STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **6**, June 2018, pp 240–247.
Published online in International Academic Press (www.IAPress.org)

POT approach for estimation of extreme risk measures of EUR/USD returns

Hakim Ouadjed*

Mascara University, Algeria

Abstract Leadbeter *et al* (*M.R.,G.Leadbetter, G.Lindgren,and H.Rootzen, Extremes and Related Properties of Random Sequences and Processes, Springer Series in Statistics. Springer-Verlag: New York, 1983.*) have generalized the extreme value theory of i.i.d. in the case of the stationary process, where it have defined an extremal index $\theta \in]0,1[$ for measuring the degree of dependence at the extremes, this parameter measures how the extremes cluster together and $1/\theta$ is interpreted as the average size of these clusters. Using this parameter and the Peak Over Threshold method which involves the Generalized Pareto Distribution we estimate in this work the extreme quantile and the conditional tail expectation for EUR/USD returns.

Keywords Theory of extreme values, Tail index estimation, Extremal index

AMS 2010 subject classifications 60G70, 62G32

DOI: 10.19139/soic.v6i2.395

1. Introduction

The risk assessment have assumed important profiles in many financial institutions. Value-at-Risk (VaR) is probably one of the most widely used measures of risk. If X is a random variable of loss with continuous df $F(x)$, and p be a probability level such that $0 < p < 1$, the Value-at-Risk at probability level p, denoted by $VaR_p(X)$, is the p-quantile of X . That is

$$VaR_p(X) = F^{-1}(p).$$

The probability level p is usually taken to be close to 1 (say, 0.95, 0.99 or 0.999).

A drawback of VaR is that it only makes use of the cut-off point corresponding to the probability level p and does not use any information about the tail distribution beyond this point, it has also been criticized for lacking a certain property that desirable risk measures should meet to be a coherent risk measure (see Artzner *et al* [1].

The conditional tail expectation (CTE) corrects for this. The idea of the CTE is to measure the average severity of the loss when the extreme loss does occurs, where the extreme loss is represented by the VaR. Formally, the CTE of the random variable X at probability level p is defined as

$$CTE_p(X) = E[X|X > VaR_p(X)]$$

Let the order statistic $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ associated to the sample $(X_1, X_2, \ldots, X_n)$ of X. The empirical estimate of $VaR_p(X)$ is

*Correspondence to: Hakim Ouadjed (Email: o_hakim77@yahoo.fr). Department of Gestion, Faculty of Science of Gestion, Economic and Commerce, Mascara university, Algeria.

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

$$\widehat{VaR}_p^{emp}(X) = X_{[np]:n}, \tag{1}$$

and the empirical estimate of $CTE_p(X)$ is

$$\widehat{CTE}_p^{emp}(X) = \frac{1}{1-p}\left(\frac{1}{n}\sum_{k=[np]+1}^{n} X_{k:n} + \left(\frac{[np]}{n} - p\right) X_{[np]:n}\right)$$

(see Rachev *et al* [8]), where $[x]$ is the integer part of x.

The approach of using $\widehat{VaR}_p^{emp}$ is a poor method in the tail of the distribution where data become sparse and it is impossible to obtain an out-of-sample VaR estimate with $\widehat{VaR}_p^{emp}$.

We can extrapolate VaR and CTE from a high threshold using the POT (Peak Over Threshold) method based on extreme value theory in order to characterize the tail of loss.

The extreme values are based on the assumption that the data are i.i.d. random variables. However, in reality extremal events tend to occur in clusters because of the serial dependence in the data. It is interesting to extend the POT method to cases in which the data form a strictly stationary time series. The basic concept of the extension is the extremal index which is a measure of tail dependence.

2. POT method

The modelling of extremes may be done in two different ways: modelling the *maximum* of a collection of random variables, and modelling the *largest values* over some high threshold.

The Fisher-Tippett theorem is one of two fundamental theorems in EVT (Extreme Value theory). It plays the same role as the Central Limit Theorem plays in the studies of sums of a random variables.

Theorem 1 (Fisher-Tippett [5])
Let (X_n) be a sequence of i.i.d. random variables with distribution F. Let $M_n = \max(X_1, \ldots, X_n)$. If there exist norming constants $c_n > 0$ and $d_n \in \mathbf{R}$ and some non-degenerate distribution function G such that

$$\frac{M_n - d_n}{c_n} \xrightarrow{D} G, \tag{2}$$

then G is one of the following three types :

(i) Gumbel

$$\Lambda(x) = \exp(-e^{-x}), x \in \mathbb{R},$$

(ii) Fréchet

$$\Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0, \alpha > 0, \end{cases}$$

(iii) Weibull

$$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^\alpha) & x \leq 0, \alpha > 0 \\ 1, & x > 0. \end{cases}$$

It is possible to combine these three laws in a single form using the parametrization of Jenkinson and von Mises. The unified form is as follows:

$$G_\xi(x) = \begin{cases} \exp(-(1+\xi x)^{-1/\xi}) & if \quad \xi \neq 0 \\ \exp(-\exp(-x)) & if \quad \xi = 0 \end{cases}$$

where the parameter ξ is called a *shape parameter* or an *extreme value index* and the support of $G_\xi(x)$ corresponds to

$$\begin{array}{lll} x > -\xi^{-1} & for & \xi > 0, \\ x < -\xi^{-1} & for & \xi < 0, \\ x \in \mathbb{R} & for & \xi = 0. \end{array}$$

However, account must be taken of the location $\mu \in \mathbf{R}$ and scale $\sigma > 0$ of the distribution. By including these parameters, this formulation can be synthesized in the form of the distribution family, called the Generalized Extreme Value (GEV) distribution.

$$G_{\xi,\mu,\sigma}(x) = \begin{cases} \exp(-(1+\xi(x-\mu)/\sigma)^{-1/\xi}) & if \quad \xi \neq 0 \\ \exp(-\exp(-(x-\mu)/\sigma)) & if \quad \xi = 0 \end{cases} \tag{3}$$

where $1 + \xi(x-\mu)/\sigma > 0$. We distingue the following three cases $\xi = \alpha^{-1} > 0$ corresponds to the distribution of Fréchet, $\xi = 0$ corresponds to the distribution of Gumbel and $\xi = -\alpha^{-1} < 0$ corresponds to the distribution of Weibull.

The more modern approach to modelling extreme events is to attempt to focus not only the largest (maximum) events, but on all events greater than some large preset threshold. This is referred to as peaks over threshold (POT) modelling, for this method we will discuss about the parametric approach based on the generalized Pareto distribution (GPD) given by

$$H_{\xi,\beta}(x) = \begin{cases} 1 - (1+\frac{\xi x}{\beta})^{-1/\xi} & if \quad \xi \neq 0 \\ 1 - \exp(-\frac{x}{\beta}) & if \quad \xi = 0 \end{cases} \tag{4}$$

where

$$\begin{array}{lll} x \geq 0 & if & \xi \geq 0, \\ 0 \leq x < -\frac{\beta}{\xi} & if & \xi < 0 \end{array}$$

Let $X \sim F$ with right-end-point $x_F = \sup\{x \in \mathbf{R};\ F(x) < 1\}$. For any high threshold $u < x_F$ define the excess distribution function

$$\begin{aligned} F_u(x) &= \mathbf{P}[X - u \leq x | X > u], \quad 0 \leq x < x_F - u \\ &= \frac{F(u+x) - F(u)}{1 - F(u)} \end{aligned} \tag{5}$$

The mean excess function of X is then

$$e(u) = \mathbf{E}[X - u | X > u]$$

Balkema and de Haan [2] proved that for a sequence $X_1, \ldots, X_n$ of i.i.d. random variables, the distribution $F_u(x)$ converge to the generalized Pareto distribution $H_{\xi,\beta}(x)$. The convergence can be described by the following expression

$$\lim_{u \to x_F} \sup_{0 < x < x_F - u} |F_u(x) - H_{\xi,\beta}(x)| = 0 \tag{6}$$

Note that by (5) above for $x > u$, we may write

$$\bar{F}(x) = \bar{F}(u)\bar{F}_u(x-u)$$

Assuming that u is sufficiently large, we may then use the approximation (6) and the empirical estimator, for $\bar{F}(u)$,

$$\widehat{\bar{F}(u)} = \frac{N}{n};\ N = \sum_{i=1}^{n} 1_{\{X_i > u\}},$$

and where n is the total number of observations. The upper tail of $F(x)$ for all $x > u$ may then be estimated by

$$\widehat{\bar{F}(x)} = \frac{N}{n}\left(1 + \widehat{\xi}\frac{x-u}{\widehat{\beta}}\right)^{-1/\widehat{\xi}} \tag{7}$$

To obtain the estimator of VaR_p, one simply inverts the estimator (7), which yields

$$\widehat{VaR}_p^{iid} = u + \frac{\widehat{\beta}}{\widehat{\xi}}\left(\left(\frac{n}{N}(1-p)\right)^{-\widehat{\xi}} - 1\right), \tag{8}$$

The parameters estimation $\widehat{\xi}$ and $\widehat{\beta}$ can be founded by maximum likelihood (see Embrechts *et al* [4]). Furthermore, for $\xi < 1$ we obtain the following CTE estimator

$$\widehat{CTE}_p^{iid} = \frac{\widehat{VaR}_p^{iid}}{1-\widehat{\xi}} + \frac{\widehat{\beta} - \widehat{\xi}u}{1-\widehat{\xi}} \tag{9}$$

3. The Extremal Index

The main assumption in EVT is that the extreme observations are independent and identically distributed. This is not always fulfilled when working with real data.

Primary result incorporating dependence in the extremes is summarized in Leadbetter *et al* [7]. For a strictly stationary time series (X_i) under some regularity conditions for the tail of F and for some suitable constants $c_n > 0$ and $d_n \in \mathbf{R}$, as the sample size $n \to \infty$

$$\frac{M_n - d_n}{c_n} \xrightarrow{D} (G)^{\theta}, \tag{10}$$

where $\theta \in]0,1[$ is the extremal index and G is the GEV distribution defined in (2). The extremal index θ is the key parameter extending extreme value theory from i.i.d. random processes to stationary time series and influences the frequency with which extreme events arrive as well as the clustering characteristics of an extreme event. The quantity $1/\theta$ has a convenient heuristic interpretation, as it may be thought of as the mean cluster size of extreme values in a large sample.

As a consequence, the maximum of n observations of a stationary series with an extremal index θ behaves like the maximum of $n\theta$ observations of the i.i.d. associated series with the same mariginal distribution.

The problem of estimating θ has received some attention in the literature (see Smith and Weissman [9], Weissman and Novak [11], Ferro and Segers [6]), Süveges [10] presents the maximum likelihood estimator as

$$\widehat{\theta}_{ML} = \frac{\sum_{i=1}^{N} qS_i + N - 1 + N_c - \left[\left(\sum_{i=1}^{N-1} qS_i N - 1 + N_c\right)^2 - 8N_c \sum_{i=1}^{N-1} qS_i\right]^{1/2}}{2\sum_{i=1}^{N-1} qS_i} \tag{11}$$

where $S_i = T_i - 1$, with T_i are the inter-exceedance times and N is the number of exceedances of a high threshold u and q is the estimate of $\bar{F}(u)$, and $N_C = \sum_{i=1}^{N-1} \mathbf{1}_{\{S_i \neq 0\}}$.

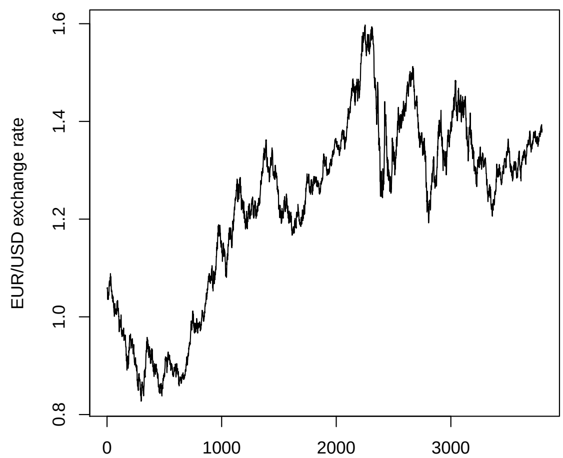

Figure 1: Exchange rate series of EUR/USD

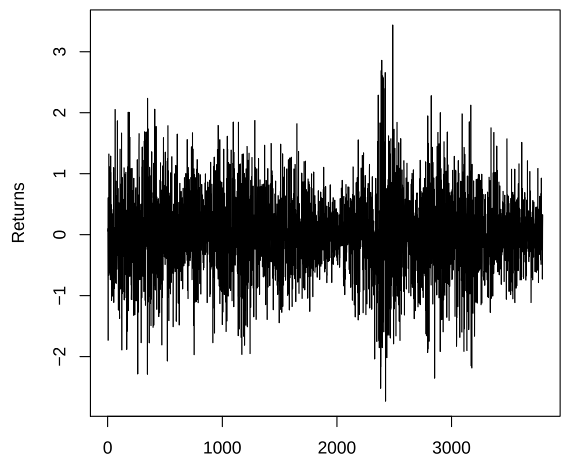

Figure 2: Returns of EUR/USD

An important property in practice is that one can first adjust a GPD as if the data were i.i.d., then estimate θ. Coles [3] generalized the extreme quantile estimator (8) to the dependent case with a high threshold $u > 0$ as follows

$$\widehat{VaR}_p^{dep} = u + \frac{\widehat{\beta}}{\widehat{\xi}}\left(\left(\frac{n\widehat{\theta}(1-p)}{N}\right)^{-\widehat{\xi}} - 1\right) \tag{12}$$

Then we have

$$\widehat{CTE}_p^{dep} = \frac{\widehat{VaR}_p^{dep}}{1-\widehat{\xi}} + \frac{\widehat{\beta}-\widehat{\xi}u}{1-\widehat{\xi}} \tag{13}$$

4. Data Analysis

The analysis of actual financial data such as the value of an exchange rate, or financial indices is very complex because of the influence of several factors. Many of these series appear heteroscedastic, they are not stationary.

If we have a series $(X_i), i = 1, \ldots, n$, then we define the returns by

$$R_i = 100 \times (\log(X_{i+1}) - \log(X_i)), i = 1, \ldots, n-1,$$

we will treat, as an example, the daily returns for the exchange rate EUR/USD during the period from 06/09/1999 to 24/03/2014, the 3796 observations are shown in figure 1. The data, taken from the website ”https://fr.investing.com/currencies/eur-usd-historical-data”.

The corresponding 3795 returns in figure 2, showing the non-regularity of the increases and the appearance of a number of extreme positive and negative variations. The series of returns whose statistical characteristics are summarized in the table 1, is distributed in figure 3.

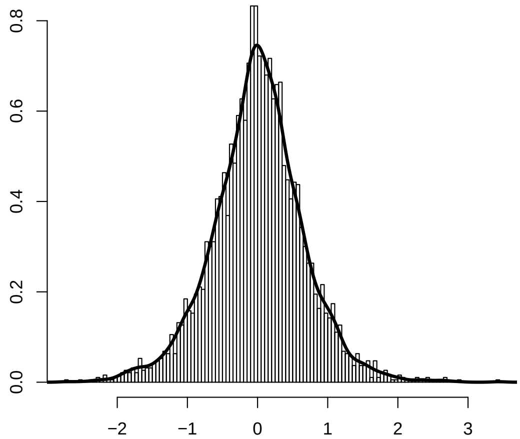

Figure 3: Distribution of EUR/USD returns

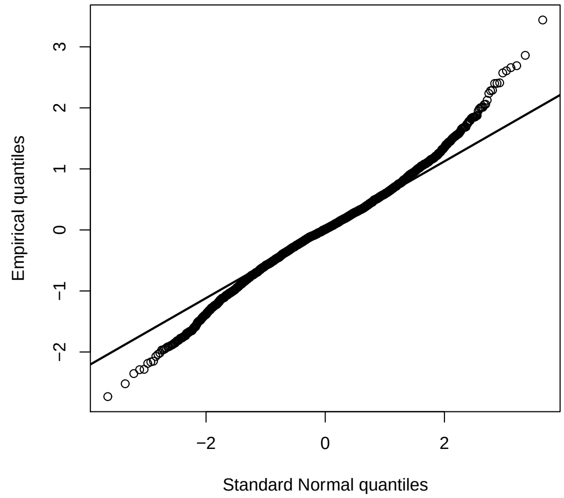

Figure 4: qq plot of EUR/USD returns

Mean	0.007071
Standard deviation	0.6506757
Min	-2.731
Max	3.44
1st Qu	-0.3743
3rd Qu	0.3814
Median	0.01086

Table 1. Empirical characteristics of EUR/USD returns

The value of the kurtosis of this financial series is $4.227 > 3$, which would lead to the invalidity of the Gaussian fit. The qq-plot in the figure 4 shows a clear deviation from the right direction especially at the extremities. On the other hand, the $p - value$ of the Shapiro-Wilk test statistic is 3.979×10^{-16}, thus confirming the rejection of the assumption that the returns would normally be distributed.

To fit the GPD to the threshold excesses of the returns series we first make a subjective choice of an appropriate threshold value.

On the figure 5, we plot the sample mean excess estimates

$$\left\{ \left(u, \frac{1}{N} \sum_{i=1}^{N} (X_i - u), \ u < X_{\max} \right) \right\}$$

where $X_1, \ldots, X_N$ consist of the N observations that exceed u and $X_{\max}$ is the largest of the X_i. We look for the point u such that the plot is a straight line passed this point. However, we know that when u is too high, the mean excess function is poorly estimated, so it is not relevant to look at it for these points. We notice a slope break from the 0.9 point, which leads to a straight line. This point is therefore the optimal threshold for the mean excess function.

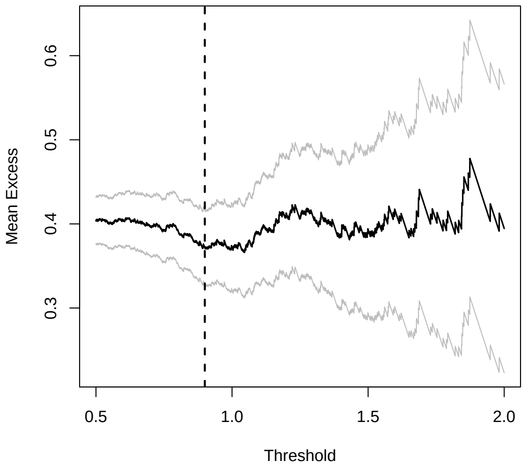

Figure 5: The mean excess function

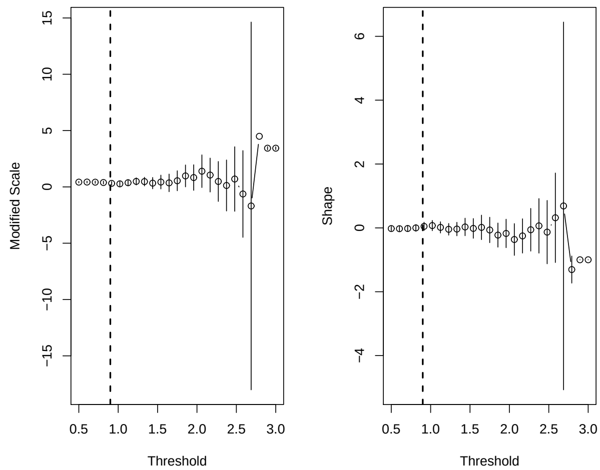

Figure 6: Threshold Choice plot

For different thresholds u, the maximum likelihood estimates for the shape and the modified scale parameter (modified by subtracting the shape multiplied by the threshold) are plotted against the thresholds (see figure 6. If the threshold u is a valid threshold to be used for peaks over threshold modelling, the parameter estimates depicted should be approximately constant above u. Based on figure 6, we choose the threshold $u = 0.9$ because the parameter estimates are approximately constant above 0.9.

The results of maximum likelihood estimation of the GPD parameters (with the chosen threshold $u = 0.9$) are $\widehat{\beta} = 0.3542872$ and $\widehat{\xi} = 0.04823426$.

Now we estimate the extremal index using the estimator (11) with $u = 0.9$ we find $\widehat{\theta}_{ML} = 0.9117315$, and we estimate $VaR_p^{iid}, CTE_p^{iid}, VaR_p^{dep}$ and CTE_p^{dep}. By way of comparison, the empirical quantiles and the empirical CTE of the returns are also presented in the following table

p	0.95	0.99	0.999
$\widehat{VaR}_p^{emp}$	1.060574	1.677017	2.606811
$\widehat{CTE}_p^{emp}$	1.440882	2.082713	2.928904
$\widehat{VaR}_p^{iid}$	1.061674	1.667644	2.620619
$\widehat{CTE}_p^{iid}$	1.442109	2.07879	3.08006
$\widehat{VaR}_p^{dep}$	1.095209	1.703886	2.661118
$\widehat{CTE}_p^{dem}$	1.477344	2.116868	3.122612

Table 2. Estimation of extreme quantiles and CTE of the returns.

The $\widehat{VaR}_p^{iid}$ and $\widehat{VaR}_p^{dep}$ are higher than the $\widehat{VaR}_p^{emp}$. We can say that the VaR estimators using the POT method is more relevant because it involves a larger number of points in the calculations. In fact, the $\widehat{VaR}_p^{emp}$ take into account only one data item, namely the greatest loss, while the POT approach takes into account all the values above a high threshold.

As expected the $\widehat{VaR}_p^{dep}$ and $\widehat{CTE}_p^{dep}$ are higher than $\widehat{VaR}_p^{iid}$ and $\widehat{CTE}_p^{iid}$ when the extremal index is neglected.

5. Conclusion

The advantage of the POT method is that we do not impose a strong hypothesis on the initial distribution. It allows us to establish explicit and simple relationships between VaR and CTE. These latter can serve as an indicator of extreme risk, through GPD model which fits tail distribution of the daily returns of EUR/USD more accurately. To not underestimate the VaR and CTE it is necessary to take into account in the POT method the extremal index .

Nevertheless, the estimators obtained by the POT method are often sensitive to the choice of the threshold, which is the reason why we must choose it well.

Acknowledgement

The authors would like to thank the referee for careful reading and for their comments which greatly improved the paper.

REFERENCES

1. P. Artzner, F. Delbaen, J. M. Eber, and D. Heath, *Coherent Measures of Risk*, Mathematical Finance, vol. 9, pp. 203–228, 1999.
2. A. A. Balkema, and L. De. Haan, *Residual life time at great age*, Annals of Probability, vol. 2, pp. 792–804, 1974.
3. S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics. Springer-Verlag: London, 2001.
4. P. Embrechts, C. Kluppelberg, and T. Mikosch, *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin, 1997.
5. R. A Fisher, and L.H.C. Tippett, *Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample*, Proceedings of the Cambridge Philosophical Society, vol. 2, pp. 180–190, 1928.
6. C. A. T. Ferro, and J. Segers, *Inference for Clusters of Extreme Values*, Journal of the Royal Statistical Society, Series B, vol.65, no.2, pp. 545-556, 2003.
7. M. R., G. Leadbetter, G. Lindgren, and H. Rootzen, *Extremes and Related Properties of Random Sequences and Processes*, Springer Series in Statistics. Springer-Verlag: New York, 1983.
8. S. T. Rachev, S. Stoyanov, and F. J Fabozzi, *Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures*, New Jersey: John Wiley & Sons., 2007.
9. R. L Smith, and I. Weissman, *Estimating the extremal index*, Journal of the Royal Statistical Society, Series B, vol. 56, pp. 515–528, 1994.
10. M. Süveges, *Likelihood estimation of the extremal index*, Extremes, vol. 10, pp. 41–55, 2007.
11. I. Weissman, and S. Y. Novak, *On blocks and runs estimators of the extremal index*, J. Statist. Plann. Inference, vol. 66, pp. 281–288, 1998.