Research Article

Extreme Value of Intraday Returns

Aranit Muja

Faculty of Economics,
Tirana University

Doi: 10.2478/ajis-2018-0069

Abstract

The aim of this research paper is to study the properties of intraday returns, in a time range from one to fifteen minutes. In order to perform this analysis, we consider four sets of historical intraday returns for FTSE-MIB index. The first series consist of intraday returns with one-minute frequency, represented in log scale, which includes the period from 01.04.2011 till 30.09.2011. The consideration period for the other series does not vary, but the frequencies which we calculate the returns with, do. In detail, we took in consideration returns generated in 1, 5, 10 and 15 minutes. First, the study analyses the distribution of intraday returns by employing both graphical methods and moments calculation on different time scales. Secondly, the study analyses the returns maximum distribution on different time scales, checking the type GEV (Generalized Extreme value) returns distribution goodness of fit. The GEV parameters estimation was made by maximum likelihood using EVIM1 toolbox in Matlab.

Keywords: *intraday returns, block maxima, extreme value theory, GEV, stylized facts*

1. Introduction

The assessment and modelling of extreme event is recently gaining momentum in actuarial and financial disciplines. In many cases, we are more interested in the probability of extreme events (events distribution minimum or maximum) rather than normal ones.

The extreme event in financial applications may bring down the company, an equity price or a portfolio and, thus, give rise to the claim amount by the reinsurance company. It is helpful to use the principles of extreme value theory to obtain a good estimate of these events. The theory deals with the study of the asymptotic distribution of extreme events, i.e. events that occur with a low frequency and are relatively large compared to most of the observations in a given sample. The statistical methods derived from this theory have had an increasing use in finance, especially in the risk assessment. The EVT application in finance, estimates, particularly, the tail shape of a return distributions and uses, only for analysis, past performance time-series“extreme” data.

The rest of the paper is organized in the following fashion. The first section describes the extreme value theory. The goal, instead, is to find a limit theorem that gives a non-trivial result on the distribution of the maximum for sufficiently large samples. We have followed the analogy with the central limit theorem, in process.

In literature, the most important result in the extreme value theory is the “Three types theorem”(Fisher, Tippett 1928; Gnedenko 1943). The theory deals with the convergence of maxima.

It was considered that, if the maximum normalized sample admits a non-degenerate limit

[1]*EVIM is a free software package for extreme value analysis in Matlab*

distribution G (x), then G (x) is necessarily represented by one of the three distribution functions: Gumbel (1935), Fréchet (1927) o Weibull (1951). The three types, for the possible landmark law are recognized as the extreme value distribution and allow a unified representation named GEV with δ parameter known as parameterization of Jenkinson-Von Mises. (Jenkinson, 1955, 158-171).

The second section describes the dataset. The empirical study is applied to a selection of four time series of intraday losses related to the FTSE MIB index. The choice of period was driven by the desire to test the performance of the model for a particularly turbulent as the sub-prime mortgage crisis and the sovereign debt of some European countries.

As shown by (Malevergne, Y., Sornette , D, 2006, 53) in order to remove the effect of overnight price jumps, the intraday returns will be determined separately for each day before they are included in the dataset. This section also shows the results of the empirical evidence on the confirmation of the stylized facts for the intraday returns.

In the last section, the classical model of extreme value based on the distribution of the Generalized Extreme Values (GEV) will be explained in detail, and interesting comments on this distribution will be shown.

2. The Limiting Distribution and Their Properties

As (Coles, 2001, 45) and (Beirlant, J., Goegebeur, Y. , Segers .J, & Teugles, 2004, 46) have shown, let *X* be a random variable that describes the return. X_1 , X_2 ... X_n are a sequence of independent realizations of *X*, where X_1 , X_2 ... X_n are independent and identically distributed (iid) ~*X.*

The distribution: $F_X(x) = P(X \leq x) = F_{xi}(x)$ for i=1…n.

The starting point for the extreme value analysis is the behavior of $M_n = \max(X_1, \dots X_n)$ called maximum order statistics. We proceed to derive the distribution of random variable M_n.

$F_{M_n}(x) = P(M_n \leq x) = P\{(X_1 \leq x) \cap (X_2 \leq x) \cap \dots \cap (X_n \leq x)\} = P(X_1 \leq x)P(X_2 \leq x) \dots P(X_n \leq x) = F_X(x)F_X(x) \dots F_X(x) = [F_x(x)]^n$

Whereby F_{M_n}(x) $= P(M_n \leq x) = [F_x(x)]^n$, x∈ $\mathbb{R}$, $n \in \mathbb{N}$. (1.1)

However, in practice, the distribution obtained is not very useful and does not have immediate application given that the underlined distribution F_x is unknown and therefore F_{M_n} (x) will be unknown.

Objective: to find a limit theorem that gives a non-trivial result on the distribution of the maximum normalized M_n for a sufficiently large (n→+∞) sample. One possibility consists in using statistical techniques to estimate F on observed data, and afterwards replace the result obtained in (1.1). But a small bias in the estimation of F can result in a huge discrepancy for $[F_X(x)]^n$. An alternative approach is based on similarities with the Central Limit Theorem, which provide evidence that the limiting distribution of the sum of random variables with common df F is of Normal type.

In our case, instead of the sum, we have the maxima, and the limit distribution is one of the three types, Gumbel, Fréchet or Weibull.

2.1 Limiting Distributions of Maxima

We apply the same procedure of Central Limit Theorem (CLT) to the random variable

$M_n = \max(X_1, \dots X_n)$

And look for some normalization factors a_n, b_n , such that

$P(\frac{M_n - b_n}{a_n} \leq x) \xrightarrow[n \to +\infty]{} G(x)$

The solution is given from Fisher-Tippett, Gnedenko theorem.

Theorem (Fisher-Tippett, Gnedenko)

Suppose $X_1, \dots, X_n$ are iid with df F and a_n, b_n are constants so that for some non-degenerate limit distribution G,

$\lim_{n\to+\infty} P\left(\frac{M_n-b_n}{a_n}\le x\right)=G(x),\ x\in\mathbb{R}$. Than G is as follows[2]:

$$\begin{aligned}&\text{Gumbel} && G_0(x)=\exp[-e^{-x}] && x\in\mathbb{R}.\\ &\text{Fréchet:} && G_{1\alpha}(x)=\begin{cases}\exp[-x^{-\alpha}] & x>0,\alpha>0\\ 0 & x\le 0\end{cases} && (1.2)\\ &\text{Weibull} && G_{2\alpha}(x)=\begin{cases}\exp[-|x|^{-\alpha}] & x<0,\ \alpha<0\\ 1 & x\ge 0\end{cases}\end{aligned}$$

By taking the re-parametrization $\delta=\frac{1}{\alpha}$ - due to von Mises, one obtains a continuous, unified model called GEV

$$G_\delta(x)=\begin{cases}\exp\left[-(1+\delta x)^{-\frac{1}{\delta}}\right] & for\ \delta\neq 0\ e(1+\delta x)>0\\ \exp[-e^{-x}] & for\ \delta=0\end{cases} \quad (1.3)$$

Alternatively, any extreme value distribution can be represented as

$$G_{\delta,\mu,\sigma}=G_\delta\left(\frac{x-\mu}{\sigma}\right)=\begin{cases}\exp\left[-\left(1+\delta(\frac{x-\mu}{\sigma})\right)^{\frac{-1}{\delta}}\right]\delta\neq 0 & 1+\delta(\frac{x-\mu}{\sigma})>0\\ \exp\left[-e^{-\left(\frac{x-\mu}{\sigma}\right)}\right]\delta=0\end{cases} \quad (1.4)$$

Where μ and $\sigma>0$ are the location and scale parameters and δ[3] indicates the shape parameter. The GEV parameter estimation is made by applying the maximum likelihood[4] principle using the toolbox EVIM in Matlab.

3. The Dataset

We have considered four sets of historical intraday returns for FTSE-MIB index. As shown by (Malevergne, Y., Sornette , D, 2006, 53), in order to remove the effect of price overnight jumps, the intraday returns will be determined separately for each day belonging to the dataset, only then, the union of all the days provides a better overall dataset of intraday returns.

In an elaborate way, the first series consists of intraday returns with a frequency of one minute, calculated in log scale, which spans the period from 01.04.2011 until 30.09.2011. Whereas the period of consideration has not varied for the other series, we have observed that the frequencies which we calculate the returns with, have. In detail, we have taken in consideration returns generated in 1, 5, 10 and 15 minutes.

The data have been retrieved from Bloomberg.

3.1 Stylized statistical properties of asset returns

Normally, the analysis of financial returns implies the same basic assumptions for the quantitative models development. Usually, in the initial hypothesis, it is assumed that the return distribution follows the Gaussian distribution, - the observed information is uncorrelated in time and the return variation per time unit remains constant in time. However, when financial time series are analysed, empirical studies have shown (McNeil, Frey& Embrechts, 2005, 117) some peculiarities which do not match with the assumptions listed above. There are some facts, so-called stylized, that have a tendency to persist over time. They are very useful for the purpose of forming hypotheses

---

[2] *we have used the same symbols as in Reiss, R and D., Thomas, M (2007, 14)*

[3] *Indicate the shape parameter:* $\begin{cases}\delta=0\ Gumbel\\ \delta>0\ Fréchet\\ \delta<0\ Weibull\end{cases}$

[4] *Let $M_1,M_2,\ldots,M_m$ be an iid random sample belonging to the GEV distribution, than if $\delta\neq 0$, the log-likelihood for the parameters, will be* $Log(\delta,\mu,\sigma)=-mlog\sigma-\left(1+\frac{1}{\delta}\right)\sum_{i=1}^{m}log\left[1+\delta\frac{M_i-\mu}{\sigma}\right]-\sum_{i=1}^{m}[1+\delta\left(\frac{M_i-\mu}{\sigma}\right)]^{\frac{-1}{\delta}}$ *,for* $1+\delta\frac{M_i-\mu}{\sigma}>0.\ i=1\ldots m$*.If* $\delta=0$ *, then* $Log(\mu,\sigma)=-mlog\sigma-\sum_{i=1}^{m}(\frac{M_i-\mu}{\sigma})-\sum_{i=1}^{m}exp\{-(\frac{M_i-\mu}{\sigma}).$

underlying the returns distribution and are valid regardless the nature of activity and the time period considered. In particular, we can make the following observations:

1. The distribution of returns is not well represented by the normal distribution because of the empirical distributions observed heavy tails and gain/loss asymmetry in the distribution
2. The returns are not independent and identically distributed (iid). In fact, the variance is not constant over time (volatility clustering, volume/volatility correlation, intermittency)

To test the hypothesis regarding the normality distribution of intraday returns, as first criterion, we report the series histograms. Although the empirical distribution is approximately symmetric (Figure1.), it is, however, noted that the intraday returns tails are heavier than the normal distribution. Based on these findings, we observe that empirical distributions are more pointed at the centre showing a higher probability of returns close to zero. As a second criterion, we observe the Q-Q[5]plot (figure 2) where it is shown that the non-linearity (with respect to the standard normal distribution) is very clear, especially for the extreme values- the quantiles at the distributions tails.

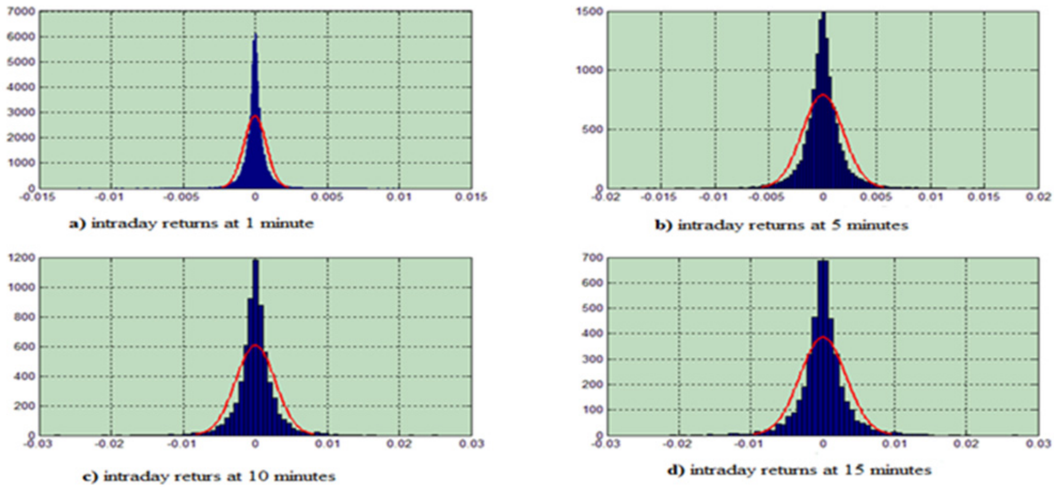

Fig.1.Histogram ofintraday distribution a) series at 1 minute, b) series at 5 minutes, c) series at 10 minutes, d) series at 15 minutes.

Table 1 (below) depicts the moment[6] for the four considered series. It is observed that the outcome confirms the empirical evidence of the intraday returns non-normality. We have also performed the JB[7] test and the outcome confirms the hypothesis of non-normality

[5] *If the parametric model fits the data well, this graph must have a linear form.The more linear the Q-Q plots, the more appropriate the model in terms of goodness of fit.*

[6] *Sample moments are used to estimate the unknown parameters of a distribution and to calculate the standard errors of such estimates.*

[7] $JB = \left(\frac{\hat{\gamma}_1}{\sqrt{\frac{6}{n}}}\right)^2 + \left(\frac{\hat{\gamma}_2}{\sqrt{\frac{24}{n}}}\right)^2$

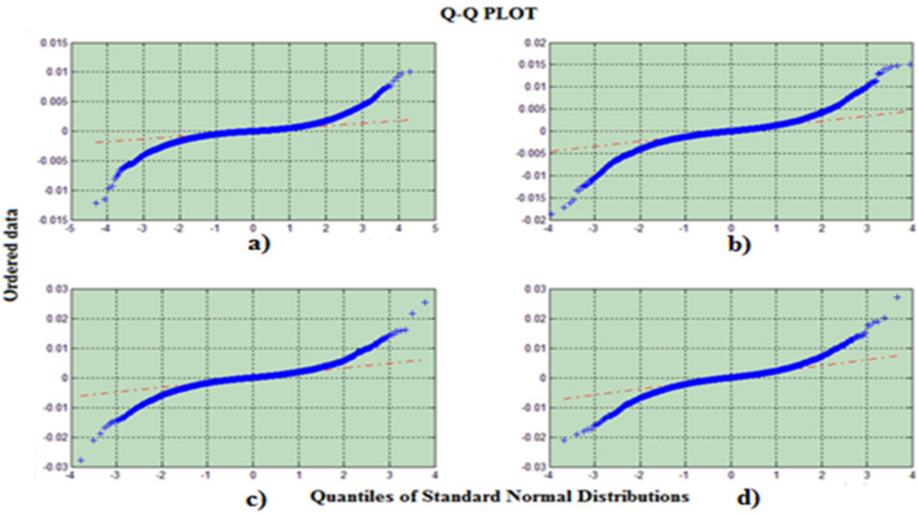

Fig.2. Q-Q plot of Intraday returns. a) series at 1 minute, b) series at 5 minutes, c) series at 10 minutes, d) series at 15 minutes.

Table 1: The Moment of intraday returns

Series	Mean	Variance	Skweness[8]	Kurtosis[9]
FTSE-MIB1	3.9119e-006	6.1100e-007	0.1625	15.6719
FTSE-MIB5	1.9560e-005	3.6262e-006	-0.0687	12.1625
FTSE-MIB10	3.9116e-005	7.5456e-006	-0.1305	12.1959
FTSE-MIB15	5.8666e-005	1.0308e-005	0.0918	9.6601

In order to verify the existence of time dependence and study its characteristics, it is helpful to look at the figures that represent the trend of intraday returns for the considered period. We have seen that the figure displays the volatility clustering. Alternatively, a statistical test that can be used to test the hypothesis that the process is a white noise is that of Box-Ljung[10]. Making the test, by choosing $\alpha = 5\%$ and $m = 50$, shows that there is linear autocorrelation.

4. Fitting the GEV distribution

4.1 The Block Maxima Method

Based on Fisher-Tippett, Gnedenko theorem, GEV provides a model for the maxima distribution. Its application consists in dividing the sample into blocks with equal length. We refer to block maximum of the *j*th block by M_{nj} so that our data are $M_{n1,\dots,}M_{nm}$. However, we have observed that the block size selection for each dataset can be critical in the model implementation process. There is a trade-off between bias and variance. The GEV distribution can be fitted using various methods. We have chosen to implement the maximum likelihood method. In this implementation, it will be assumed that block size *n* is quite large, and block maxima observation is independent.

---

[8] $\gamma_1 = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$

[9] $\gamma_2 = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] - 3$

[10] $BL(m) = n(n+2)\sum_{h=1}^{m}\frac{\hat{\rho}^2(h)}{n-h}$, *we reject hypothesis H_0 if BL(m)* $>\chi^2_{(m,1-\alpha)}$.

4.2 Block Maxima analysis of FTSE-MIB intraday returns.

We fit the GEV distribution to intraday maximum data for FTSE[11]-MIB series, dividing it into 128 blocks, each one including a sample of 505 intraday data, in so doing, building a sample with 128 intraday maximum losses. In figure 3 we have reported the maxima[12] time series for FTSE-MIB series with frequency of 1 minute and FTSE-MIB with frequency of 5 minutes.

In this maximum likelihood implementation, the estimated parameter values $\hat{\delta} = 0.3025$, $\hat{\mu} = 0.2272$, and $\hat{\sigma} = 0.1146$ have a standard error of 0.0968, 0.0121, and 0.0103 respectively.

The shape parameter $\hat{\delta} = 0.3025$ indicates that the fitted distribution is a heavy-tailed Fréchet distribution with infinite fourth moment[13].

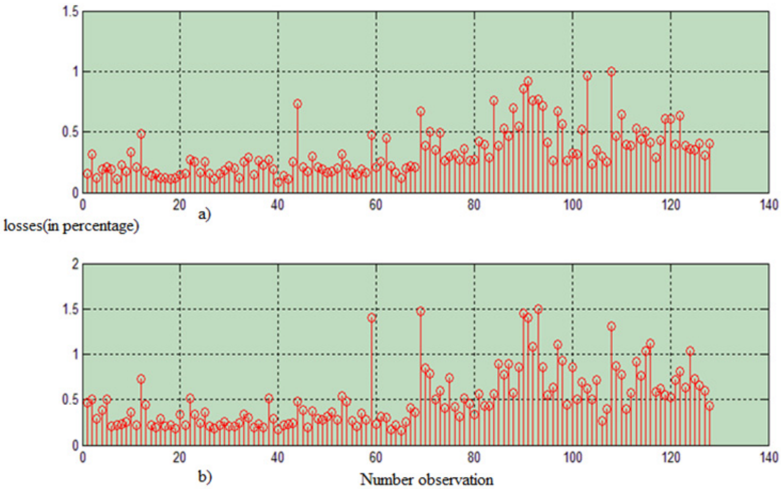

Fig.4 Maxima time series. a) series at 1 minute, b) series at 5 minutes

As empirical studies have shown (McNeil, Frey& Embrechts, 2015, 144) to increase the number of blocks we also fit a GEV model to 256 day maxima[14] and obtain the parameter estimates $\hat{\delta} = 0.3562$, $\hat{\mu} = 0.1814$ and $\hat{\sigma} = 0.0914$ with standard error 0.0698, 0.0068 and 0.006. We notice that in case the number of blocks is increased, by reducing the number of observation for each block, the estimates are more accurate. In fact, the standard_deviations for the shape parameter, when the number of blocks is increased, decrease from 0.0968 to 0.0698.

We have obtained the parameters for the other considered series in a similar way and we have noticed a bias-variance trade-off.

4.3 Extreme scenarios

One of the major objectives of extreme value analysis is the analysis of stress losses. Like in (McNeil, Frey& Embrechts, 2015, 144), we have focused in two related quantities: in the first approach, we have defined return level; and in the second the return period problem.

---

[11] *The intraday returns data with frequency of 1 minute, expressed in percentage, for FTSE-MIB index. The loses will be denoted as positive, and we will consider the maximum losses for each day.*
[12] *Maximum losses for each day.*
[13] $\hat{\alpha} = \frac{1}{\hat{\delta}} = 3.305$
[14] *We consider two maxima for each day*

Return level. Taking definition (1.4)[15] for GEV df, the extremes quantiles estimation for maxima limiting distribution is obtained by inverting equation (1.4):

$$x_p = \begin{cases} \hat{\mu} + \frac{\hat{\sigma}}{\hat{\delta}}\left((-\ln(1-\frac{1}{k})^{-\hat{\delta}} - 1\right) & for \quad \delta \neq 0 \\ \hat{\mu} + \hat{\sigma}\ln\left(\ln\left(1-\frac{1}{k}\right)\right) & for \quad \delta = 0 \end{cases}$$

Where $G_{\delta,\mu,\sigma}(x_p) = 1 - \frac{1}{k}$, and x_p indicate the exceeded one in every k n-blocks, on average.

Return period. Denoting by G the df of maxima, the return period is the time we would expect to observe a single block in which a particular level was exceeded. The event return period $\{M_n > \mu\}$ is given by $k_{n,\mu} = \frac{1}{1-G(\mu)}$ where *n* is the number of observations in a block while μ indicates the threshold.

Stress Losses, application. By estimating the return level for series FTSE-MIB, (intraday data with frequency at 1 minute), choosing k = 20, the point estimator of the monthly return level is 0.7787%,[16] with 95% confidence interval.

If we make an effort to estimate the return period of μ = 1.5% loss, considering the FTSE-MIB series with data sampling frequency of one minute of intraday returns, the point estimation is 130 days. In a similar way we can derive the return level and return period for the other series.

In figure 4 we have reported the comparison between empirical[17] and theoretical distributions of GEV and we have observed that the empirical distribution is well approximated by GEV distribution, especially for extreme values of the distributions.

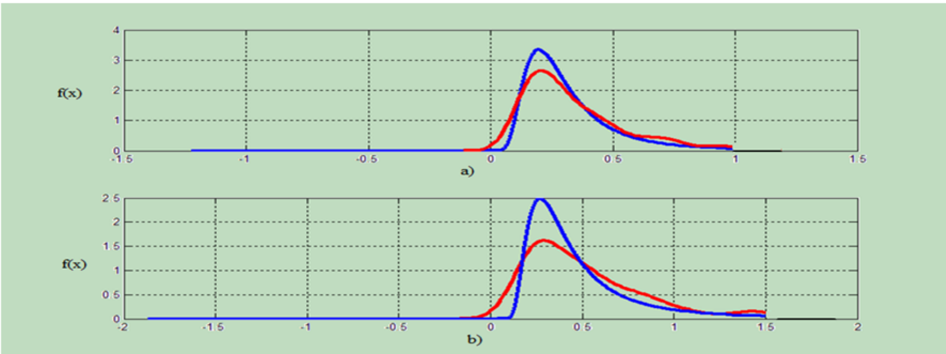

Fig.4. Comparison between empirical and theoretical distributions of GEV, a) series at 1 minute, b) series at 5 minutes

5. Conclusion

The present work has dealt with the BLOCK MAXIMA approach, the theory of extreme values, from tow points of view: a) the theoretical approach related to EVT, and b) the application of EVT in intraday returns. It is observed, from the theoretical point of view, that EVT demonstrates some significant advantages: it suggests methods with strong theoretical base for modelling extreme

---

[15] $$G_{\delta,\mu,\sigma} = G_\delta\left(\frac{x-\mu}{\sigma}\right) = \begin{cases} exp\left[-\left(1+\delta(\frac{x-\mu}{\sigma})\right)^{\frac{-1}{\delta}}\right] & \delta \neq 0 \quad 1+\delta(\frac{x-\mu}{\sigma}) > 0 \\ exp\left[-e^{-(\frac{x-\mu}{\sigma})}\right] & \delta = 0 \end{cases} \quad (1.4)$$

[16] *Represent the losses which are exceeded in one out every month on average (considering that one day represent one block). We are dividing the series into 128 blocks of the same length (505 observations) and choosing the maximum (losses) from each block. We are considering k = 20.*

[17] *The probability density function is estimated by Kernel estimator.*

events; flexibility is also the accuracy of the modelling enhanced by the key feature of EVT, which includes examining exclusive of the tail of the distribution of intraday data, ignoring the centre of the distribution. In the end, the possibility to use parametric methods consents to the forecasting of extreme events.

The major disadvantages of EVT are: - the choice of trade-off given that there is still no consensus in the literature so far, on how the choice is to be made; - secondly, the theory assumes that the data are not serially correlated[18]. By applying maximum likelihood approach we obtained the estimation of the parameters and it is observed that by providing the shape parameter greater than zero, the GEV appropriate is the Fréchet. We have also considered the return value, i.e. the value that is expected to be exceeded on average once every k-block, and the return period which expresses the time you have to wait so that a threshold value is exceeded.

As already suggested from literature, the extreme value analysis is of interest in many areas such as telecommunications, ecology, geology and meteorology. In relation to insurance, EVT can be used in the pricing of reinsurance contracts, particularly in Excess of loss treaty. Lately, Albania is considering the opportunity to implement the catastrophe (re)insurance treaty In order to quantify the premium of this insurance coverage, the extreme value analysis could be used.

References

Bali, T. G. (2007). “A generalized extreme value approach to financial risk measurement”. *Journal of Money, Credit and Banking39(7):*1613–1649.

Balkema, G., A. A. Balkema and P. Embrechts (2007). High Risk Scenarios and Extremes. A Geometric Approach. Zurich: European Mathematical Society.

Beirlant, J., Y. Goegebeur, J. Segers and J. Teugels (2004). Statistics of Extremes: Theory and Applications. *Chichester (England): Wiley.*

De Haan, L. e Ferreira, A (2006). Extreme Value Theory: An Introduction. Springer Series in Operations Research and Financial Engineering. Boston.

Embrechts, P,Klüppelberg, C and Mikosch. Th (1997) .Modelling extremal events, for Insurance and Finance. Springer-Verlag, Berlin.

Fisher, R.A. and Tippett, L.H.C. Limiting forms of the frequency distribution of the largest and smallest member of a sample, *Proc. Camb. Phil. Soc.,24, 180-190*, 1928.

Gençay, R, Selçuk, F, Ulugulyagcı, A (2001). EVIM: A Software Package for Extreme Value Analysis in MATLAB

Gumbel, E.J (1958). Statistics of Extremes, Columbia University Press, New York.

Kotz, S., Nadarajah, S (2000). "Extreme Value Distributions: Theory and Applications." London: Imperial College Press.

Malevergne, Y., Sornette (2006). Extreme Financial Risks From Dependence to Risk Management (revised edition). Springer Berlin Heidelberg New York.

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

McNeil, A. J., R. Frey and P. Embrechts (2015). Quantitative Risk Management. Princeton (New Jersey): *Princeton University Press*.

N. Balakrishnan, Jose Maria Sarabia, Enrique Castillo and Ali S. Hadi (2004). “Extreme Value and Related Models with Applications in Engineering and Science”. Wiley.

Pickands, J. III (1975). “Statistical inference using extreme order statistics”. *Ann. Statist. 3:119–131.*

Reiss, R.-D., Thomas, M (2007). Statistical Analysis of Extreme Values, with Application to Insurance, Finance, Hydrology and Other Fields, 2nd edition; 3rd edition, *Birkhuser Verlag*.

Stuart Coles(2001). An introduction to statistical modeling of extreme values. Springer Series in Statistics. Springer-Verlag London Ltd., London.

Zhou, C. (2009b). “Existence and consistency of maximum likelihood estimator for the extreme value index”. Journal of Multivariate Analysis 100:794–815

[18] *The problem can be overcome by increasing the number of the sample*