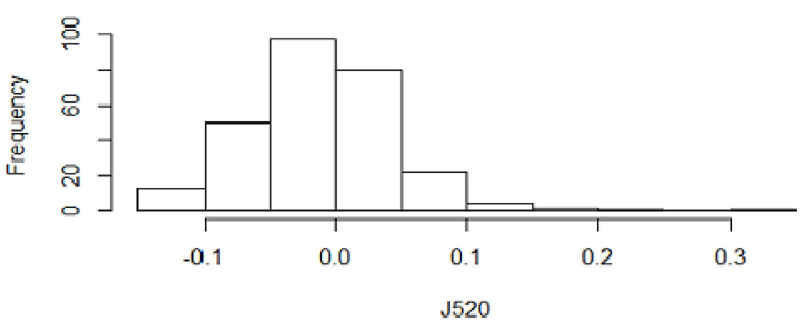

Figure 6. Histogram of monthly log returns of r_t, the South African Industrial Index (J520)

In Figure 6, the histogram shows that observations are relatively asymmetrical about the mean. The histogram shows that observations are asymmetrical about the mean which is consistent with the positive skewness value obtained. A high peak observed corresponds to the large kurtosis value obtained. The findings confirm that the data is fat-tailed.

4.6. Threshold determination results

The data was separated into two sets, negative returns (representing losses) and positive returns (representing gains). The negative and positive were modelled separately using the GPD. The Pareto QQ plot, Hill plot and Mean excess plot were used to determine the threshold for both the right and the left tail.

4.6.1. The Pareto QQ Plot

In this study, the Pareto QQ plot is one of the methods used to estimate the threshold.

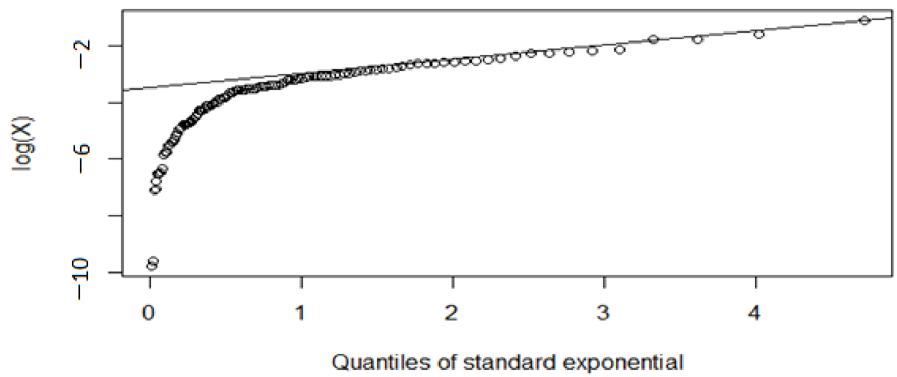

Figure 7. Pareto QQ diagram used for estimating the threshold for positive returns (gains).

Using the tangent line to the data points in Figure 7, we determined that it crosses the y-axis at -3.5 and therefore its exponent is 0.03. This is the threshold for the positive returns (gains).

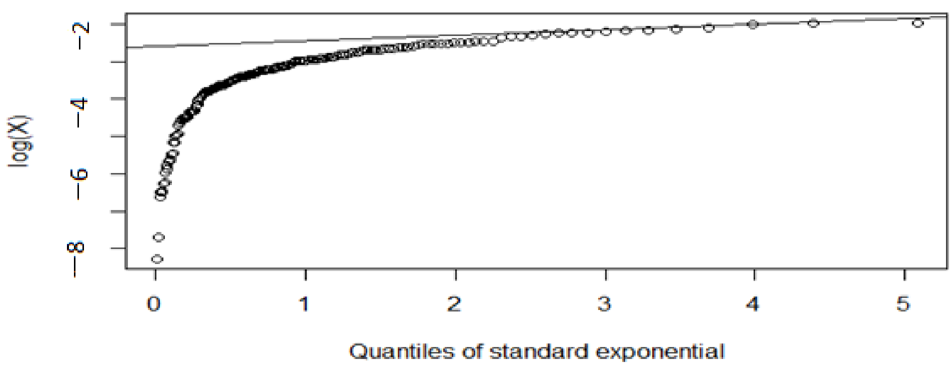

Figure 8. Pareto QQ diagram used for estimating the threshold for negative returns (losses)

In Figure 8, using the tangent line to the data points, we determined that it crosses the y-axis at -2.6 and therefore its exponent is 0.07 meaning that is the threshold estimate for the negative returns (as a loss function, Y_t).

4.6.2. Hill plot

The Hill plots for the right and left tail are in Figure 7 and Figure 8. The threshold is chosen from an area which is relatively steady on the plot, which provides enough excesses above the threshold for the GPD to be fitted. For the left tail of the loss function, Y_t, the threshold was observed to be between 57 and 64 order statistics (Figure 7), which implies the optimal threshold will to lie between this range, confirming our threshold of 0.03.

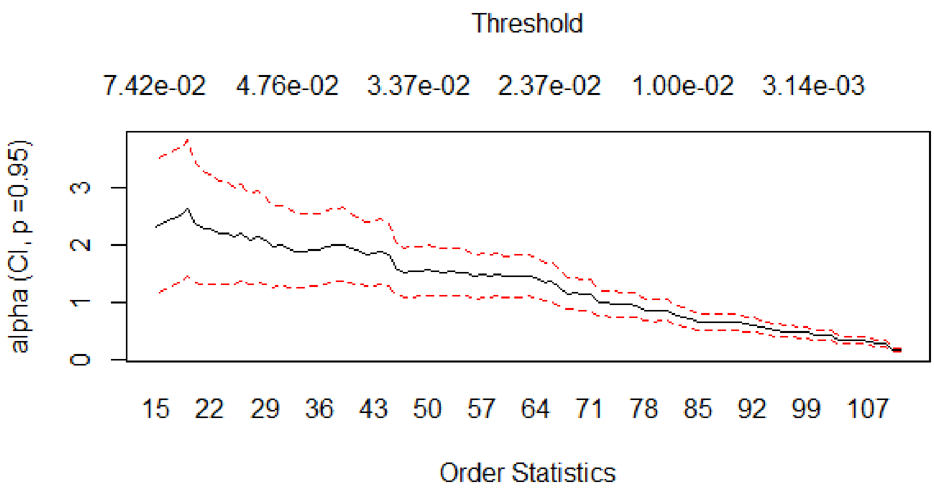

Figure 9. Hill plot for the positive returns, representing the gains (left tail of the loss function)

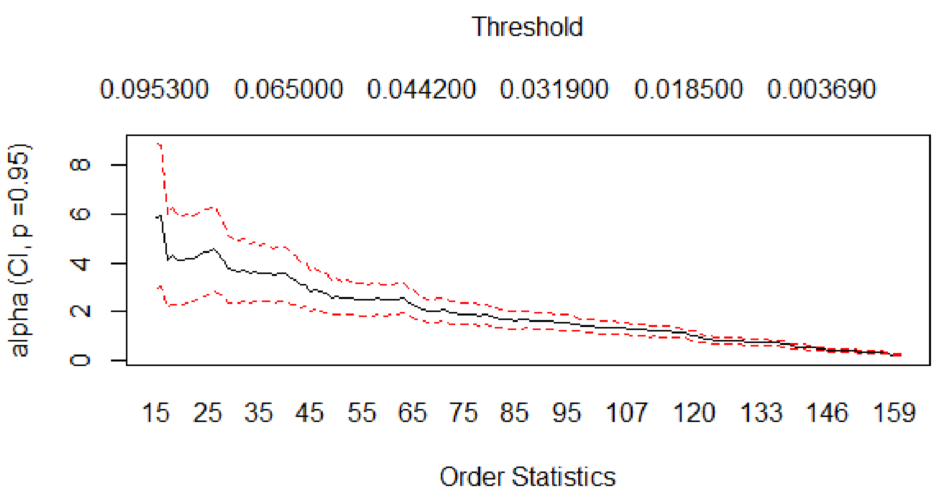

Figure 10. Hill plot for the negative returns, representing the losses (right tail of the loss function).

For the right tail of the loss function, the threshold was determined to be between 25 and 35 order statistics. This implies that the threshold lies within this range, confirming our threshold of 0.07.

4.6.3. Mean Residual Plots

The graph of the mean excess or mean residual live against possible thresholds plot is in Figure 11. The aim of the mean excess line is to relate it to the shape parameter ξ, thereby facilitating the fitting of a GPD distribution to the empirical excess distribution. The mean excess function is a constant when $\xi = 0$ (e.g. an exponential loss distribution). However, for other loss distributions, the mean excess function will slope upwards if ξ is positive and downwards if ξ is negative. The graph in figure 9 slopes upwards from about $\mu = 0.03$. The up and down movement from $\mu = 0.12$ can be attributed to the scarcity of data in the tails. The threshold is again maintained at 0.03 where the graph has a positive slope and is steady.

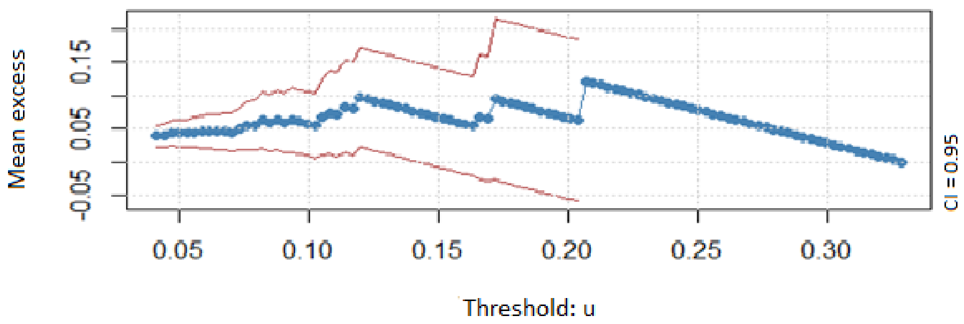

Figure 11. Mean Residual plot for the positive returns representing the gains (left tail of the loss function).

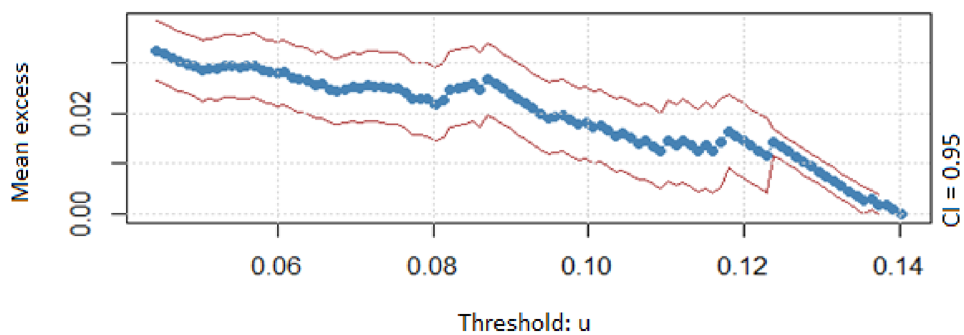

Figure 12. Mean Residual plot for the negative returns, representing the losses (right tail of the loss function)

The thresholds were determined to be 0.07 for the right tail of the loss function (negative returns) in Figure 12.

4.7. Model diagnostics plots for the right and the left tail

Figure 13 and Figure 14 provides the diagnostic plots for the positive and negative returns.

The diagnostic plots indicate a good fit to the models. The Quantile (QQ) plots for both right and the left tail of the loss function, the plots do not move away from a straight line significantly. The density plots confirm that the GPD models provide good fits to the excesses above the threshold. The Return Level plots in Figure 13 and Figure 14 indicate that there are no significant departures from the curve.

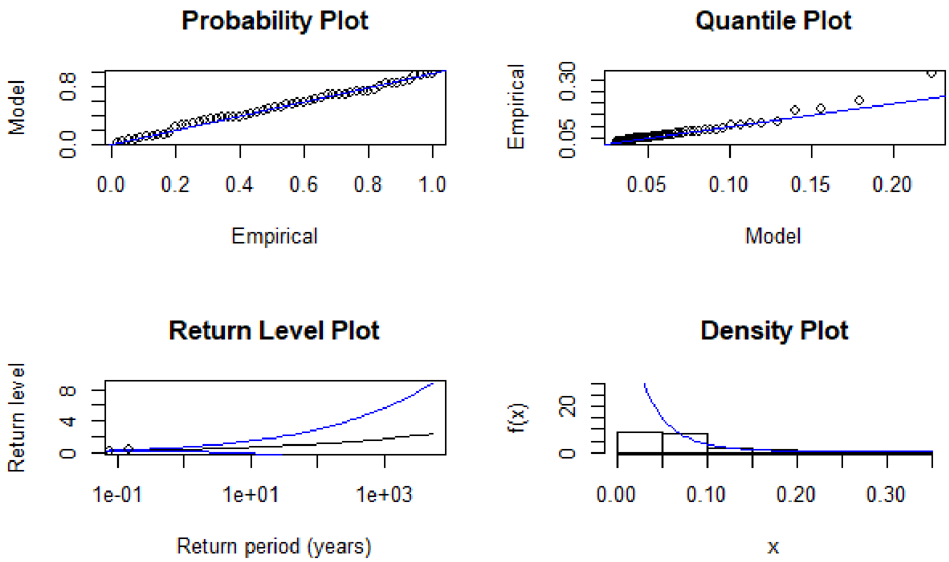

Figure 13. Diagnostic plots for the positive returns, representing the gains (left tail of loss function).

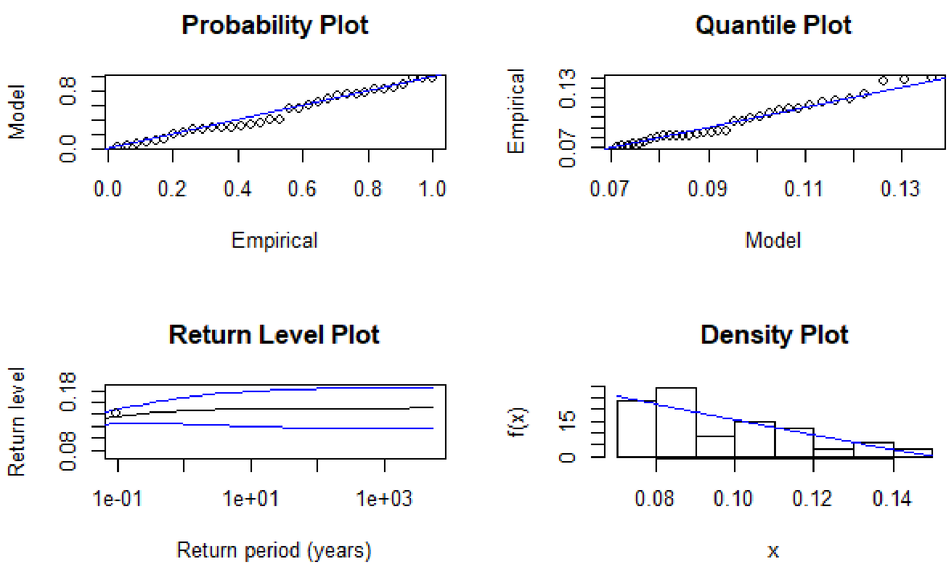

Figure 14. Diagnostic plots for the negative returns, representing the losses (right tail of loss function).

4.8. Fitting the Generalised Pareto Distribution model and estimation of parameters.

The GPD model was fitted separately to the positive and negative returns. The shape parameter and the scale parameter from the fitted GPD model and their corresponding standard errors are shown in Table 2.

Table 2. Estimation of parameters using maximum likelihood estimate (MLE).

	Right tail of loss function (Negative returns, losses)	Left tail of loss function (Positive returns, gains)
	$\mu = 0.07$	$\mu = 0.03$
Shape Parameter	-0.48207096	0.182889166
Standard error	0.22388147	0.149031473
Scale parameter	0.03894339	0.03262922
Standard error	0.01044674	0.006492347

The shape parameter for the left tail of the loss function (gains) is $\xi = 0.1829 > 0$, showing the distribution is fat-tailed on this side of the distribution. The maximum likelihood estimate for standard errors for the right tail (0.22) is greater than the left tail (0.18). This indicates that the left tail is a more desirable model compared to the right tail.

The shape parameter for the right tail of the loss distribution (losses) is $\xi = -0.4821 < 0$ showing the distribution belongs to the short-tailed class. This has a finite upper bound indicating an absolute maximum. In finance one cannot fix an upper bound for losses and can only infer that losses are unlikely to exceed this upper bound. The estimated tail parameters for the monthly South African Industrial Index (J520) returns for both the right and left tails indicate the prospects of significant extreme losses (downside risk) and significant extreme gains (upside risk).

4.9. Financial Risk Measures

The tail-related risk measures for the right and left tails tail are presented in Table 3.

Table 3. Financial risk measures for VaR and ES for the right and left tail.

Right tail of loss function (Negative returns, losses)		
Level of significance	Value at Risk (VaR)	Expected shortfall (ES)
0.950	0.1317134	0.1379103
0.990	0.1419954	0.1448474
0.995	0.1444861	0.1465277
Left tail of loss function (Positive returns, gains)		
0.950	0.1602247	0.2293516
0.990	0.2659208	0.3587309
0.995	0.3219556	0.4273214

VaR and ES are used for estimating risk quantities at a given levels of significance (Table 3). ES is a measure of downside risk, which estimates the potential size of the loss/gain exceeding VaR. In this study the expected monthly loss will not go beyond 13.17% and if it does, the loss will be 13.79% at the 95 per cent level of significance. The expected monthly gain will not exceed 16.02% and if it does, the gain will be 22.94% at the 95% level of significance. The interpretation is the same at 99% and 99.5% level of confidence.

The results reveal that the prospects of potential extreme losses on the South African Industrial Index (J520) is less than the prospects of potential extreme gains. There is an upper bound on losses where losses do not seem to exceed a certain value. This is consistent with the findings from [22], [19] and [17] who found that the prospects of potential extreme losses is lower than the prospects of potential extreme gains on different sets of data. Gilli and Këllezi (2006), found that the downside risk is higher than the upside potential gains for DJ Euro Stoxx 50, Swiss stock Market and Hang Seng stock market indices, this is in contrast with this study, but the markets are very different.

5. Conclusion and areas of further study

5.1. Conclusion

The main objective of the study was to extend the application and use of the EVT framework in modelling the monthly South African Industrial Index (J520) return distribution. The study applied the PoT method which fitted the excesses above a certain threshold to the GPD model. The estimated thresholds of $\mu = 0.07$ and $\mu = 0.03$ for the negative returns (right tail of loss function/losses) and positive returns (left tail of loss function/gains) respectively. The parameters were estimated using the MLE method. Results indicated that the prospects of potential extreme losses and prospects of potential extreme gains are significant. It was concluded that both thresholds provide a good fit for the right and left tails as confirmed by the diagnostic plots.

The tail-related risk measures of VaR and ES are used to evaluate the prospects of potential extreme losses/gains. The estimation of the tail-related risk measures are determined at high quantiles of 95%, 99% and 99.5%. The study revealed that the prospects of potential extreme losses (downside risk) for the South African Industrial Index (J520) investment is less than the prospects of potential extreme gains (upside risk).

Summary of main findings

- the GPD provides an adequate fit to the distribution of extreme financial returns.
- low frequency but very high or very low returns impact on investment decisions.
- the prospects of potential extreme losses are less than the prospects of potential extreme gains.

An accurate estimation of tail-related risk measures of VaR and ES is necessary to determine extreme risk exposure so that investors and risk analysts do not overestimate or underestimate the risk level. EVT provides the appropriate models. This study will contribute to the knowledge and development of EVT models that help to protect financial systems against unpredictable fluctuations and losses of an extreme nature. These models will help investors, researchers, practitioners and policy makers in South Africa to estimate risk and optimise the returns of financial investments.

5.2. Areas of further study

Suggestions for further studies may include modelling bivariate loss distributions of the South African equity market using the EVT approaches and Copula methods.

REFERENCES

1. W. Walls, and W. Zhang *Using Extreme Value Theory to Model Electricity Price Risk With an Application to the Alberta Power Market*, Energy exploration and exploitation, vol 23, no.5, pp. 375-404, 2006.
2. E. Brodin, and C. Kluppelberg *Extreme value theory in finance: A survey* Journal of Economic Surveys, vol. 28, no. 1, 2006, doi: 10.1111/j.1467-6419.2012.00744.x.
3. A.A. Balkema, and L. de Haan *Residual Life Time at Great Age* The Annals of Probability, vol 2, no. 5, pp. 792-804, 1974. Available at: http://projecteuclid.org/euclid.aop/1176996548.
4. Y. Bensalah *Steps in Applying Extreme Value Theory to Finance:* A Review, Working Paper 2000-20, Bank of Canada., 2000.
5. A. Bucher, and C. Zhou *A horse racing between the block maxima method and the peak-over-threshold approach*, pp. 1-19, 2018, http://arxiv.org/abs/1807.00282
6. B. O. Apudo et al. *An Application Of Extreme Value Theory In Modelling Electricity*, Mathematical Theory and Modelling, vol 4, no. 4, pp. 151-162, 2014.
7. P. Artzner et al. *Coherent measures of risk* Mathematical Finance, vol 9. no. 3, pp. 203-228, 1999, doi: 10.1111/1467-9965.00068.
8. C.K. Chege, J.K. Mungatu, and O. Ngesa *Estimating the Extreme Financial Risk of the Kenyan Shilling Versus Us Dollar Exchange Rates*, Science Journal of Applied Mathematics and Statistics, vol. 4, no. 6, p. 241-249, 2017, doi: 10.11648/j.sjams.20160406.11.

9. F.X. Diebold, T. Schuermann, and J.D. Stroughair *Pitfalls and opportunities in the use of extreme value theory in risk management*, Journal of Risk Finance, vol. 1, no. 2, pp. 30-35, 2000, doi: 10.1108/eb043443.
10. N.J. de Dieu, N.P. Mwita, and N.J. Mungatu *Estimation of Extreme Value at Risk in Rwanda Exchange Rate* European Journal Statistics and Probabilit, vol. 2, no. 3, pp. 14-22, 2014.
11. P. Embrechts, C. Klppelberg, and T. Mikosch *Modelling Extremal Events for Insurance and Finance*, Modelling Extremal Events, 2012, doi: 10.1007/978-3-642-33483-2.
12. H. Flugentiusson *Push it to the limit. Testing the usefulness of extreme value theory in electricity markets*, 2012, Lund University Publications.
13. M. Gilli, and F. Këllezi *An application of extreme value theory for measuring financial risk*, Computational Economics, vol. 27, no. 2-3, pp. 207-228, 2006, doi: 10.1007/s10614-006-9025-7.
14. N. G ülpinar, and K. Katata *Modelling oil and gas supply disruption risks using Extreme Value Theory and copula*, Journal of Applied Statistics, vol. 41, no. 1, pp. 2-25, 2014, DOI: 10.1080/02664763.2013.827160
15. O. Hakim *POT approach for estimation of extreme risk measures of EUR/USD returns*, Statistics, Optimization & Information Computing, vol. 6, no. 2, pp. 240-247, 2018, doi: 10.19139/soic.v6i2.395.
16. A. Heymans, and B, Santana *How efficient is the Johannesburg Stock Exchange really?*, South African Journal of Economic and Management Sciences, vol. 21, no. 1, 2018,a1968. https://doi. org/10.4102/sajems. v21i1.1968.
17. O. Jakata and D. Chikobvu *Modelling extreme risk of the South African Financial Index (J580) using the generalised Pareto distribution*, Journal of Economic and Financial Sciences, vol. 12, no. 1, 2019, a407. https://doi.org/ 10.4102/jef.v12i1.407.
18. A. Jenkinson *The Frequency Distribution of the Annual Maximum (or Minimum) of Meteorological Elements*, Quarterly Journal of the Royal Meteorological Society, vol. 81, pp. 158-171, 1955.
19. K. Kiragu, and J. Mungatu *Extreme Values Modelling of Nairobi Securities Exchange Index*, American Journal of Theoretical and Applied Statistics, vol. 5, no. 4, p. 229-234, 2016, doi: 10.11648/j.ajtas.20160504.20.
20. J.K. Kolma, P.N. Mwita, and D.K. Nassiuma *Application of extreme value theory in the estimation of value at risk in Kenyan stock market*, vol.2, pp. 276-284, 2013.
21. G. Magnou *An application of extreme value theory for measuring financial risk in the Uruguayan Pension Fund 1, Department of Legal, Compliance and Risk, University of the Republic, Uruguay.* 2016.
22. E.N.N. Nortey, K. Asare, and F.O. Mettle *Extreme value modelling of Ghana stock exchange index* SpringerPlus. Springer International Publishing, vol. 4, no. 1, pp. 1-17, 2015. doi: 10.1186/s40064-015-1306-y.
23. C.O. Omari, P.N. Mwita, and A.G. Waititu *Using Conditional Extreme Value Theory to Estimate Value-at-Risk for Daily Currency Exchange Rates*, Journal of Mathematical Finance, vol. 07, no. 04, pp. 846-870, 2017. doi: 10.4236/jmf.2017.74045.
24. I.A. Onour *Extreme Risk and Fat-Tails Distribution Model: Empirical Analysis*, Journal of Money, vol. 13, no. 965, 2010.
25. O. Oyewole, O. Adesina, and L. Adekola *On Extreme Value Theory in Modeling Nigeria Marine and Aviation Insurance Class of Business Researchgate.Net.*, 2018
26. J. Pickands *Statistical Inference Using Extreme Order Statistics*, The Annals of Statistics, vol. 3, no. 1, pp. 119-131, 2007. doi: 10.1214/aos/1176343003.
27. F. Ren, and D.E. Giles *Extreme value analysis of daily canadian crude oil prices*, Applied Financial Economics, vol. 20, no. 12, pp. 941-954, 2010. doi: 10.1080/09603101003724323.
28. M. Rydman *Application of the Peaks-Over-Threshold Method on Insurance Data*, Uppsala Universitet U.U.D.M. Project Report, pp. 1-21, 2018.
29. C. Sigauke, R. Makhwiting, and M. Lesaoana *Modelling conditional heteroskedasticity in JSE stock returns using the Generalized Pareto Distribution*, African Review of Economics and Finance, vol. 6, no. 1, pp. 41-55, 2014.
30. J. Cerović, and V. Karadžić *Extreme value theory in emerging markets: Evidence from the Montenegrin stock exchange*, Economic Annals, vol. 60, no. 206, pp. 87-116, 2015. https://doi.org/10.2298/EKA1506087C
31. T. Bali *A generalized extreme value approach to financial risk measurement*, Journal of Money, Credit and Banking, vol. 39, no. 7, 1613-1649, 2007. https://doi.org/10.1111/j.1538-4616.2007.00081.x
32. K. Tolikas *Value-at-risk and extreme value distributions for financial returns*, The Journal of Risk, vol. 10, no. 3, pp. 31-77, 2016. https://doi.org/10.21314/jor.2008.174
33. A.K. Singh, and D.E. Allen *Value at Risk Estimation Using Extreme Value Theory. 19th International Congress on Modelling and Simulation, 500(December), 12-16*. 2011.
34. F. Szubzda, and M. Chlebus *Comparison of Block Maxima and Peaks Over Threshold Value-at-Risk models for market risk in various economic conditions*, Central European Economic Journal, vol, 6, no. 53, pp. 70-85. 2019. https://doi.org/10.2478/ceej-2019-0005
35. O. Vladimir, E.T. Sergey, P.O. Oksana, and P. Gennady *Extreme value theory and peaks over threshold model in the Russian Stock Market*, Journal of Siberian Federal University. Engineering & Technologies vol. 1, no. 5, pp. 111-121, 2016.
36. A. Velasco, and D. Lapuz, *Extreme Value Modelling for Measuring Financial Risk with Application to Selected Philippine Stocks*, Journal of Applied & Computational Mathematics, vol. 07, no. 3, pp. 1-13, 2018, doi: 10.4172/2168-9679.1000404.
37. L. Xiu-min, and L. Fao-chao *Extreme Value Theory: An Empirical Analysis of Equity Risk for Shanghai Stock Market*, International Conference on Service Systems and Service Management, Troyes, pp. 1073-1077, 2006, doi: 10.1109/ICSSSM.2006.320657.
38. South African Financial Index (J520) *Data* (1995-2018), iress expert, viewed 04 February 2018, from https://expert.inetbfa.com.