QUANTITATIVE
&
FINANCIAL MODELLING
(QFM)

ADVANCED STATISTICAL METHODS FOR MODELLING AND FINANCE/
CAPSTONE PROJECT REPORT - 2023

Extreme Values Distributions

EWINSOU Damilola Roméo
NDAO Alioune Badara
SIANOU Ezéckiel Houénafa
TINA Djara Olivier

Professor:
Ravi Prakash Ranjan

Abstract

Extreme Value Distributions (EVD) and Extreme Value Theory (EVT) in general are very important tools for understanding and analyzing extreme value data and for understanding the risks and impacts associated with extreme events. This work aims to show why these distributions are important and how the EVD can be used in the field of modeling and particularly in financial modeling. We have used different exploratory statistical analysis techniques to show that these distributions are heavy tail and have indeed been developed to deal efficiently with extremes events. In order to analyze the losses risk that can occur in the investment in Meta stock, the risk measures we have considered are the Value at Risk (VaR) and the expected shortfall (ES). The Block Maxima (BM) approach which is based on EVD and the Peak Over Threshold (POT) approach based on Generalized Pareto Distribution (GPD used quite a lot in the EVT) have been implemented to find the VaR and ES associated, as done in the article: Modeling Bank of Kigali Stock Risks in Rwanda Stock Exchange Using Extreme Value Distribution (Katu D. E. & Marcel N. (2021)). The results show that the Fréchet distribution corresponds reasonably well to the losses data in the case of Block Maxima.

Keywords

Extremes values distribution; Block Maxima; Peak Over Threshold; Extreme Value Theory; Value at Risk; Expected Shortfall

Contents

1 Introduction

Rare events, also known as extreme events, are occurrences that have a low probability of happening but can have significant consequences. Examples of rare events include natural disasters, technological failures, and large financial losses. Modeling rare events is valuable in many fields because they can have significant impacts on individuals, communities, and organizations. So, tools like Extreme value distributions and Extreme Value Theory have been developed to deal efficiently with extremes events.

Extreme value distributions are a vital tool in the field of finance for modeling and predicting rare events, such as financial crises or extreme market movements. These distributions are characterized by their "heavy tails", which means that probability of occurrence of the most extreme events (both positive and negative) is higher than what is predicted by a reference distribution, such as a normal distribution. These is also called exponentially unbounded tail distributions.

In finance, extreme value distributions are often used to model the potential losses or gains of a financial portfolio, to estimate the probability of rare events such as extreme market crashes, and to measure the impact of these events on the portfolio.

There are three types of extreme value distributions, each with their own specific characteristics and applications. In this report, we will delve into the details of extreme value distributions and their importance in the field of finance.
We will discuss how they are used to model and predict rare events in finance and in addition, use Meta stock prices for different applications.

2 Brief summary on Extreme Values Distributions

2.1 Extreme Value

Extremes values are rare events that lead to significant losses. For example, we have natural disasters, stock market crashes,...
An extreme value is either very small or very large values in a probability distribution. These extreme values are found in the tails of a probability distribution (i.e. the distribution's extremities). The term "extreme value" can mean something slightly different depending on where you read about it. Some authors use the term "extreme value" as another name for the minimum value and/or the maximum value of a function (i.e. the single smallest and/or largest number in the dataset), and others use as a synonym for an outlier. In calculus, the points where you find maximum and minimum values are called extrema, so some authors will refer to these points as "extreme values" as well. However, in most cases, when people talk about extreme values, they're usually talking about values associated with the Extreme Value Theory.

2.2 Extreme Value Distribution (EVD)

In probability and statistic, an extreme value distribution is a type of distribution that is used to model the occurrences of extreme values (i.e., the smallest and largest values) in a dataset. More simply, it is a limiting model for the maximums and minimums of a data set. It models how large (or small) your data will probably get.
The extreme value distributions are often used in engineering and financial applications to model the occurrence of extreme events, such as the failure of a structural component or the maximum loss that a portfolio can incur. These distributions are often used in conjunction with other statistical techniques, such as hypothesis testing and regression analysis, to make predictions and draw conclusions about the data.

We have three types of Extreme Value Distribution: EVD type I, II, and III.
If you generate any number of datasets, take the minimums and maximums from those sets, and generate a new distribution, it will follow one of three model types: no upper or lower limits (EVD I), bounded on the lower end (EVD II), or bounded on the upper end (EVD Type III).

** **EVD Type I: Gumbel Distribution**

The Gumbel distribution is the most common EVD and has two forms: one for the minimum, and one for the maximum, although it is unbounded (not restricted to a range). The probability density function has only one unchanging shape which shifts according to the location parameter μ.
Its pdf is:

$$f(x;\mu,\beta) = \frac{1}{\beta}\exp\left(-\frac{x-\mu}{\beta}\right)\exp\left[-\exp\left(-\frac{x-\mu}{\beta}\right)\right] \tag{1}$$

with $x \in \mathbb{R}$; $\mu \in \mathbb{R}$ and $\beta \in \mathbb{R}_+^*$ the location and the scale parameters respectively.

** **EVD Type II: Fréchet Distribution**

The Fréchet Distribution is used to model maximum values in a dataset . The Fréchet has three parameters: scale parameter β, location parameter μ and shape parameter α. It is bounded (restricted) on the lower side. A wide range of phenomena like flood analysis, horse racing, human lifespans, maximum rainfalls and river discharges in hydrology, Bank stock risks, gold price analysis,... can be modeled with the Fréchet Distribution.

Its pdf is:

$$f(x,\mu,\beta,\alpha) = \frac{\alpha}{\beta}\left(\frac{x-\mu}{\beta}\right)^{-1-\alpha} \exp\left[-\left(\frac{x-\mu}{\beta}\right)^{-\alpha}\right] \quad \text{for } x \geq \mu - \frac{\beta}{\alpha}, \tag{2}$$

where $\alpha \in \mathbb{R}_{+}^{*}$, $\beta \in \mathbb{R}_{+}^{*}$ and $\mu \in \mathbb{R}$.

** **EVD Type III: Weibull Distribution**

The Weibull distribution is used in assessing product reliability to model failure times and life data analysis. It appears practical when we observe data that deal with minima values. It can also fit a wide range of data from many fields, including: biology, finance, economics, engineering sciences, and hydrology. The Weibull is actually a family of distributions that can take on many shapes, depending on what parameters you pick. It is bounded (restricted) on the upper side.
Two versions of the Weibull probability density function (pdf) are in common use: the two parameter pdf and the three parameter pdf. Different authors use different notation, which makes the notation a little confusing if you're looking at different texts.
The pdf of the three parameters general Weibull distribution is given by:

$$f(x;\mu,\beta,\alpha) = \frac{\alpha}{\beta}\left(\frac{x-\mu}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{x-\mu}{\beta}\right)^{\alpha}\right] \quad \text{for } x \leq \mu - \frac{\beta}{\alpha}, \tag{3}$$

where $\alpha \in \mathbb{R}_{+}^{*}$, $\beta \in \mathbb{R}_{+}^{*}$ and $\mu \in \mathbb{R}$. They are respectively the shape (also known as the Weibull slope or the threshold parameter), scale (also called the characteristic life parameter) and location parameter (also called the waiting time parameter or sometimes the shift parameter). The standard Weibull distribution is given by the case where $\mu = 0$ and $\beta = 1$.

Now, the formula of two parameter Weibull is practically identical to the three parameter Weibull, except that μ isn't included:

$$f(x;\alpha,\beta) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{\alpha-1} \mathrm{e}^{-(x/\beta)^{\alpha}} \qquad x \leq -\frac{\beta}{\alpha} \tag{4}$$

The table 1 presents the key statistical properties of each of these three EVDs.

Table 1: Statistical properties

	Gumbel	**Fréchet**	**Weibull**
Mean	$\mu + \beta\gamma$	$\begin{cases} \mu + \beta * \Gamma\left(1 - \frac{1}{\alpha}\right) & \text{if } \alpha > 1 \\ \infty & \text{otherwise} \end{cases}$	$\mu + \beta\Gamma(1 + 1/\alpha)$
Median	$\mu - \beta\ln(\ln 2)$	$\mu + \frac{\beta}{\sqrt[\alpha]{\log(2)}}$	$\mu + \beta(\ln 2)^{1/\alpha}$
Mode	μ	$\mu + \beta\left(\frac{\alpha}{1+\alpha}\right)^{1/\alpha}$	$\mu + \beta\left(\frac{\alpha-1}{\alpha}\right)^{1/\alpha}$
Variance	$\frac{\pi^2}{6}\beta^2$	$\begin{cases} \beta^2\left(\Gamma\left(1-\frac{2}{\alpha}\right) - \left(\Gamma\left(1-\frac{1}{\alpha}\right)\right)^2\right) & \text{if } \alpha > 2 \\ \infty & \text{otherwise} \end{cases}$	Var
MGF	$\Gamma(1-\beta t)e^{\mu t}$	$e^{\mu t}\sum_{r=0}^{\infty}\frac{(\beta t)^r}{r!}\Gamma(1-\frac{r}{\alpha})$	$M_X(t)$

$Var = \beta^2\left[\Gamma\left(1+\frac{2}{\alpha}\right) - \left(\Gamma\left(1+\frac{1}{\alpha}\right)\right)^2\right]$ and $M_X(t) = \sum_{n=0}^{\infty}\frac{t^n\beta^n}{n!}\Gamma\left(1+\frac{n}{\alpha}\right)$ for $\alpha \geq 1$ in the table, are for two parameter Weibull. γ is the Euler–Mascheroni constant and is given by:

$$\gamma = \lim_{n\to\infty}\left(-\log n + \sum_{k=1}^{n}\frac{1}{k}\right) = \int_{1}^{\infty}\left(-\frac{1}{x} + \frac{1}{\lfloor x \rfloor}\right) dx. \tag{5}$$

$\lfloor x \rfloor$ represents the floor function.

2.3 Generalized Extreme Value (GEV) Distribution

The generalized extreme value (GEV) distribution is a family of continuous probability distributions developed within extreme value theory to combine the Gumbel, Fréchet and Weibull families also known as type I, II and III extreme value distributions. It is used to represent extreme value phenomena (minimum or maximum). Note that a limit distribution needs to exist, which requires regularity conditions on the tail of the distribution. Despite this, the GEV distribution is often used as an approximation to model the extrema of long (finite) sequences of random variables.
Its pdf is defined as:

$$f(x;\mu,\beta,\alpha) = \frac{1}{\beta}\left[1+\alpha\left(\frac{x-\mu}{\sigma}\right)\right]^{(-1/\alpha)-1}\exp\left\{-\left[1+\alpha\left(\frac{x-\mu}{\beta}\right)\right]^{-1/\alpha}\right\} \tag{6}$$

$$f(x;\mu,\beta,0) = \frac{1}{\beta}\exp\left(-\frac{x-\mu}{\beta}\right)\exp\left[-\exp\left(-\frac{x-\mu}{\beta}\right)\right] \tag{7}$$

$\alpha \in \mathbb{R}$ is also called the extreme value index. The values $\alpha = 0$ (intermediate or light tail), $\alpha > 0$ (fat tail) and $\alpha < 0$ (thin tail) correspond, respectively, to the Gumbel (1), Fréchet (2) and Weibull (3 and 4) distributions.

2.4 Domain of attraction

When a distribution converges to one of three EVD, it is said to belong to its domain of attraction. That implies by example that the distributions belonging to the Gumbel domain of attraction have the light tails (Moderately heavy tails), while the distributions belonging to domain of attraction of Weibull have an upper bounded support (S. El Adlouni et al, 2007). We can determine on the basis of these characterizations, the domain of attraction of the majority of the usual distributions. For example, the table 2 presents the domains of attraction of each of the three EVDs.

Table 2: EVDs's Domain of attraction

Gumbel	**Fréchet**	**Weibull**
Normal	Cauchy	Uniform
Exponential	Pareto	Beta
Log-normal	Chi-squared	
Gamma	Student	

3 Exploratory analysis in R

In this section, we first randomly generate data of size $n = 100000$ folllowing each EVD. Subsequently, the pdf of these distributions are plotted and also some key descriptive statistics are calculated. We used these generated data to construct histograms, quartile plots (Boxplot). The last part is devoted to the influence of each parameter on the probability densities.

3.1 Generalized extreme value densities and some key statistics

The *revd* function of the *evd* package of R allowed us to generate data following the three types of EVD. Figure 1 shows us the probability density functions of each of the EVDs. We plotted them using the same location parameters μ, dispersion parameters β and varying the shape parameter α, which normally defines each type of EVD. We can clearly see that the Gumbel distribution has light tails while the Weibull and Fréchet distributions have respectively a thin and a fat tail and moreover have respectively an upper bounded and an lower bounded support.

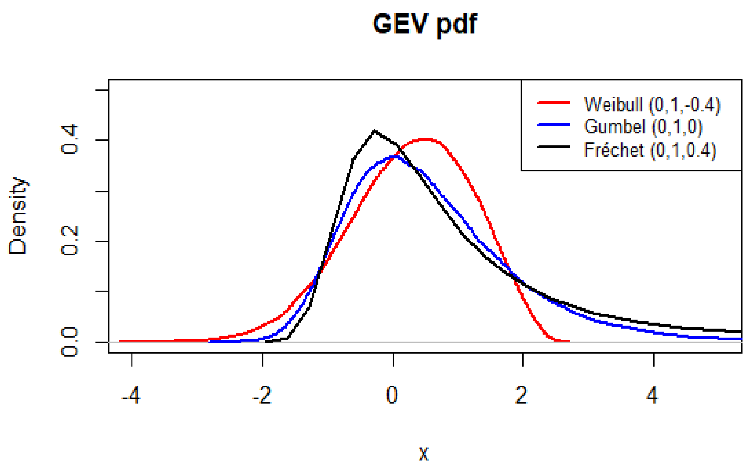

Figure 1: GEV densities with $\mu = 0$ and $\beta = 1$

Using the data generated and the function *summary* in R, we have computed some key statistics such as the minimum, maximum, mean, median, first and third quartiles of each EVD. These statistics are presented on the figure 2.

```
> summary(data_gumbel)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-2.4960 -0.3311  0.3669  0.5777  1.2463 12.2004
> summary(data_weibull)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-3.9657 -0.3413  0.3480  0.2889  0.9881  2.4694
> summary(data_frechet)
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
 -1.5709  -0.3067   0.3936   1.2022   1.6100 169.3623
```

Figure 2: some key statistics

From these statistics (especially the min and max), we can start by noticing the characteristics of the tails of the three distributions. These characteristics will be better explained by the histograms and boxplots that we will plot in the following.

3.2 Histograms and Boxplots

As we know, histograms are visual representations of a data set that show the frequency of occurrence of each value in the data set.
The histograms of our data are represented in figure 3. So, we can easily see which values are most common and which values are least common in the three cases. In addition, we can clearly see that each histogram shows the shape of each of the densities as in the figure1.

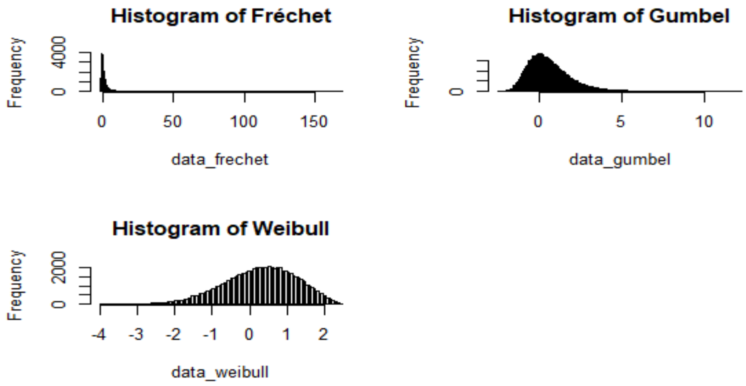

Figure 3: Histograms of the three EVD

The figure 4 displays the boxplots of the three families of GEV distribution. Boxplots are a useful way to visualise the range (like the case of histograms) and other characteristics of the data. To begin with, each data of EVD are sorted . Then, four equal sized groups are made from the ordered values. That is, 25% of all values are placed in each group. The lines dividing the groups are called quartiles, and the groups are referred to as quartile groups. Usually we label these groups 1 to 4 starting at the bottom.
Considering each middle box colored red, it represents half of values. The median (middle quartile) is represented by the line that divides these box into two parts. Half of the total values are greater than the median and the other half is less. The upper limit of each middle box is called upper quartile and 25% of the values are above this limit. Similarly, the lower boundary is called the lower quartile and 25% of scores fall below it. Values outside the box are represented by upper and lower whiskers.

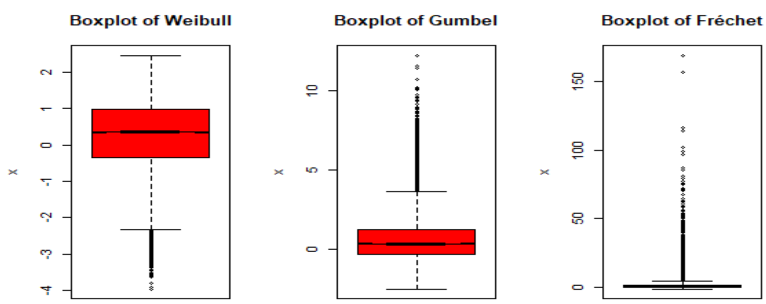

Figure 4: Boxplots of the three EVD

Now, we can also see that the figure 4 shows that all the three boxplots indicate some number of values above upper or lower whisker which are referred to as the potential outliers. These values are usually values that if removed may significantly distort the accuracy of some parameter estimates of the empirical models about the dataset.

We note a strong display of characteristics that enable to distinguish each type of distribution family. The Weibull boxplot is the first one shown. It shows a sign of skewness on the left from the lower whisker range, with outliers from the lower whisker range. The second boxplot is for the Gumbel distribution. The last one is the boxplot of the obtained Fréchet distribution. It is asymmetric or left bounded. The long fat tail can be observed from the upper whisker spread and the marked outliers.

3.3 Identify of the heavy tail using QQ-plot

We consider only the Fréchet sample in the rest of this section. We generally talk about a heavy tail distribution when the tail is not exponentially bounded. Exponential QQ-plot in the figure 5 shows that the data are not straight line. The points on the upper side of the Q-Q diagram are much farther from the straight line than the other points, this may indicate a right heavy tail. So, we can say that the Fréchet distribution is not exponentially bounded and has a heavy tail.

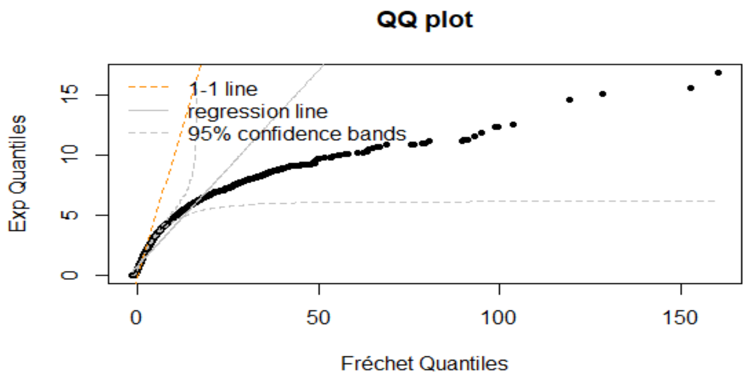

Figure 5: QQ-plot of Fréchet data with exponential distribution

3.4 Influence of each parameter on the EVD densities

We have the figures 6, 7 and 8. We keep two parameters constant and we vary the third one. The

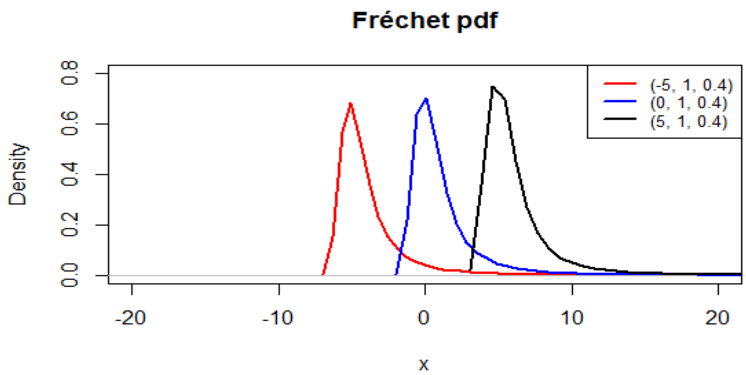

Figure 6: Influence of location parameter μ

first figure (figure 6) shows different positions of the distribution along the x-axis. This confirms that the location parameter μ determines the position of the distribution along the x-axis.
On the figure 7, we see that the larger the scale parameter β is, the more dispersed the values are. The α positive values characterize the Fréchet type. One can also clearly notice on the figure 8 that the larger α is, the more the distribution is lower bounded.

Finally, we can notice that the parameters are too influential on the probability density. So, we can deduce that it is much more logical to use a very efficient method to estimate the parameters in order to allow a better representation of the phenomena with the EVD.

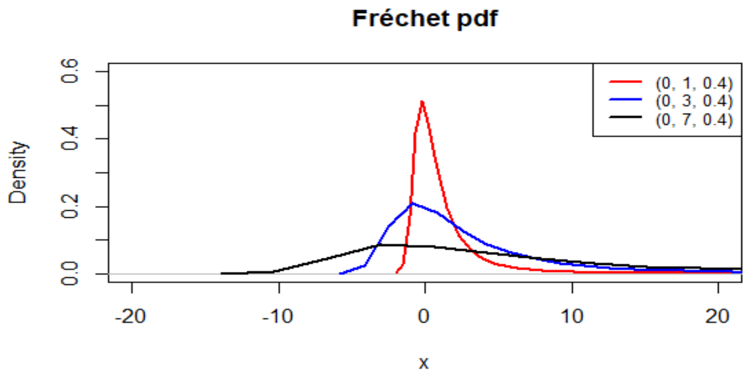

Figure 7: Influence of scale parameter β

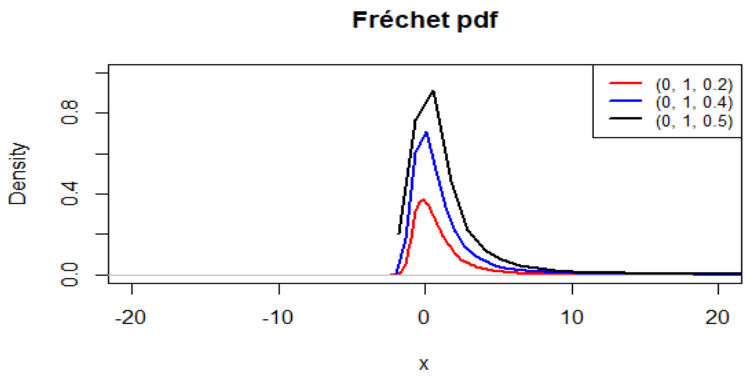

Figure 8: Influence of shape parameter α

4 Summary of the article : Modeling Bank of Kigali Stock Risks in Rwanda Stock Exchange Using Extreme Value Distribution

This study has been done to model the Bank of Kigali's (BK) stock risks in Rwanda stock exchange using Extreme Value Distribution. Two major approaches are used.

To model Bank of Kigali stock risks, the Generalised Extreme Value Distribution (GEVD), precisely the Block Maxima is implemented.

To examine its associated exceedances, the Generalised Pareto Distribution (GPD) is also implemented. Risk measures considered are the Value at Risk (VaR) and the Expected Shortfalls (ES). Findings reveal that the Weibull distribution fits reasonably well the distribution of the BK stock returns and GPD the exceedances.

Also, the risk measures such as Value at Risk and Expected shortfall were computed with high level (99.5%) quantiles to serve as a guide to investors to make a decision as to whether to invest in Bank of Kigali's stock or not. The findings show that GPD fits the tail of the data well.

4.1 The problem statement

The last decades have been exposed to instabilities in financial markets worldwide and for that matter, Rwanda. There comes the need to require tools to actually assess the probability of rare financial extremes. Different periods of financial markets have been affected by several crises such as the Stock Market Crash (1987), the Crisis (1997-1998) and Global Financial Crisis (2007-2008). The Global Financial Crisis in 2008, for example, led to an unforgettable mark of unemployment and debt.

Then the question one will ask is what the right thing being done in finance? What was the cause for this? What went wrong?

It is for these reasons that prompted looking for appropriate ways to deal with rare events that have a big effect on financial market. So, it comes naturally to try to anticipate and predict these kinds of events,

Thus, the research questions to this essay are articulate around 4 points:

1) Which particular case of the GEVD could fit reasonably well the BK daily stock loses?

2) Which size of exceedances are associated with the BK daily losses when GPD is used?

3) What is BK stocks VaR size?

4) To what extent can the de-cluster the exceedances be informative of the behaviour of VaR and ES?

4.2 Relevance of the problem

The problem being addressed above in this paper is the need to develop tools to assess the probability of rare financial extremes and to anticipate and predict these events for the Bank of Kigali. This is relevant because financial crises and other rare events can have significant impacts on financial markets, including unemployment and debt, and it is important to be able to anticipate and manage these risks.

The research questions focus on identifying the appropriate statistical models and methods for analyzing financial data and understanding the behavior of financial risks. This is relevant because accurate risk assessment is critical for informed decision-making in finance, and the results of this research could help to improve the ability of individuals and organizations to anticipate and manage financial risks. It could help investors to decid if whether or not they should invest in the Bank of Kigali.

Overall, the problem is relevant to anyone interested in understanding and managing financial risks, particularly in the context of past financial crises and the potential for future events to have significant impacts on financial markets.

4.3 Approach to solve the problem

4.3.1 Data description

In the paper, they used the adjusted closing daily stock market prices for Bank of Kigali investment in the Rwanda Stock Exchange. The sample consists of 1226 daily closing prices of RSE starting from 3rd January 2012 to 16 November 2017. 2012-2017 closing prices because those were the data available as at the time the researcher was conducting the study. There are 229 missing data or 18.7% of the total data collected. They deal with missing data using na.interpolation method.

4.3.2 Research design

Quantitative research was used to describe procedures that are followed to model Bank of Kigali stock risks using the GEV distribution. To identify the threshold, the GPD is used. The following steps are performed:

1) Observation of stock prices from RSE in terms their daily returns;

2) Cleaning of data and imputation of missing stock prices using appropriate smoothing method, precisely, na.interpolation (spline);

3) Data Processing using R software, version 3.4.4. With the help of R software, dataset is processed to:

3.a) Measure bank of Kigali stock risks using GEVD specifically using Block Maxima approach;

3.b) Determine the exceedances by fitting the daily loses using GPD precisely Peak Over Threshold;

3.c) Assess the measures of risk (VaR and ES).

3.d) De-cluster the exceedances occurring in the stock's returns.

4.3.3 Methodology

Two approaches are used, mainly the Block Maxima method and the Peak over Threshold.

4.3.3.1 Block Maxima

Block Maxima theory deals with the convergence of maxima, that is, the limits law for the maxima. To actually illustrate this method (figure 9), assume that $r_t, t = 1, \ldots, n$, is a sequence of *iid* observations which have a distribution function $H(x) = P_r(r_t \leq x)$ and let the sample maximum be denoted by $M_n = max\{r_1, \ldots, r_n\}$ where $n \geq 2$. GEVD represented by $H(x)$ describes the limiting distribution of suitably normalised maxima.

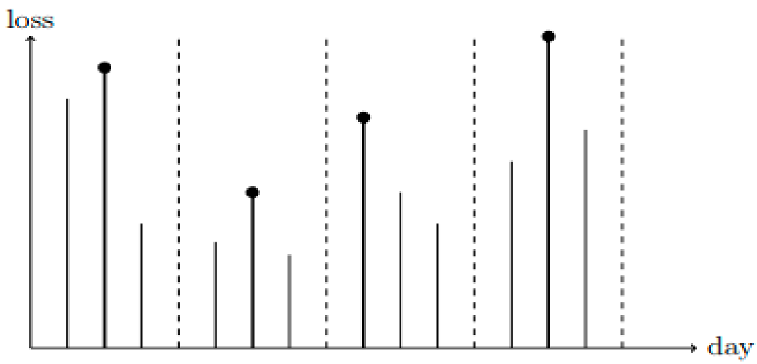

Figure 9: Block maxima method

Block Maxima theory deals with the convergence of maxima, that is, the limits law for the maxima. To actually illustrate this method (figure 9), assume that $r_t, t = 1, \ldots, n$, is a sequence of *iid* observations which have a distribution function $H(x) = P_r(r_t \leq x)$ and let the sample maximum be denoted by $M_n = max\{r_1, \ldots, r_n\}$ where $n \geq 2$.

4.3.3.2 Peak Over Threshold

Considering Peak over Threshold, the focus is on set of realized values exceeding a certain threshold. One technique of removing extremes from a sample of observations $r_t, t = 1, 2, \ldots, n$ with distribution function $F(x) = Pr(r_t \leq x)$ is to take exceedances over a deterministic high threshold μ. Exceedances over threshold occur when $r_t > \mu$ for any t in $t = 1, \ldots, n$.

An excess over μ is defined by $x = r_i - u$. Given a threshold μ, the probability distribution of excess values of r_t over threshold is defined as $F_u(x) = Pr\{r - \mu \leq x | r > \mu\}$. This represents the probability that the value of r exceeds the threshold μ by at most an amount x given that r exceeds the threshold μ.

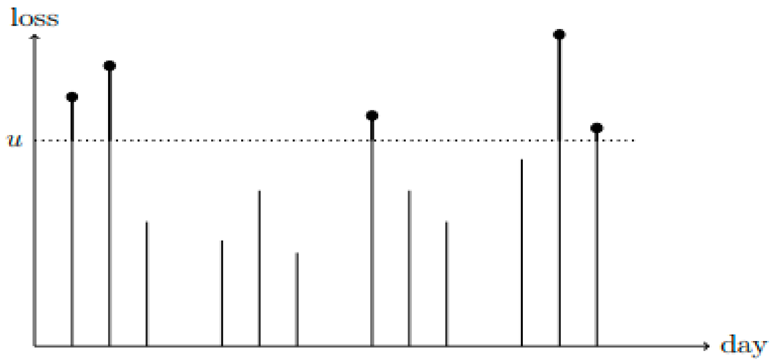

Figure 10: Peak Over Threshold method

Measuring risks use VaR and ES. The mathematical tool to deal with risk measures are the Value at Risk (VaR) and Expected Shortfall (ES). VaR is a tool for risks assessment. Its utilization in Banks reflects their fear of a liquidity emergency, where a low-probability catastrophic occurrence is made. Assume an arbitrary variable X, with continuous distribution function F, models return of an asset over a certain time horizon. The VaR_p is the p^{th} quantile of the distribution F defined as $VaR_p = F^{-1}(1-p)$, where F^{-1} is the pth quantile function. The ES is defined as the expected of the return that exceeds VaR_p: $ES = E(X|X > VaR_p)$.

4.3.4 Declustering

Declustering systems are strategies that look to distinguish occasions in a data record and filter out extreme values from them.

4.4 Conclusion and Business implications

This study showed how Extreme Value Distributions and Extreme Value Theory in general, can be a useful tool to model Bank of Kigali's stock risks. Also, it was revealed how it can be used to model tail-related measures such as Value at Risk, expected shortfall and applying it to Bank of Kigali stock prices. The study exhibited heavy-tail behaviour in returns of the Bank's data and provided a better fit to the tails of the distribution of results.

The study recommends Rwandan investors, especially risk-averse ones, who would prefer not to take up risky investments in light of the propensity of losing their investments to go on and invest.

This comes forth from the result of risk measures (VaR and ES) as obtained in the study. This is because VaR is used to quantify the risk of a portfolio over a specific time frame.

VaR is a very useful tool for investment as it provides a means of assessing how risk exposure managers are taking on to achieve their portfolio return. This work also can serve as a guide to insurance companies and financial economists in modelling risks associated with them.

5 Application. Modelling Meta stock losses: VaR and ES analysis

5.1 META adjuced closing price and return

The data we have used, are for the Meta(Facebook) group adjusted closing prices , over 6 years (2017-2022) and are downloaded from yahoo finance. We have no missing values in the dataset. We have computed the return using the formala:

$$return = \frac{Price_t - Price_{t-1}}{Price_{t-1}} \tag{8}$$

We can see the adjuced closing prices and return on figure 11.

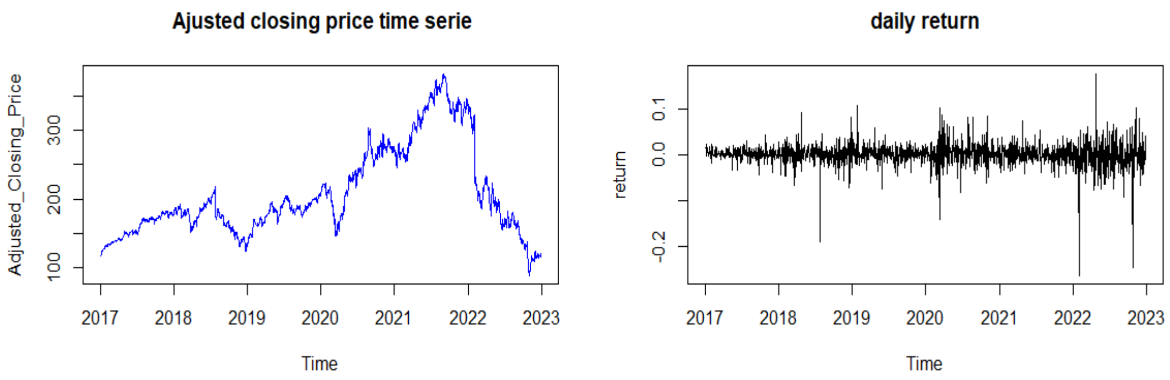

Figure 11: Adjuced closing prices and return

From the daily return plot on the above figure, we can notify two tails of the return distribution. The left tail and the right tail. The left tail corresponds to the maximum losses. The methodology applying to right tails, in the left tail case we change the sign of the returns so that positive values correspond to losses.

5.2 Block Maxima and GEV fitting

We used Block maxima methods to identify the extreme value corresponding to the maximum losses. With a total of 1509 data, we divide the data into 76 blocks. Each block contains 20 values. We can see the maximum of each block on the figure 12.

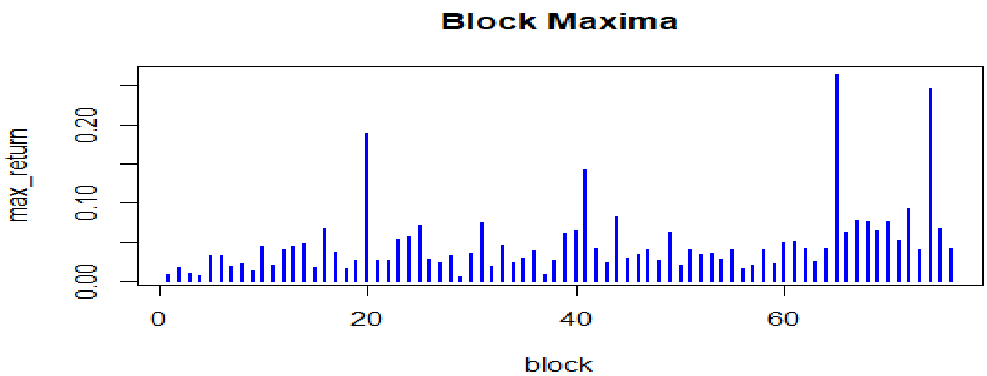

Figure 12: Block Maxima depicting maximum Meta losses

Using the function *fevd* of R library *extRemes*, we fitted an EVD to Block Maxima data. This function use Maximum Likelihood Estimator (MLE) method to estimate the parameters of the

Table 3: EVD Estimation

Parameters	location(μ)	scale(β)	shape(α)
Estimates	0.02932643	0.01824550	0.30072201
Standard error	0.002360138	0.001954631	0.094424579

corresponding EVD. The table 3 presents the three parameters.
Considering the shape parameter α from the table 3, since $\alpha > 0$, it follows that the GEV is of Fréchet type (type II). The figure 13 depicts diagnostic plot for fitting GEV to the block maxima data. It can be seen that the top two plots do not deviate much from the straight line. Empirical observations (marked with a circle) are close to the diagonal, representing a perfect fit in both the PP and QQ plot. For the density plot, it can be observed that the curve does not exactly match up with the histogram very well. However, it can be seen that it is right-skewed (positive skewness) with a right fat tail. The return level presents a graph of the empirical estimates of the return level plotted against the estimated return levels from the fitted model. It can also be observed from the plot that all the points lie approximately on the line.

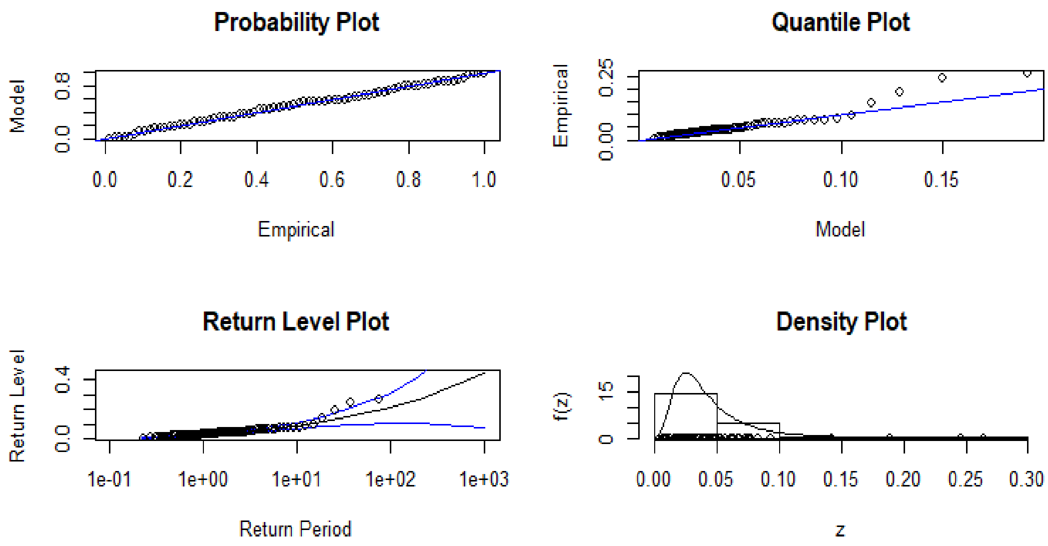

Figure 13: Diagnostic plot depicting GEV fitting to the block maxima data

5.3 VaR and ES

Using the block maxima results obtained, we can compute the VaR using the following equation (equation 9):

$$\mathrm{VaR}_p = \begin{cases} \mu - \frac{\beta}{\alpha}\left(1 - (-n\ln(1-p))^{-\alpha}\right) & \text{for} \quad \alpha \neq 0 \\ \mu - \beta\ln\left(-n\ln(1-p)\right) & \text{otherwise} \end{cases} \tag{9}$$

where $n = 76$, the block number.
As remember, ES is defined as the expected of the losses that exceed VaR_p: $ES = E(loss|loss > VaR_p)$. To determine it, we think to consider the Peak Over Threshold (POT) approach based on GPD (ϵ, δ) and used in the article we have summarized:

$$ES_p = E(loss|loss > VaR_p) = \frac{VaR_p}{1-\epsilon} + \frac{\delta - \epsilon s}{1-\epsilon} \tag{10}$$

$$VaR_p = s + \frac{\delta}{\epsilon}\left[\left(\frac{1-p}{1-F(s)}\right)^{-\epsilon} - 1\right] \tag{11}$$

where F is the cdf of GPD and s is the threshold (Klara, A. 2020). We then require an estimate of $F(s)$ which according to Davison, A. C.& Richard L. S.(1990), can be approximated by the empirical distribution function $\bar{F}(s) = \frac{n-N_s}{n}$ where n is the total number of observations and N_s the number of observations above the threshold. In this way the tail estimator can be written as

$$VaR_p = s + \frac{\delta}{\epsilon}\Big[\Big(\frac{n}{N_s}(1-p)\Big)^{-\epsilon} - 1\Big] \tag{12}$$

To determine the threshold, we used the quantile analysis method and therefore we consider that the values above this threshold are exceptional. So, we have chosen the 95_{th} quantile of the loss (negative return changed to positive) distribution as the threshold used. The R function $findThreshold$ of library $fExtremes$ is used to find this threshold and the figure 14 shows all these values above the threshold. The table 4 presents the parameters of the GPD fitting to the POT data.

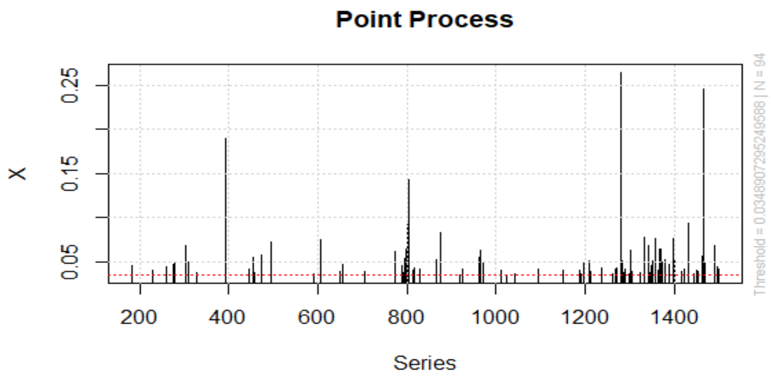

Figure 14: POT depicting maximum Meta losses

Table 4: GPD Estimation

Parameters	**scale(δ)**	shape(ϵ)
Estimates	0.0110902	0.5173850
Standard error	0.002197907	0.176031128

Using the equations 9, 5.3 and 11, we computed the VaR and the ES which are presented on table 5.

Table 5: VaR and ES

Confidence level	95%	97%	99%
Block Maxima VaR	0.05999799	0.04674735	0.02460466
POT VaR	0.02215508	0.02210514	0.02205674
ES	0.02714837	0.02704489	0.02694461

Using POT approach, we have obtained $VaR = 2.205674\%$ with 99% confidence level. It means that the next days, maximum potential loss will be 2.205674% at the confidence level of 99%. For example, suppose an investor invests 1 million dhs, then the maximum loss that may be obtained in the next days is 22056.74 dhs. The ES implies that the corresponding loss in the situation where loss exceeds 2.205674% is 2.694461% at 99%. So for the 1 million dhs, the loss one should expect is equal to 26944.61 dhs.

6 Conclusion

This work had a double objective: to present extreme value distributions and show how to use them to model financial rares events. This statistical tool is particularly suitable for studying booms and crashes in financial markets. We have shown how to get the distribution of extreme values for a given data-set using Block Maxima and Peak Over Threshold.

One of common applications of EVT is risk management, for this reason we provided two different methods to get the risk measures.

In the application part, we showed that the expected loss of a portfolio or financial asset under stressful market conditions is around 2.7% with a confidence level of 99%. This can help decision-making for investors.

Despite their usefulness, there are also limitations and challenges to the use of EVDs in finance. For example, EVDs rely on the assumption that the data follows a particular distribution, which may not always be the case in reality. Additionally, EVDs may not always accurately capture the complexity of real-world financial systems and the interactions between different market variables. As a result, it is important to carefully consider the limitations and assumptions of EVDs when using them to model and predict rare events in finance.

7 references

1- Antoine B. & Ashkan N. (2000). *La théorie des extrêmes et la gestion des risques de marché*
2- Christian Y. R. (2006). *Théorie des Valeurs Extrêmes*,ISFA - Université Lyon 1)
3- Davison, Anthony C. & Richard L. Smith. (1990) *Models for exceedances over high thresholds. Journal of the Royal Statistical Society. Series B (Methodological)*
4- Edem, K. D., & Ndengo, M. (2021). *Modeling Bank of Kigali Stock Risks in Rwanda Stock Exchange Using Extreme Value Distribution. Journal of Financial Risk Management, 10, 225-240. https://doi.org/10.4236/jfrm.2021.103013*
5- Erik B.& Claudia K. *Extreme Value Theory in Finance*
6- Ferreira, A., De Haan, L. & al. (2015). *On the Block Maxima Method in Extreme Value theory: PWM Estimators*, The Annals of Statistics, 43, 276-298
7- Javier J. G. V. & Alberto R. A. (2014), *Use of Gumbel and Weibull functions to model extreme values of diameter distributions in forest stands*, Annals of Forest Science 71:741–750, DOI 10.1007/s13595-014-0369
8- Klara A. (2020).*A comparative study of VaR and ES using extreme value theory*, Department of Economics Lund University, Sweden
9- Longin, F. (2005). *The Choice of the Distribution of Asset Returns: How Extreme Value Theory Can Help?* Journal of Banking and Finance, 29, 1017-1035
10- Manfred G.& Evis K. (2006). *An Application of Extreme Value Theory for Measuring Financial Risk*, (DOI: 10.1007/s10614-006-9025-7)
11- Nortey, E. N., Asare, K., & Mettle, F. O. (2015). *Extreme Value Modelling of Ghana Stock Exchange Index*, SpringerPlus, 4, Article No. 696
12- Pratiwi N., Iswahyudi C. & Safitri R. I. (2019). *Generalized extreme value distribution for value at risk analysis on gold price*, J. Phys.: Conf. Ser. 1217 012090
13- Stéphane G. (2008). *Introduction à la statistique des valeurs extrêmes*, INRIA Rhône-Alpes, projet Mistis http ://mistis.inrialpes.fr/girard
14- Stephanie G. *"Extreme Value Distribution & the Extreme Value Theory"* From StatisticsHowTo.com: Elementary Statistics for the rest of us! https://www.statisticshowto.com/extreme-value-distribution/
15- https://en.wikipedia.org/wiki/Generalized extreme value distribution

8 Appendix

```
library(evd)
library(dplyr)
library(fitdistrplus)
library(extRemes)
library(evir)
library(fExtremes)
library(ismev)
#GEV pdf
data_gumbel <-revd(100000,loc=0,scale=1,shape=0)
data_weibull <-revd(100000,loc=0,scale=1,shape=-0.4)
data_frechet <-revd(100000,loc=0,scale=1,shape=0.4)
plot(density(data_weibull), xlim=c(-4,5), ylim=c(0,0.5), main='GEV pdf',
xlab = 'x', col="red", lwd= 2, type='l')
lines(density(revd(100000,loc=0,scale=1,shape=0)), col="blue", lwd= 2, type='l')
lines(density(data_frechet), col = "black", lwd = 2, type='l')
legend("topright", legend=c("Weibull (0,1,-0.4)","Gumbel (0,1,0)",
"Fréchet (0,1,0.4)"),col=c("red", "blue","black"), lwd = 2, cex=0.8)
#Key statistics computing , one case
summary(data_gumbel)
#Histograms , one case
par(mfrow = c(2,2))
hist(data_frechet, breaks = "FD", main = "Histogram of Fréchet")
#boxplot , one case
par(mfrow = c(1,3))
boxplot(data_weibull, col = "red", ylab = "x", main = "Boxplot of Weibull",
notch = TRUE, varwidth = TRUE)
#qqplot
qqplot(data_frechet,rexp(length(data_frechet),1/mean(data_frechet)),
xlab = 'Fréchet
Quantiles',ylab = 'Exp Quantiles', main="QQ plot")
#influence of each parameter, one case
par(mfrow = c(1,1))
frechet1 <-revd(100000,loc=-5,scale=1,shape=0.4)
frechet2 <-revd(100000,loc=0,scale=1,shape=0.4)
frechet3 <-revd(100000,loc=5,scale=1,shape=0.4)
plot(density(frechet1), col="red", xlim=c(-20,20), ylim=c(0,0.8),
main = "Fréchet pdf", xlab = "x", lwd=2)
lines(density(frechet2), col="blue",lwd=2)
lines(density(frechet3), col="black", lwd=2)
legend("topright", legend=c("(-5, 1, 0.4)", "(0, 1, 0.4)", "(5, 1, 0.4)"),
       col=c("red", "blue","black"), lty = 1, lwd=2, cex=0.8)

#data importation
facebook_data <- read.csv("C:/Users/LENOVO L470/Desktop/QFM/Advanced Stat/
presentation/META.csv")
#adjuced price  plot
par(mfrow = c(1,1))
Adjusted_Closing_Price = facebook_data$Adj.Close
plot.ts(Adjusted_Closing_Price,xaxp=c(2017,2022,5),col="blue",main=" Ajusted
closing price
time serie")
axis(side=1,at=c(1,252,252*2,252*3,252*4,252*5,252*6),labels=c("2017",
```

```
"2018","2019","2020",
"2021","2022","2023"))
#daily losses and plots
return =NULL
n = length(Adjusted_Closing_Price)
x = seq(1,n, by = 1)
for (i in (2:n) ){
return[i-1] = (Adjusted_Closing_Price[i] - Adjusted_Closing_Price[i-1])/
Adjusted_Closing_
Price[i-1]}
plot.ts(return,xaxp=c(2017,2022,5),main="daily return")
axis(side=1,at=c(1,252,252*2,252*3,252*4,252*5,252*6),labels=c("2017",
"2018","2019","2020","2021","2022","2023"))
#block_maxima
max_return=blockMaxima(-return, block = 20, doplot = FALSE)
plot(max_return,type = 'h',lwd='2',col='blue',xlab ='block',
main = "Block Maxima")
#fitting and GEV parameters
fit = fevd(max_return,type="GEV")
summary(fit)
# Diagnostic of GEV
gev.diag(gev.fit(max_return))
#VaR GEV
loc = 0.02932643;scale=0.01824550
shape=0.30072201;n=76;alpha=0.01
var = loc - (scale/shape)*(1 - (-n*log(1-alpha))^shape)
#GPD
return=-return
findThreshold(return, n = floor(0.05*length(as.vector(return))), doplot = FALSE)
pot=pointProcess(return, u = quantile(return, 0.95), doplot = TRUE)
fitgpd=gpd.fit(return,0.03893246)
gpd.diag(fitgpd)
# VAR et ES GPD
seuil=0.03893246;epsilon=0.5173850;delta=0.0110902
n= length(return)
var = seuil + (delta/epsilon)*(((n/Ns)*(1-alpha))^(-epsilon)-1)
ES= var/(1-epsilon)+ (delta-epsilon*seuil)/(1-epsilon)
```