Master Thesis : Extreme Value Theroy Applied to Securitizations Rating Methodology

Tarek Barbouche

March 23, 2017

Sammanfattning

Värdepapperisering är en av dagens finansiella trender. Att utvärdera värdepapperisering risk kräver starka kvantitativa kunskaper och en förståelse för både kredit- och marknadsrisk. För internationell värdepapperisering är det obligatoriskt att hänsyn tas till valutarisker. Vi kommer att se de olika metoder för att utvärdera extrema variationer i valutakurser med hjälp av extremvärdesteori och Monte Carlo-simuleringar.

Keywords: Extremvärdesteori, Valutakurser, Block Maxima, Peaks-over-Threshold, Värdepapperisering

Abstract

One of today's financial trends is securitization. Evaluating Securitization risk requires some strong quantitative skills and a deep understanding of both credit and market risk. For international securitization programs it is mandatory to take into account the exchange-rates-related risks. We will see the different methods to evaluate extreme variations of the exchange rates using the Extreme Value Theory and Monte Carlo simulations.

Keywords: Extreme Value Theory, Exchange rates, Block Maxima, Peaks-over-Threshold, Securitization

Acknowledgements

I would like to thank every single colleague that helped me writing this master thesis. I would especially like to show my gratitude to Samir Kasdi that supervised and strongly contributed in the elaboration of each model and in my training during the entire time of my master thesis, Tiphaine Le Ludec that was also very involved in the research. I would also like to thank Taieb Eliraqui, Pierre-Edouard Bienvenu and Ali El Hamidi for receiving me into their team and managing and validating our work. I would also like to express mu gratitude to Henrik Hult, my supervisor at the Royal Institute of Technology, for his availability, guidance and ideas.

Paris, March 2017

Tarek Barbouche

Contents

Introduction

Risk management is a very wide and complex field that is at the core of every financial activity. The importance of this area of expertise is emphasized by the regulatory framework which is currently at the center of many discussions and negotiations. As a matter of fact, the Basel committee is shortly going to publish the new amendments that will constitute Basel 4 and that will strongly impact all financial institutions and will without a doubt impact the European economy. The complexity of risk management relies on the combination of very advanced quantitative methods that should always be combined to a deep qualitative understanding of the market and the notions that are hidden behind every computation.
Risk analysis is based on estimating extreme losses and covering all kind of risks. In the following case, we will focus on managing the exchange rate in the quantitative risk analysis of securitizations (considered as a very specific credit risk that needs to have its own rating methodology).
In this thesis we will try to analyse how the extreme value theory can be applied in order to find extreme scenarios for exchange rates. For foreign transactions we want to find a stressed exchange rate in order to cover the exchange rates risks.

Chapter 1

Context and goals

1.1 Preliminaries

As part of my Master Thesis, I integrated a Rating Methodology Team in a french Investment Bank. This team is responsible for all the Credit Risk internal methodologies of the Bank. These methodologies require having a strong quantitative and qualitative knowledge of the Credit Risk. We will focus on the methodologies concerning Securitizations. There are several kinds of securitizations in this bank (e.g. Trade Receivables, Auto Loan/Lease, Dealer Floor Plan, Consumer Loan ...). We will focus on the Trade Receivables that represents almost half of the securitization activity and is the only one where the Bank has to take some exchange rates risk. An Investment Bank can take equity risks (or sometimes mezzanine risks) for Trade Receivable Securitization Transactions. These operations are generally off balance-sheet from the seller's standpoint. Thus this part of the portfolio presents specific securitization transactions risks. In the current portfolio, there are several currencies: USD, GBP, AUD, DKK, SEK and NOK. To convert all of them in EUR, the FX rate at the calculation date is stressed with an AA scenario. We are interested in the estimation of an "extreme" 99,985% percentile of the FX rate for several horizons (from one to twelve months), using a 17 year historical data base (from 03/01/1999 to 31/07/2016). The daily FX rates values are directly taken from the website www.quandl.com [2]. The percentile we are trying to estimate is extremely high, that is why we have to be cautious in our choices in modeling the variations of the FX rates.

1.2 Description of the Data

In order to study the variations of the FX-rates we need to find the historical variations. Based on the history of one FX-rate EUR/CUR (where CUR is any other currency), we will distinguish the daily variations and variations over longer periods (1 to 12 months). From now on we will let X_i denote the EUR/CUR FX rate's value for day "i", V_i^n the relative variation over n months at day "i" and V_i the log-return over 1 day at day "i" (here CUR designs any of the considered currencies). We have the following formulas:

$$V_i = \ln\left(\frac{X_{i+1}}{X_i}\right)$$

$$V_i^n = \ln\left(\frac{X_{i-30*n}}{X_i}\right)$$

We took these formulas so that when V_i^n is high this means the domestic currency (EUR) is devalued. That will allow us, once we have modeled V_i^n to take the 99.985% percentile to estimate the worst relative variation of the FX-rate. For V_i, we start at day 2 and we compute all the variations possible with the extracted data. We notice that the variation over a non-working day (week-ends or holidays) is automatically 0. These values are predictable and they may bias the results (if they are not removed that would mean that almost one third of the data would be zeros). On top of that, removing the zeros is being more conservative in the simulations (by simulating 30 days over one month we allow the FX rate to decrease every day even during the non-working days of the month).

1.3 Main goal

We want to find an extreme percentile for the V_i^n so we will try to fit classical statistical models and to apply the Extreme Values Theory. We will also try to fit classical statistical models to the one day variations and then generate randomly millions of days to create variations over n months and then take an empirical percentile over a sufficient amount of data. Once we have a 0.99985 percentile that we will call Q. The stress factor is given by 1-Q.

Chapter 2

Mathematical Background

2.1 Classical distributions

The distributions detailed in this part are distribution that have been tested to fit the daily variations of the FX-rates.

2.1.1 Normal Distribution

It is the most common probability distribution. It is entirely determined by its mean and standard deviation. It is often used to model financial random variables but it has a very light tail so it is not exactly the best to model extreme values. It has the following PDF :

$$\frac{1}{\sqrt{2\sigma^2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

And the following CDF:

$$\frac{1}{2}\left[1 + erf\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)\right]$$

Where erf is is the "error function":

$$erf(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2}\,\mathrm{d}t$$

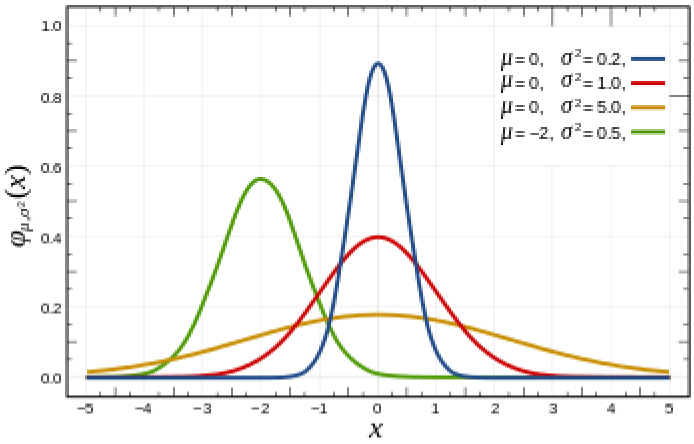

Figure 2.1: Plot of Normal PDF

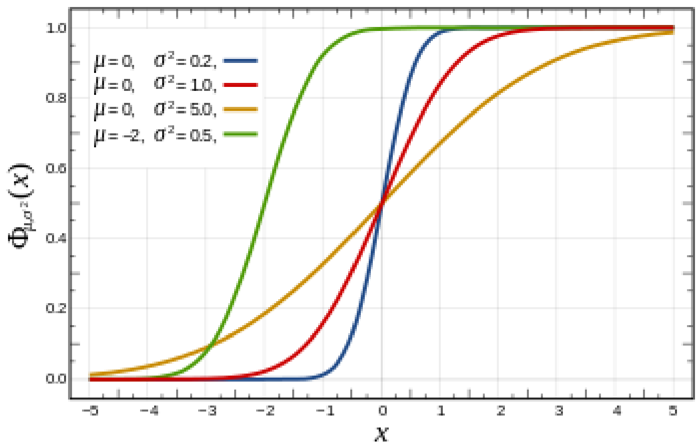

Figure 2.2: Plot of Normal CDF

2.1.2 Log-normal Distribution

It is a continuous probability distribution of a random variable whose logarithm is normally distributed. So if X is log-normally distributed then $\ln(X)$ has a normal distribution of mean μ and standard deviation σ.
It has the following PDF:

$$\frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{\ln(x-\mu)^2}{2\sigma^2}\right)$$

And the following CDF:

$$\frac{1}{2}\left[1 + erf\left(\frac{\ln(x-\mu)}{\sqrt{2}\sigma}\right)\right]$$

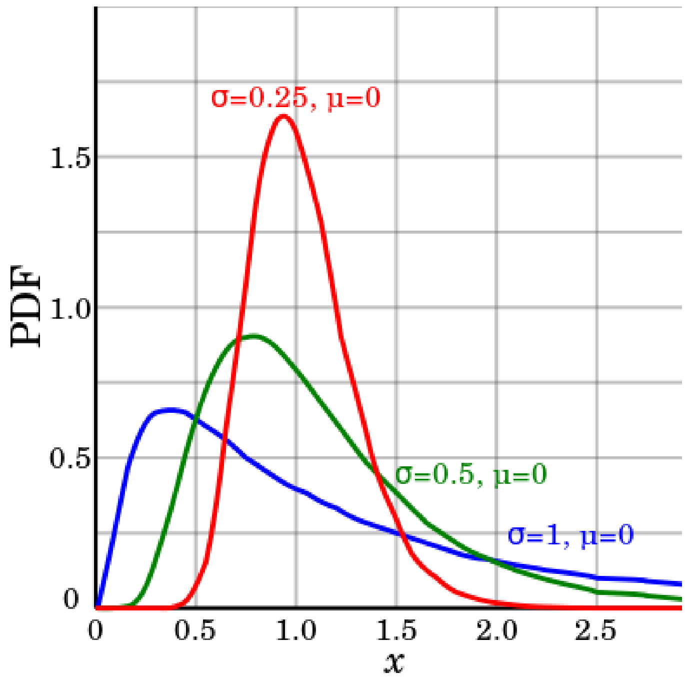

Figure 2.3: Plot of Log Normal PDF

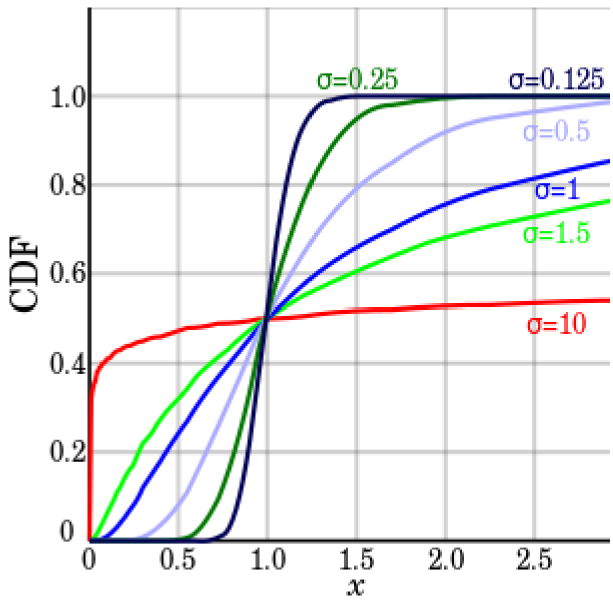

Figure 2.4: Plot of Log Normal CDF

2.1.3 Laplace Distribution

It is sometimes called the double exponential distribution[1]. It is determined by the location parameter μ and the scale parameter $b > 0$. It has the

[1] [9] p.19

following PDF:

$$\frac{1}{2b} exp \left(- \frac{|x - \mu|}{b} \right)$$

And the following CDF:

$$\begin{cases} \frac{1}{2} exp \left(\frac{x-\mu}{b} \right) & if x < \mu \\ 1 - \frac{1}{2} exp \left(- \frac{x-\mu}{b} \right) & if x \geq \mu \end{cases}$$

[7]

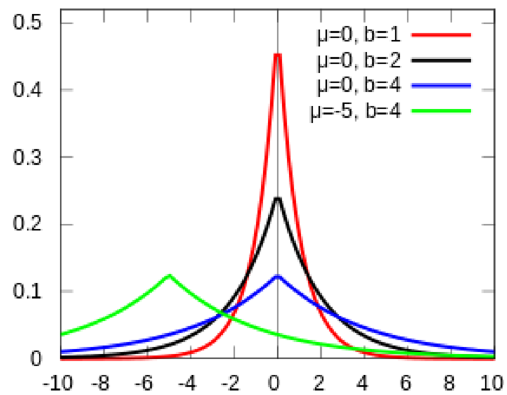

Figure 2.5: Plot of Laplace PDF

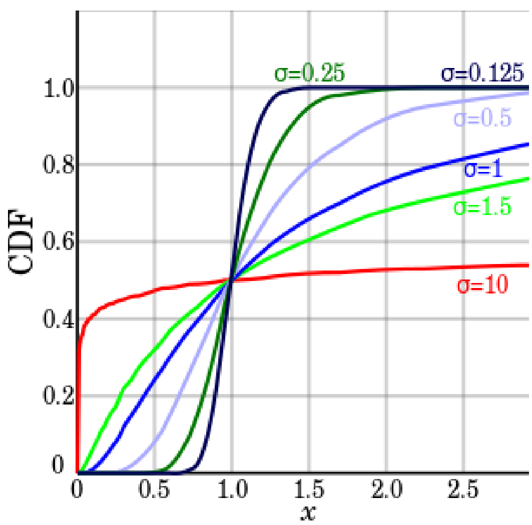

Figure 2.6: Plot of Laplace CDF

2.1.4 Student t Distribution

The larger the sample size gets the closer to the normal distribution the Student's t distribution gets. The t-distribution is symmetric and bell-shaped but has heavier tails than the normal distribution. That is why it is interesting to take it into account in our studies because it is more likely to producing values that fall far from its mean. The standard Student's t is only determined by $\nu > 0$ the number of degrees of freedom but we define non-standardized Student's t-distributions by: The Standard Student's t has the following PDF:

$$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

And the following CDF:

$$F(t) = 1 - \frac{1}{2} I_{x(t)}\left(\frac{\nu}{2}, \frac{1}{2}\right)$$

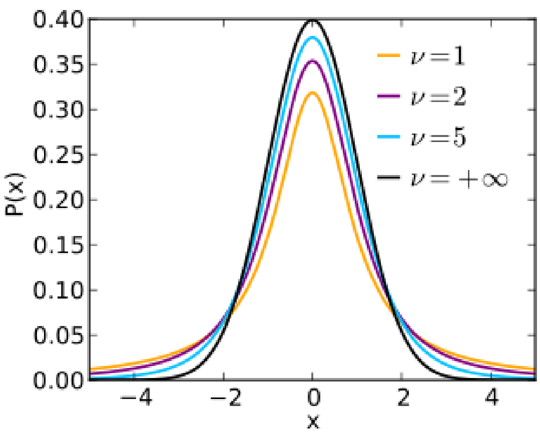

Figure 2.7: Plot of Student PDF

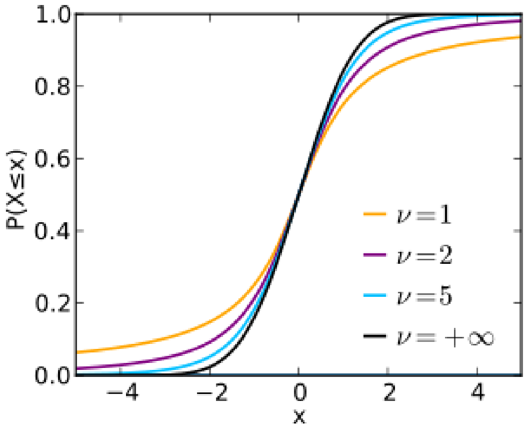

Figure 2.8: Plot of Student CDF

2.1.5 Cauchy Distribution

The Cauchy distribution is particular since it does not have finite moments since it has no moment generating function. It is nonetheless a stable distribution and has probability density function that can easily be expressed analytically. It is determined by its location x_0 and its scale $\gamma > 0$. It has the following PDF:

$$\frac{1}{\pi\gamma\left[1 + \left(\frac{x - x_0}{\gamma}\right)^2\right]}$$

And the following CDF:

$$\frac{1}{\pi}\arctan\left(\frac{x - x_0}{\gamma}\right) + \frac{1}{2}$$

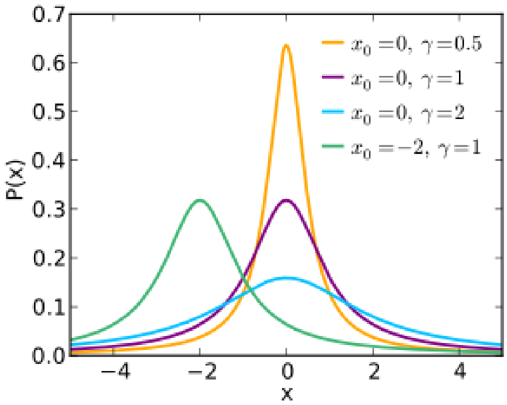

Figure 2.9: Plot of Cauchy PDF

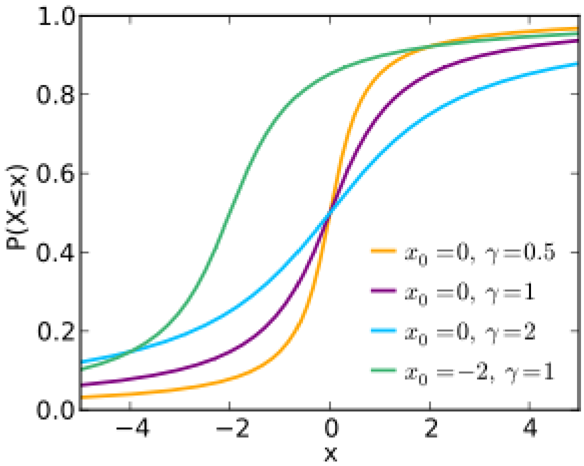

Figure 2.10: Plot of Cauchy CDF

2.1.6 Hyperbolic Distribution

The hyperbolic distribution[2] is a continuous probability distribution characterized by the logarithm of the probability density function being a hyperbola. Thus the distribution decreases exponentially, which is more slowly than the normal distribution (i.e. heavier tails). It is defined by 4 parameters:

- Location parameter: usually denoted by μ
- Parameter of tail: usually denoted by α.
- Asymmetry parameter: generally denoted by β.

[2] [7]

- Scale parameter: generally represented by δ.

These parameters are real: valued, provided that $\alpha > |\beta|$. It has the following PDF:

$$\frac{\gamma}{2\delta\alpha K_1(\gamma\delta)} \exp\left(-\alpha\sqrt{(x-\mu)^2+\delta^2}+\beta(x-\mu)\right)$$

Where K_λ is Bessel's function of second kind and $\gamma = \sqrt{\alpha^2 - \beta^2}$

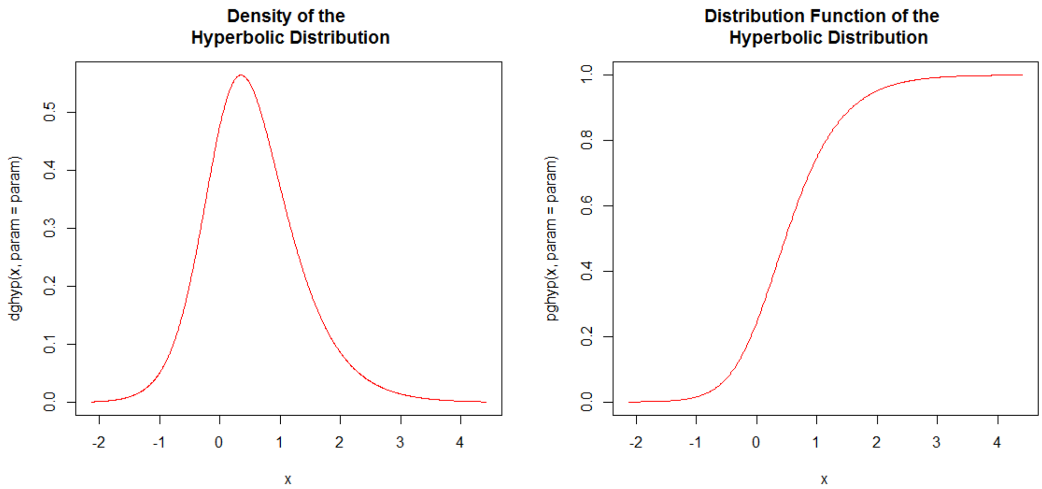

Figure 2.11: Plot of Hyperbolic PDF and CDF

2.2 Goodness-of-fit tests and Information Criteria

2.2.1 Kolmogorov-Smirnov Test

Kolmogorov-Smirnov Test[3] is a non-parametric test of the equality of continuous one-dimensional probability distribution. It quantifies the distance between the empirical distribution function of a sample and the cumulative distribution function of the reference distribution. The statistic that is computed in this test is:

$$D_n = \sup |F_n(x) - F(x)|$$

If this distance is greater than a specific threshold (given by Kolmogorov tables and depending on sample sizes) we will refuse the hypothesis that the two distributions are equal. In the computations we are given a p-value that represents the probability that D_n is smaller than the matching threshold. If the p-value is less than 0.05 we reject the hypothesis that the two distributions are equal. When the p-value is greater than 0.05 we will assume that there is reasonable doubt in admitting the hypothesis. The choice of 0.05 is classical in statistical testing and it is most of the time the convention used is statistical modeling. We can visualize the statistic D_n in the following graph.

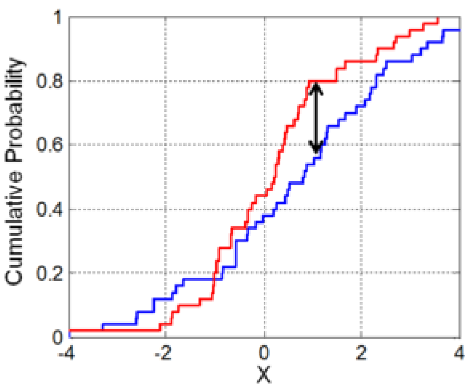

Figure 2.12: Example of the Kolmogorov-Smirnov distance

[3] [14],p.731

2.2.2 Anderson-Darling Test

The Anderson-Darling test[4] is also based on the empirical distribution function. Similarly to the Kolmogorov-Smirnov test it quantifies the distance between the empirical distribution of the data and the theoretical cumulative distribution function that is tested. It is based on the distance

$$A = n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 \omega(x) \, \mathrm{d}F(x)$$

with $\omega(x) = [F(x)(1 - F(x))]^{-1}$ that is a weighing function that places more weight on observations in the tails of the distributions. This is interesting in our case since we are interested in modeling extreme values. The statistic actually computed is an estimation of $A^2 = -N - S$ where :

$$S = \sum_{i=1}^{N} \frac{2i-1}{N} \left(\ln(F(Y_i)) + \ln((1 - F(Y_{N+1-i}))\right)$$

and N is the number of observations in the tested dataset and F is the theoretical cumulative distribution function of the tested distribution. Once this statistic computed the reasoning in order to say if the model is approved is similar to the one described in the Kolmogorov-Smirnov test. A p-value over 0.05 will be accepted.

2.2.3 Cramer Von Mises Test

This test is very similar to Anderson-Darling's test because it is based on the same distance but with a different weighing function. This time we take $\omega(x) = 1$. The distance becomes:

$$\int_{-\infty}^{\infty} (F_n(x) - F(x))^2 \, \mathrm{d}F(x)$$

but the statistic actually computed in the test is:

$$T = n\omega^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left[\frac{2i-1}{2n} - F(x_i) \right]^2$$

Once this statistic computed the reasoning in order to say if the model is approved is similar to the one described in the Kolmogorov-Smirnov test. A p-value over 0.05 will be accepted.

2.2.4 Maximum Likelihood Estimation

Definition of the Likelihood

The likelihood is a criterion that is very commonly taken into account in order to compare a theoretical model with an empirical data set. A likelihood

[4] [14],p.731

function L(a) is the probability or probability density for the occurrence of a sample configuration $x_1, ..., x_n$ given that the probability density $f(x; a)$ with parameter a is known, then $L(a) = f(x_1; a)...f(x_n; a)$.

Parameter estimation

MLE is choosing a vector of parameters which maximize the Likelihood L (or more often ln(L)). This allows us to choose (among all models possible for a given family) the closest one to the empirical data (in terms of PDF and CDF). The optimization method most commonly used is Nelder-Mead's algorithm. It is a non-linear, multi-dimensional method based on simplex theory.

2.2.5 Akaike Information Criterion (AIC)

The Akaike information criterion (AIC) is a measure of the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. The model with the lowest AIC is preferred Let L be the maximum value of the likelihood function for the model; let k be the number of estimated parameters in the model. Then the AIC value of the model is the following :

$$AIC = 2k - 2ln(L)$$

([4], p.268)

2.2.6 Bayesian Information Criterion (BIC)

In statistics, the Bayesian information criterion (BIC) or Schwarz criterion (also SBC, SBIC) is a criterion for model selection among a finite set of models; the model with the lowest BIC is preferred. It is based, in part, on the likelihood function and it is closely related to the Akaike information criterion (AIC). Let L be the maximum value of the likelihood function for the model; let n be the number of observations in the data set. Then the BIC value of the model is the following :

$$BIC = kln(n) - 2ln(L)$$

([4], p.275)

Extreme Value Theory Reasoning

Two important results from Extreme Values Theory are the limit distributions of a series of block maxima (BM) and of excesses over a threshold, called "peaks over threshold" (POT), given that the distributions are non-degenerate and the sample is independent and identically distributed.

2.3 Block Maxima

2.3.1 Main result

The following result is one of the most important theorem in extreme value theory. Let M_n denote the distribution of the maximum of a size n subsample of (X_n) a sequence of i.i.d random variable. Learning more about the M_n distribution would give substantial information about the extreme values of (X_n).

Theorem 2.1 *([8], p.4)*

> *"Let (X_n) be a sequence of i.i.d. random variables. If there exist constants $c_n > 0, dn \in R$ and some non-degenerate distribution function H such that*
>
> $$\frac{M_n - d_n}{c_n} \to H$$
>
> *then H belongs to one of the three standard extreme value distributions:*
>
> $$\begin{array}{lll} \textit{Frechet:} & \Phi_\alpha(x) = \begin{cases} 0 & if x \leq 0 \\ e^{-x^{-\alpha}} & if x > 0 \end{cases} & \alpha > 0, \\ \textit{Weibull :} & \Psi_\alpha(x) = \begin{cases} e^{-(-x)^\alpha} & if x \leq 0 \\ 1 & if x > 0 \end{cases} & \alpha > 0, \\ \textit{Gumbel :} & \Lambda(x) = e^{-e^{-x}}, x \in \mathbf{R}\,'' & \end{array}$$

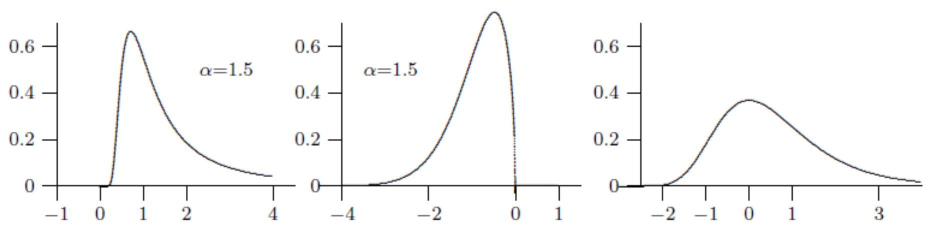

Figure 2.13: Plot of Frechet, Weibul and Gumbel density functions

2.3.2 Interpretation

In practice we will use this results slightly differently. Instead of considering three distinct distribution families we will gather them into one more general family. The main idea is to take several equally sized blocks in the dataset and to keep only one value for each block. The value we will keep is the maximum of the block in order to model extreme values. The size of the blocks is to be determined carefully. As a matter of fact taking blocks that contain a lot of value will leave few observations for modeling the maxima (we divide the number of observations by the size of the blocks). On the contrary taking small blocks will probably lead to value that are not extreme and thus alter the modeling. Once the blocks done we can model the distribution of the maxima by a Generalized Extreme Value (GEV) distribution. The GEV distribution is determined by 3 parameters, the location μ, the scale σ and the shape ξ. We define the function:

$$t(x) = \begin{cases} \left(1 + \left(\frac{x-\mu}{\sigma}\right)\xi\right)^{-\frac{1}{\xi}} & if \xi \neq 0 \\ e^{-\frac{(x-\mu)}{\sigma}} & if \xi = 0 \end{cases}$$

Then the GEV has the PDF:

$$\frac{1}{\sigma} t(x)^{\xi+1} e^{-t(x)}$$

And the CDF:

$$e^{-t(x)}$$

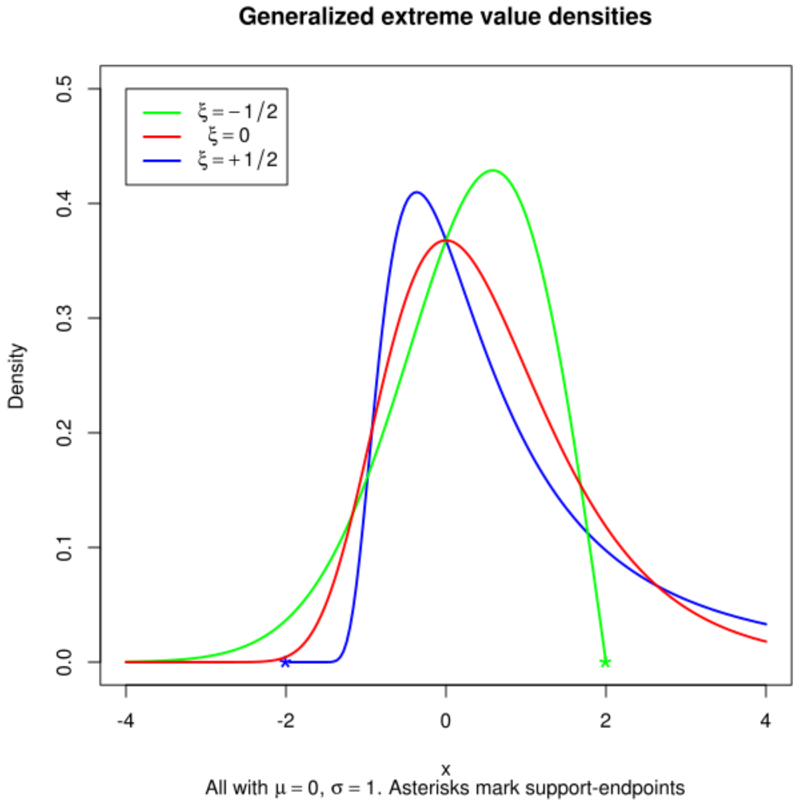

Figure 2.14: Plot of GEV PDF

2.3.3 Choice of the Block Size

The hardest part in this model is to select the block size. As a matter of fact taking blocks that contain a lot of value will leave few observations for modeling the maxima (we divide the number of observations by the size of the blocks). On the contrary taking small blocks will probably lead to value that are not extreme and thus alter the modeling. That is why in general the choice is arbitrary or has a qualitative meaning (in meteorology the blocks often represent months or years).

2.4 Peaks-Over-Threshold

2.4.1 Definition

In this model, we focus on values greater than a threshold.

Definition 2.2 *"Suppose we have a sequence of i.i.d. observations $X_1; ...; X_n$, from an unknown distribution function F. We are interested in excess losses over a high threshold u. Let x_0 be the finite or infinite right endpoint of the distribution F. We define the distribution function of the excesses over the threshold u by*

$$F_u(x) = P\left(X - u \leq x | X > u\right) = \frac{F(x+u) - F(u)}{1 - F(u)}$$

for $0 \leq x < x_0 - u$. $F_u(x)$ is thus the probability that a loss exceeds the threshold u by no more than an amount x, given that the threshold is exceeded."([11], p.2)

Once the threshold chosen, the values above this threshold are generally modeled by a Generalized Pareto Distribution (GPD). It is commonly used to model the tails of another distribution. It has three parameters; the location μ, the scale σ and the shape ξ. We define $z = \frac{x-\mu}{\sigma}$ It has the following PDF:

$$\frac{1}{\sigma}(1 + \xi z)^{-(\frac{1}{\xi}+1)}$$

And the CDF:

$$1 - (1 + \xi z)^{-\frac{1}{\xi}}$$

2.4.2 Choice of the Threshold

The choice of the threshold raises the same questions as the choice of the blocks' size in the BM method. If the threshold is too high there will not be enough values to correctly estimate a model and if the threshold is too low there will be values that could not be considered as extreme.

> "[...]determination of the optimal threshold [...] is in fact related to the optimal determination of the subsamples size" ([10], p.48)

That is why in many cases a fixed threshold is set to the 95% percentile of the empirical distribution that is studied. The Mean Residual Life can also be used.

Definition 2.3 *"Assuming the niteness of the variable mean, i.e. $E[X] < +\infty$, the mean excess function associated to X is defined as $e(u) = E[X - u|X > u]$, that is, the expected exceedance of the threshold u given that exceedance occurs."([5],p.7)*

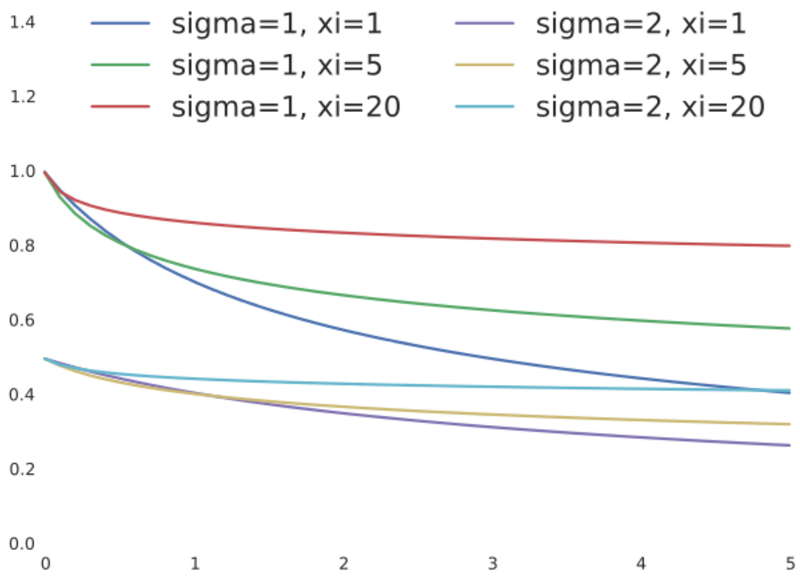

Figure 2.15: Plot of GPD PDF

This mean excess function also known as the mean residual life can be plotted in order to identify the heaviness of the tail and find the right threshold. However sometimes it does not give that much information about the correct threshold to consider.

2.5 Monte-Carlo Simulations

2.5.1 Distribution fitting

The idea is to fit the best distribution among the ones presented in the previous part to a vector of historical data. Here we call the "best distribution" the one that will perform best to the three tests developed in the goodness-of-fit part. So we want the model that has the smallest statistics computed by the test. For each test we will compare the statistic for every model and select the model that minimizes the statistic. This could potentially give us 3 models. If a model is chosen for 2 tests or 3 tests then we will consider that it is the one we will keep. If 3 distinct models are chosen by the tests it is the one chosen by Anderson-Darling that we keep. The reason why we prioritize Anderson-Darling over the others is that it focuses on the tails of the distributions (unlike Cramer Von Mises) and that Kolmogorov-Smirnov is computing a regular distance and not a quadratic distance. We are able to compute these statistics and the p-values thanks to the R package "goftest" and the functions "ad.test, ks.test and cvm.test". So for each model we will estimate the parameters by maximum likelihood thanks to the R package "fitdistrplus" and the function "fitdist" that automatically computes the estimations by maximum likelihood. We use the R function gofstat in order to compute all the statistical test and comparing the values for the 7 models. We created a R function called BestFit that has for input a numerical vector of historical data and gives the best model, the estimated parameters and the 99.985% percentile of the model and the p-values of each test. It also displays the pdf, cdf, qq-plot and pp-plot thanks to the functions "qqcomp, cdfcomp and ppcomp". The idea is to use this function to model the 1 day variations with the purpose of simulating days to deduce as many monthly variations as required to estimate a very precise percentile.

Chapter 3

Securitizations : Rating Methodology

3.1 Trade Receivables securitization

Definition 3.1 *"Securitization is the process of pooling various types of debt – mortgages, car loans, or credit card debt, for example – and packaging that debt as bonds, pass-through securities, or collateralized mortgage obligations (CMOs), which are sold to investors.*

The principal and interest on the debt underlying the security is paid to the investors on a regular basis, though the method varies based on the type of security. Debts backed by mortgages are known as mortgage-backed securities, while those backed by other types of loans are known as asset-backed securities." (Dictionary of Financial Terms. Copyright©2008 Lightbulb Press, Inc.) [3]

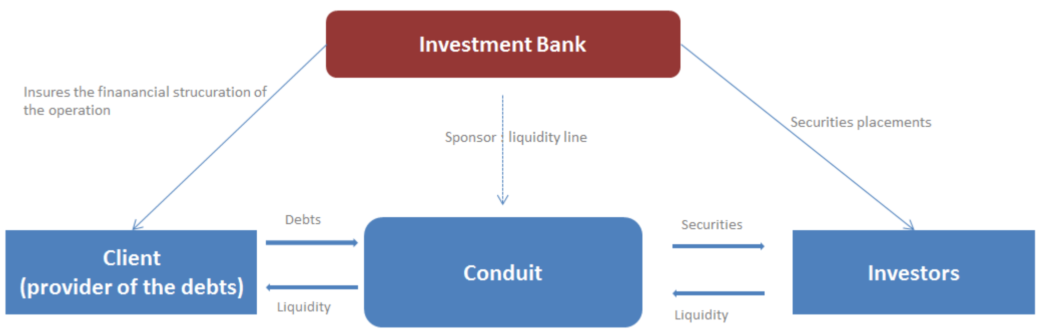

Figure 3.1: Securitization diagram

In our case the investment bank arranges for its core clients securitization transactions backed by trade receivables and funded through a sponsored Asset Backed Commercial Paper (ABCP) conduit on the capital markets. The interest for the client that provides the pool of debts are to diversify his financing sources, reduce the costs of his financing and above all to transfer the risk. In fact, the risk of the debtors defaults is transferred to the Investment Bank in exchange of a premium or a Credit Discount that is usually used as a First Reserve in order to partly cover this risk. For each target rating we have to compute the reserves that the Bank has to put aside in order to cover the sufficient amount of risk for this rating. The sum of all reserves is called the credit enhancement. In general the Bank completes the Credit Discount with a guarantee provided by an insurance company (it is often part of the banking group in order to keep the risk). For example we have the diagram in the next figure. We will now see how the reserves

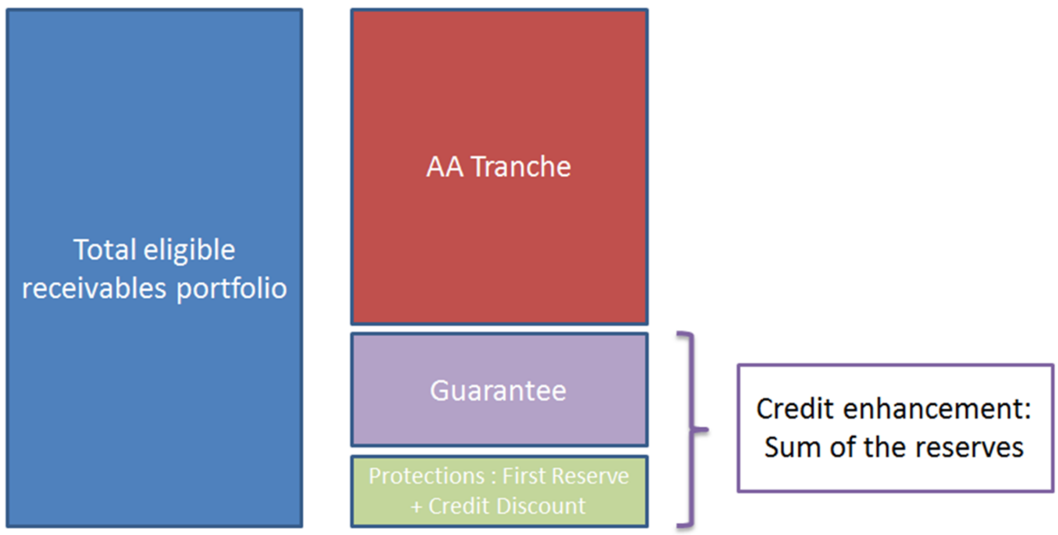

Figure 3.2: Credit enhancement for a AA target rating

are computed in the methodology of some the rating agencies. We keep in mind that the internal methodology of the Investment Bank at stake here is slightly different but mainly based on the agencies' methodologies.

3.2 Methodology for computing the reserves

The credit enhancement is mainly used to know how much the insurance company has to cover. There are four kind of reserves that we will develop :

- Loss reserve

- Dilution reserve
- Yield reserve
- Servicing reserve

We will develop the methodologies of computation of the reserves by the 3 following agencies : Standard & Poors, DBRS and Fitch Ratings. [12] [6] [13]

3.2.1 Loss reserve

This reserve is related to the risk that obligors of underlying receivable pool default (i.e. don't pay within a contractually determined time frame called the "Loss Horizon").

<table>
<tr><th>Fitch</th><th>S&P</th><th>DBRS</th></tr>
<tr>
<td>Loss Reserve = (A x B x C) + D
A = Stress Factor
B = Loss Ratio
C = Loss Horizon Ratio
D = Default Volatility Factor
Benchmark stress factors :
<table>
<tr><th>Target rating</th><th>Stress factor</th></tr>
<tr><td>AAA</td><td>2,5</td></tr>
<tr><td>AA</td><td>2,25</td></tr>
<tr><td>A</td><td>2</td></tr>
<tr><td>BBB</td><td>1,75</td></tr>
<tr><td>BB</td><td>1,35</td></tr>
</table>
</td>
<td>Loss Reserve = (A x B x C)
A = Stress Factor
B = Loss Ratio
C = Loss Horizon Ratio
Note that Loss Reserve calculated per the above formula is then compared to the loss reserve floor per S&P methodology
Benchmark stress factors :
<table>
<tr><th>Target rating</th><th>Stress factor</th></tr>
<tr><td>AAA</td><td>2,5</td></tr>
<tr><td>AA</td><td>2,25</td></tr>
<tr><td>A</td><td>2</td></tr>
<tr><td>BBB</td><td>1,5</td></tr>
</table>
</td>
<td>Idem S&P</td>
</tr>
</table>

Figure 3.3: Benchmarking rating agencies on Loss reserves

3.2.2 Dilution reserve

A dilution is any non-credit-related reduction in the value of a receivable. Dilution usually occurs where: (i) a credit note is issued from the seller to a debtor to compensate for a wrongly billed receivable or (ii) for faulty goods; a discount offer is made to a customer for early payment of invoices; or (iii) a volume discount provided retrospectively to regular customers. Dilution can also arise as a result of set-off and provisional pricing used in some commodities markets. The effects of certain forms of dilution are particularly severe, since once a credit note has been issued against a securitised receivable, the value of the related receivable could potentially be reduced to zero.

Fitch	S&P	DBRS
Dilution Reserve = ((SF x F) +G) x H SF = Stress factor (see above) F = Dilution Ratio G = Dilution Volatility Factor H = Dilution Horizon Ratio	Dilution Reserve = [(SF x F) + ((DS - F) x DS/F)] x H SF = Stress factor (see above) F = Expected dilution DS = Dilution spike H = Dilution horizon ratio It is to be noted that quantifiable items of dilutions are not stressed[1]	Idem Fitch

Figure 3.4: Benchmarking rating agencies on Dilution reserves

3.2.3 Yield reserve

In every trade receivable transaction, a yield reserve is needed to cover interests which are expected to be born over an assumed stressed amortization period. Trade receivables are noninterest-bearing assets. Therefore, the discount applied to the receivables that were purchased by the special-purpose entity (SPE) before amortization is sufficient to cover the increased debt servicing cost throughout the assumed amortization period for a given scenario. This cost includes mainly two risks: increase of interest rates and increase of DSO.

<table>
<tr><th>Fitch</th><th>S&P</th><th>DBRS</th></tr>
<tr>
<td>Yield Reserve =
(1 x 2 x 3) / 360
1 = Annual coupon: a + b + c
a = LIBOR / Euribor
b = Margin
c = Trade Receivable Interest Rate Stress (formula detailed in Fitch methodology stressing the interest rate to be reserved up to the max of (i) 60% * spot Libor and (ii) 3%, for AAA deal, the max of (i) 52% * spot Libor and (ii) 2.6% for AA deal etc…)
2 = Day Sales Outstanding (DSO)
3 = Stress factor (see above)</td>
<td>Yield Reserve =
(1 x 2 x 3) / 360
1 = Annual coupon: a x 1.5 + b
a = LIBOR / Euribor
b = Margin
2 = Day Sales Outstanding (DSO)
3 = Stress factor (not precise, “at least 2” according to documentation)</td>
<td>Identical to S&P but with the following stress factors:
<table>
<tr><th>Target rating</th><th>Stress factor</th></tr>
<tr><td>AAA</td><td>2</td></tr>
<tr><td>AA</td><td>1,75</td></tr>
<tr><td>A</td><td>1,5</td></tr>
<tr><td>BBB</td><td>1,25</td></tr>
</table>
</td>
</tr>
</table>

Figure 3.5: Benchmarking rating agencies on Yield reserves

3.2.4 Servicing reserve

In most transactions the originator acts as the servicer, collecting and administering the receivables. A servicing agreement typically contains a provision that the servicer will apply its customary and usual servicing procedures for securitized receivables in accordance with its own policies and procedures. The servicing reserve aims at covering servicing expenses incurred in the context of default of the initial servicer and of its replacement by a new servicer.

Fitch	S&P	DBRS
Servicing Reserve = (1 x 2 x 3) / 360 1 = Annual Senior Expenses: a + b + c a = Higher of current servicer or back␣up servicer fees b = Trustee fees c = Other fees 2 = Days of Sales Outstanding (DSO) 3 = Stress factor (see above)	Identical to Fitch	Not detailed

Figure 3.6: Benchmarking rating agencies on Servicing reserves

3.3 Stress factor on foreign currencies

The core of this thesis is to manage the exchange rates risks in the computations. The reserves are computed in the currency of the transaction but the insurance company only gives a guarantee in EUR. On top of that the reserve is computed now but has to cover the risk after the maturity of the receivables (in general between 2 and 6 months). The idea is to find a exchange rate that takes into account a catastrophic scenario between the day of the computation and the maturity. To do so we will consider a stress factor which will have to be multiplied to the spot exchange rate at the computation date. The stress factor is necessarily between 0 and 1. The smaller the stress factor is the worse the impact would be on the portfolio (as a matter of fact if the EUR/USD fx-rate is small, the outstanding in USD will be worth less in EUR). For example for a 2 months stress factor of 0.75 would mean that the outstanding we are interested in could lose 25% of its value in 2 months in the worst case scenario only due to currency risk. From now on we will focus on estimating this stress factor.

Chapter 4

Computations and Results

4.1 Descriptions of the R functions

In order to apply the different methods for estimating a stress factor on the exchange rates, I have computed some R functions in order to synthesize all the computations in one main function for each method.

4.1.1 Block Maxima

For the block maxima method we can see in Figure 4.1 the description of what the function BMgev does. The idea is to enter a historical data set and a block size and the important output for the Bank is the stress factor that is computed using the block maxima method. In this case the historical data is the overlapping log-returns on n months. That is to say :

$$V_i^n = \ln\left(\frac{X_{i-30*n}}{X_i}\right)$$

The bloc size is the parameter that will somehow determine the level of extremeness of the study. As described in the Mathematical Background the Block Maxima method requires to form blocks and only keep the maximum of each block. The block size denotes the number of elements we keep in the blocks in this method. In order to simplify the computations we will form equally sized blocks and the blocks are chosen chronologically and not randomly (random blocks were tried but the stress factors obtained were to volatile because of the randomness of the blocks). We could also object that we are considering overlapping log-returns and hence there could be redundancy and correlations in the data but taking non-overlapping date and performing the block maxima method on this basis would lead to severe lack of data in order to find a GEV model.

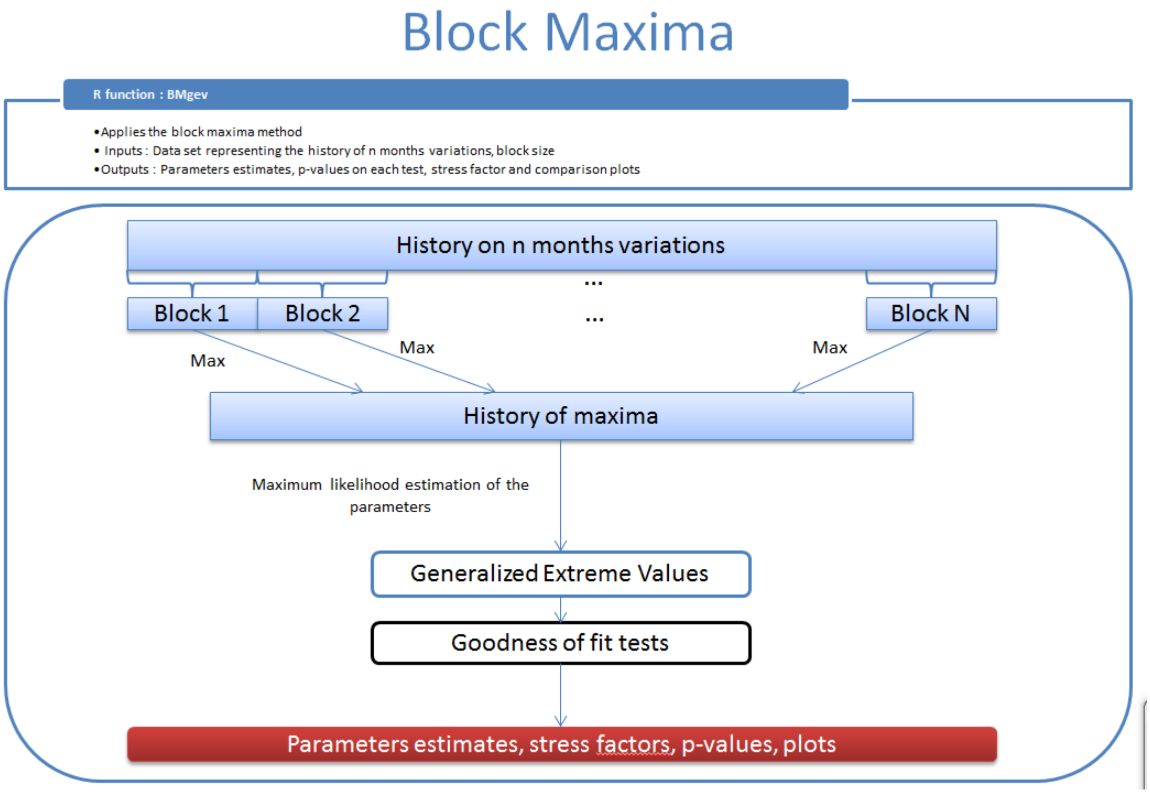

Figure 4.1: Description of the function BMgev

The Figure 4.2 is an example of a computation in R.

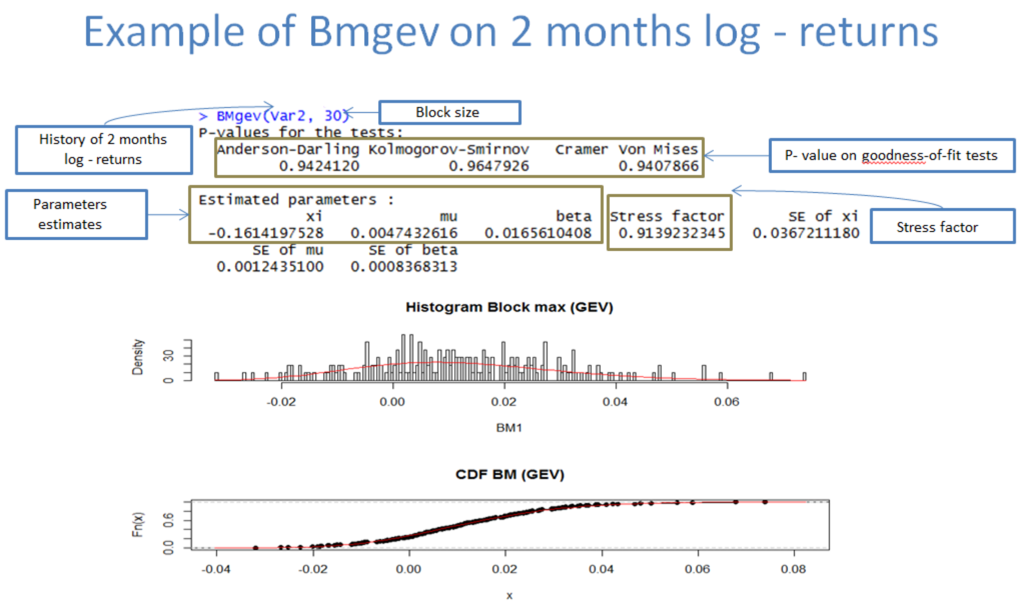

Figure 4.2: Example of BMgev on EUR/USD 2 months log-returns

Choice of the block size

In our project we wanted to see the effect of the block size on the stress factor we wanted to study and on the parameters estimates' standard error of the GEV distribution. We see here that after 30 the outcome starts to lose its consistency. This is probably caused by the lack of data for too big block sizes. That is why we choose to keep 30 as block size in all the following computations.

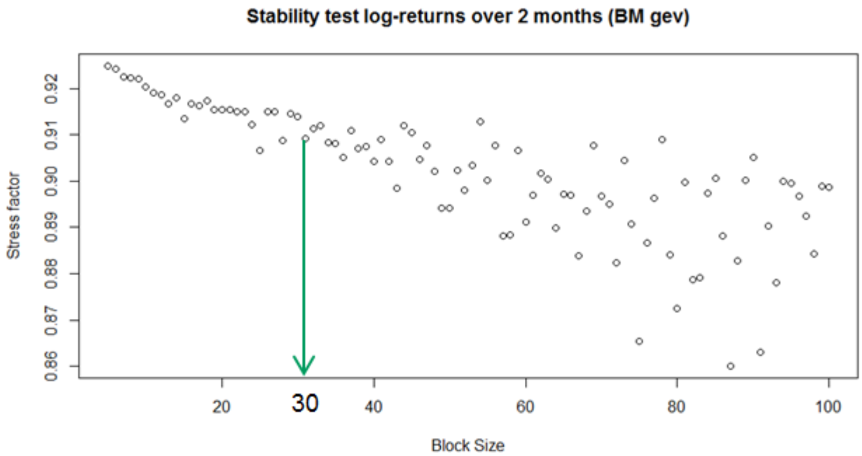

Figure 4.3: Stability of the 2 months stress factor depending on the block size

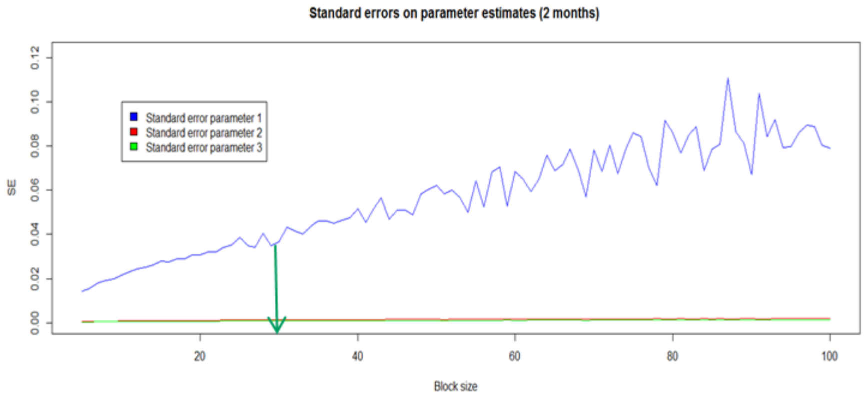

Figure 4.4: Standard error of the parameters estimates depending on the block size

4.1.2 Peaks-Over-Threshold

For the Peaks-Over-Threshold method we can see in Figure 4.3 the description of what the function POT does. The idea is to enter a historical data set and the percentage of data the threshold should eliminate. The important output for the Bank is the stress factor that is computed using the POT method. In this case the historical data is also the overlapping log-returns on n months. That is to say :

$$V_i^n = \ln\left(\frac{X_{i-30*n}}{X_i}\right)$$

The threshold is defined by an empirical quantile of the dataset. For example if the second parameter of the R function is set to 0.85, then we take the 85% empirical percentile as threshold (that is to say that only the higher 15% of the data is kept for modeling the extreme values). The Figure 4.6 is an example of a computation in R.

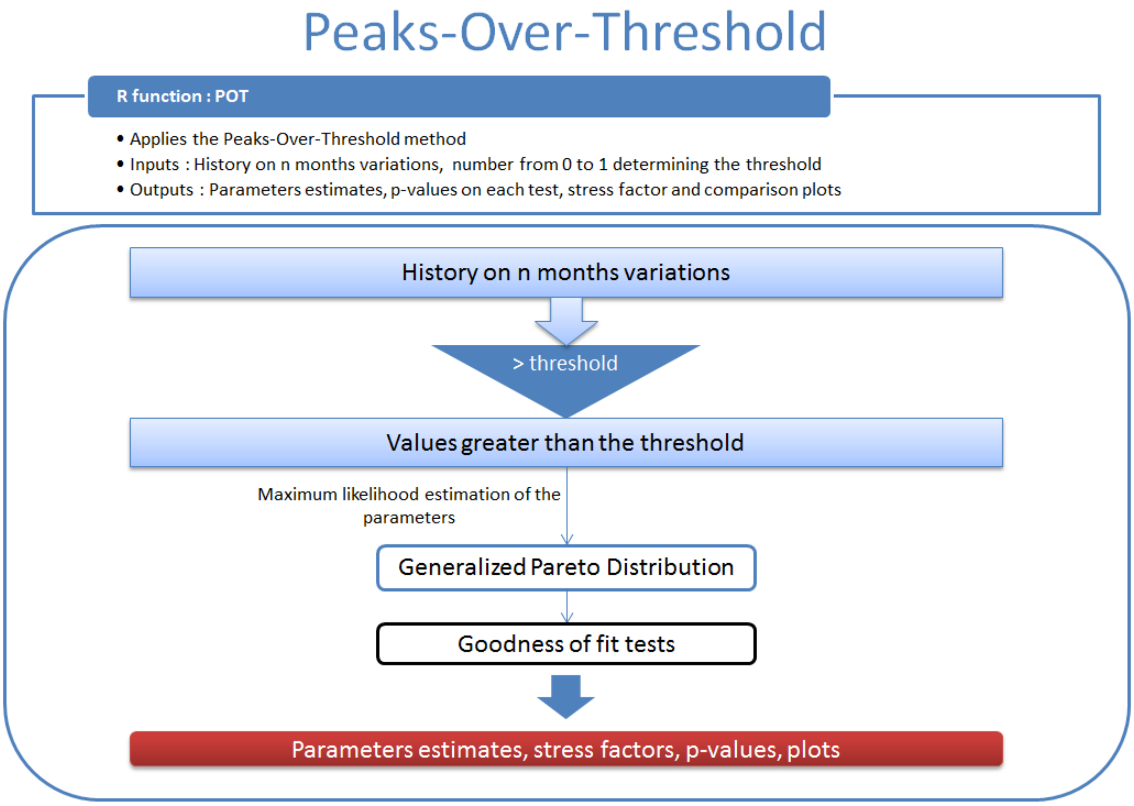

Figure 4.5: Description of the function POT

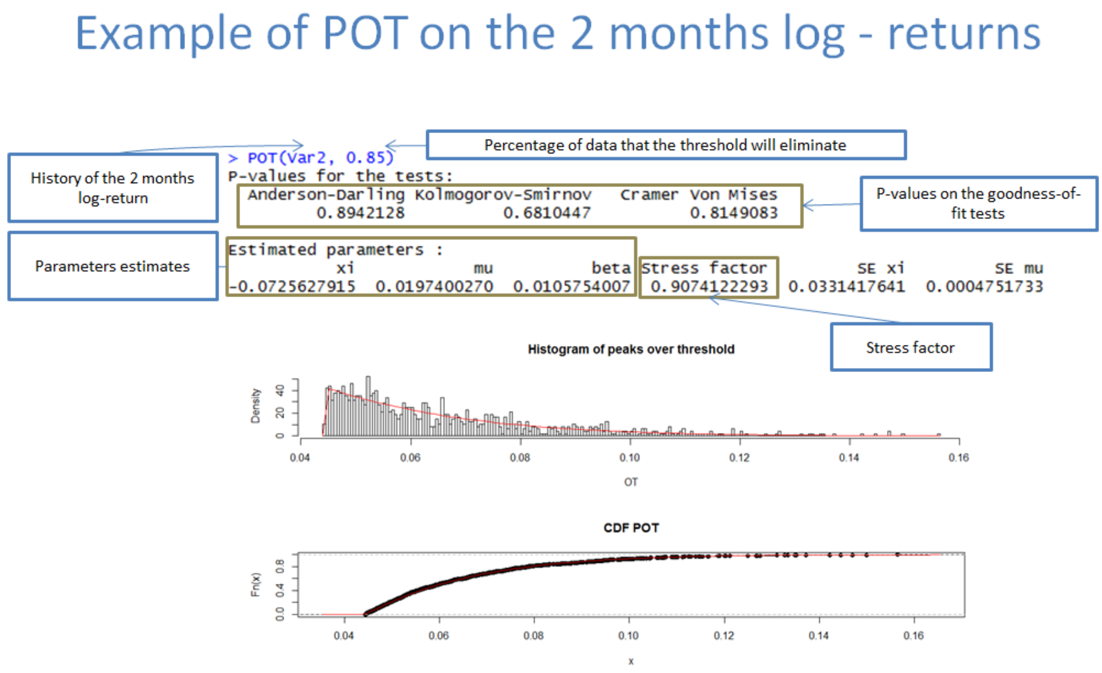

Figure 4.6: Example of POT on EUR/USD 2 months log-returns

Choice of the threshold

We will consider thresholds u in the form quantile(X,q). We use the R function gpdFit to estimate the parameters by maximum likelihood. We slightly modified gpdFit into gpdFitbis that stores the value instead of just displaying it. All this is synthetized in the function POT that we developed that takes X and q for inputs and gives the p-values, the 0.99985 percentile of the model the estimated parameters and the plot of the pdf and cdf. In our case the mean residual life did not give any conclusive explicit value for the threshold. As we can see in the Figure 4.7. In order to choose the threshold we will proceed as in the block sizes choice. We consider the 2 months variations for the EUR/USD and we compute the extreme percentile computed for several threshold of the form quantile(X,q). We take q in [0.50 ; 0.99] (taking more than 50% of the values would mean we are not modeling extreme values).

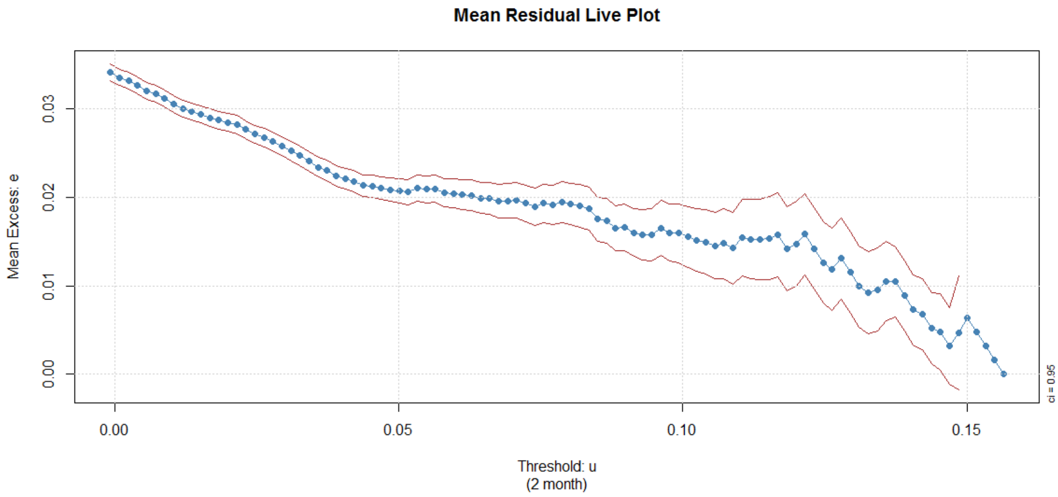

Figure 4.7: Mean residual plot for the 2 months log-returns of EUR/USD

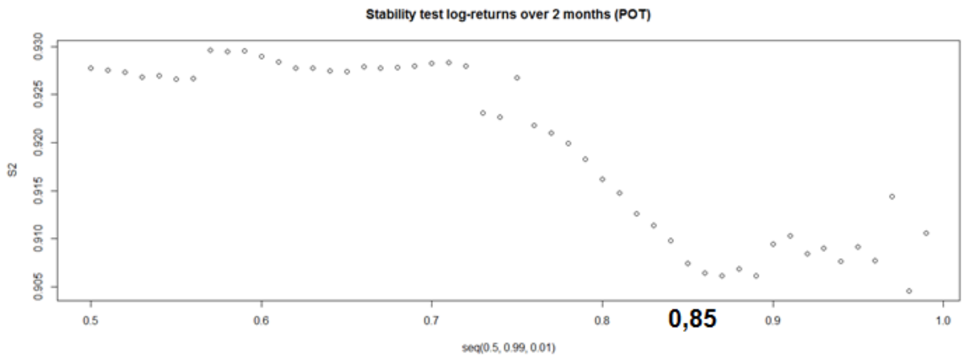

Figure 4.8: Stability of the 2 months EUR/USD stress factor depending on the threshold

We can see that choosing a threshold as the empirical 85% percentile is a logical choice and we will keep that for the computations to come.

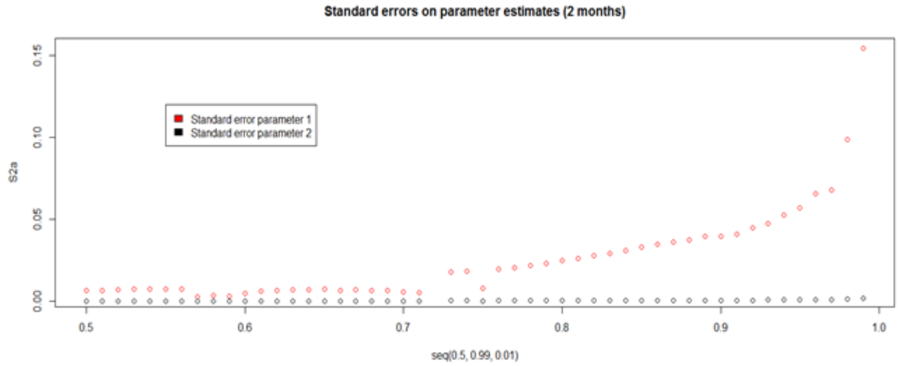

Figure 4.9: Standard error of the 2 months EUR/USD parameters estimates depending on the threshold

4.1.3 Fitting distributions

For the Monte Carlo method the first step is to fit the data on the daily log-returns, we can see in Figure 4.10 the description of what the function BestFit does. The idea is to enter a historical data set and the important output will be the model that fits best to the input data set. The Figure 4.11 is an example of a computation in R.

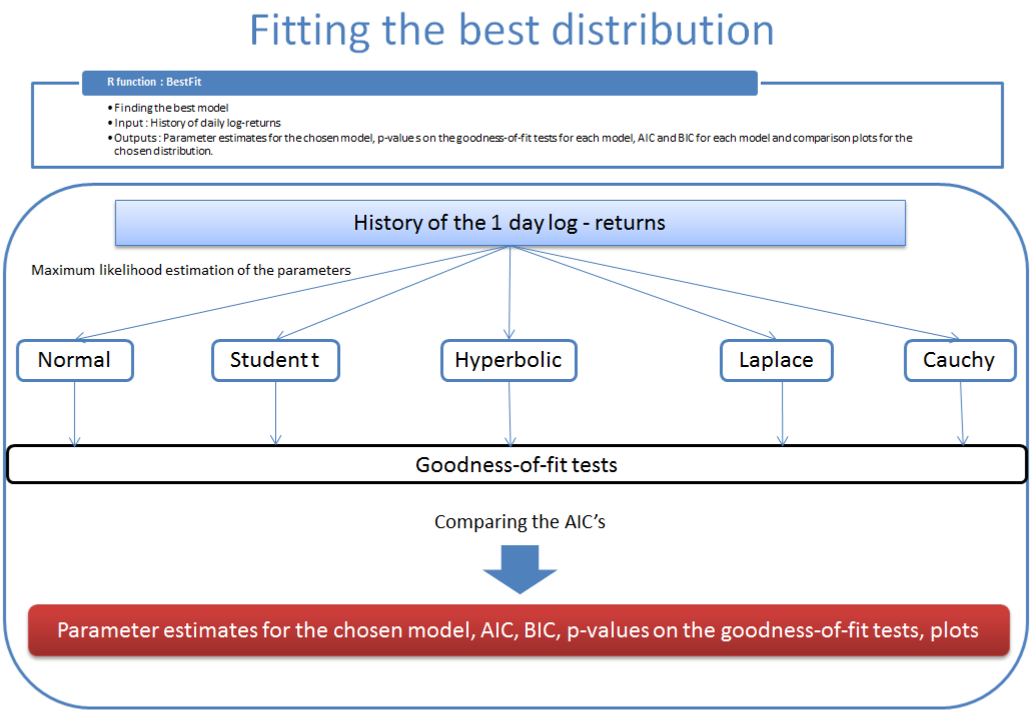

Figure 4.10: Description of the function BestFit

4.1.4 Simulations

Using the BestFit

Once we managed to have a sufficiently good fit for the one day variations we can produce a random walk in order to estimate the variations over n months. In order to have 1 observation of a variation over n months we will consider that we need 30*n observations on 1 day. The stress factor is given by the 0.00015 percentile of the value of :

$$\exp\left(\sum_{1}^{30*n-1} \ln\left(\frac{X_{i+1}}{X_i}\right)\right)$$

The idea is to compute this stress factor a sufficiently large amount of time and then take the 0.00015 percentile. We programmed the R function SIMlog(X,n,T) where n is the number of months over which we want to estimate the stress factor and T is the number of observation we want for the variations over n months. Now we will try to estimate how many observation we need in order to have a steady estimation of the stress factor.

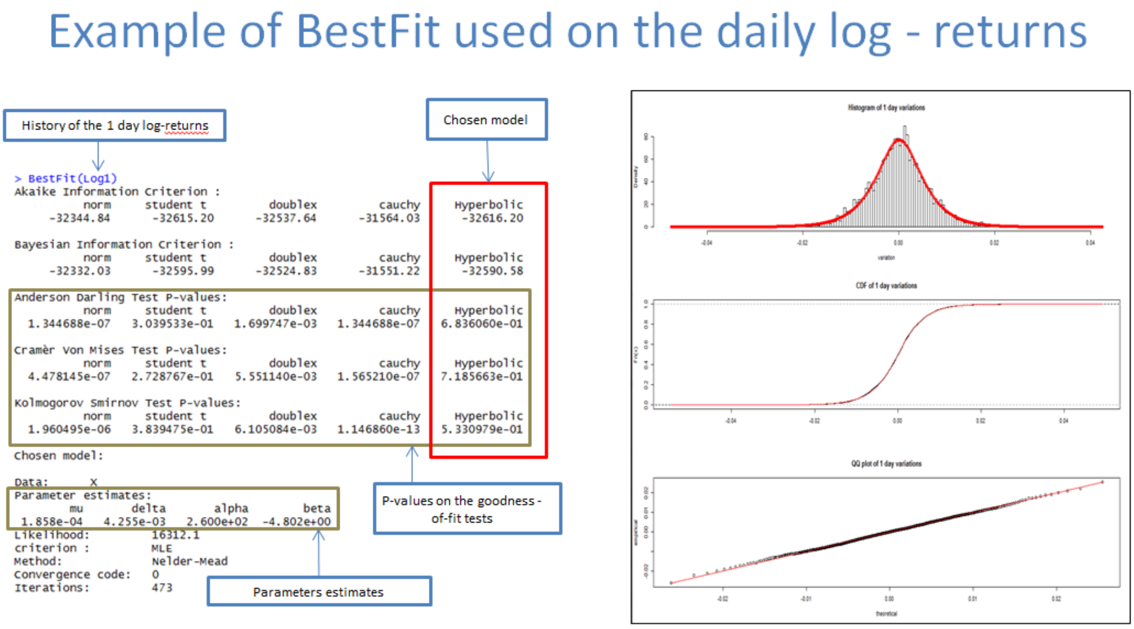

Figure 4.11: Example of BestFit on EUR/USD 1 day log-returns

Using the empirical distribution

We could have exactly the same reasoning but instead of using a theoretical distribution we would use the empirical distribution. The issue with this approach is that we do not allow the daily log return simulated to be greater than the empirical maximum nor being less than the empirical minimum. This could lead to slightly underestimating the variations and it could be hard to justify to the regulators.

Using a body-tail model

We could also use a body-tail model in order to approximate the daily log-returns distribution, this would mean modeling the body of the distribution by the BestFit function (up to a threshold) and then model the tail of the distribution by the Peaks-Over-Threshold model.

4.2 Assumption checking

In the models described earlier we have several assumptions to check before selecting the method that would be better in order to estimate a stress factor on the FX-rates.

4.2.1 Independence of the data

One of these assumptions is the independence of the data that is used. It is essential for the simulations to consider that the 1 day log-returns are independent to one another. In order to check this assumption we will compute the Auto-Correlation Function (ACF) of the historical data set of the daily log-returns.

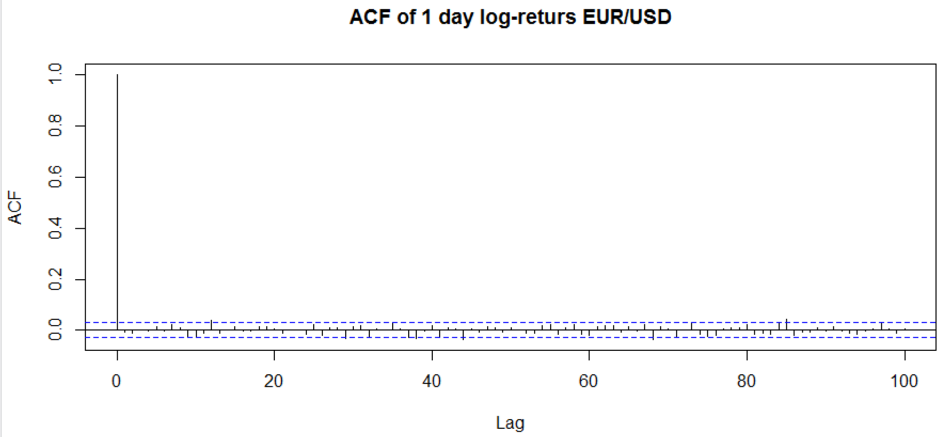

Figure 4.12: ACF of the EUR/USD 1 day log-returns

The ACF have been computed for several different FX-rates' 1 day log-returns and the result is always similar to the Figure 4.12. This allows us to consider that the daily variations of the FX-rates are generally independent. This assumption is also necessary if we want to use the Block Maxima method. However it is very doubtful that the monthly variations of the FX-rates are independent. On one hand, if we compare everyday's FX-rate with the one 30 days later we will have a lot of data but in general consecutive observations are strongly correlated. One the other hand if we keep only one variation per month (i.e. one variation for each month) then we have a significant lack of data in order to use the Block Maxima method.
One solution could be to randomly form the blocks, how ever this makes the stress factor extremely random. Thus, this solution is not acceptable for a Bank in general.
The Peaks-Over-Threshold method does not depend on the order of the data set but having a strong correlation could bias the model.
The two extreme value theory methods describe cannot be properly applied and used in the Bank's methodology so we will focus more about the Monte Carlo method. However in the next section, we will compute the stress factors given by the extreme value theory methods in order to compare them to the stress factors given by the Monte Carlo method.

4.2.2 Robustness of the BestFit function

In order to check that our models are consistent enough and that the identically distributed assumption is correct we will randomly divide the historical 1 day log-returns into a training set and a test set. It is a common statistical process to keep 70% of the data for the training set and fit a model. We will then check that the model can be used to describe the test set. (cf. Figure 4.13)

```
> ad.test(test,null="pghyp",mu=2.722e-04,delta=4.643e-03,alpha=2.654E+02,beta=-7.240)

 Anderson-Darling test of goodness-of-fit
 Null hypothesis: distribution 'pghyp'
 with parameters mu = 0.0002722, delta = 0.004643, alpha = 265.4, beta =
 -7.24

data:  test
An = 0.5465, p-value = 0.6999

> ks.test(test,"pghyp",mu=2.722e-04,delta=4.643e-03,alpha=2.654E+02,beta=-7.240)

 One-sample Kolmogorov-Smirnov test

data:  test
D = 0.0195, p-value = 0.681
alternative hypothesis: two-sided

> cvm.test(test,null="pghyp",mu=2.722e-04,delta=4.643e-03,alpha=2.654E+02,beta=-7.240)

 Cramer-von Mises test of goodness-of-fit
 Null hypothesis: distribution 'pghyp'
 with parameters mu = 0.0002722, delta = 0.004643, alpha = 265.4, beta =
 -7.24

data:  test
omega2 = 0.0824, p-value = 0.6782
```

Figure 4.13: Goodness-of-fit on the test set

There is also an important time dependence when we consider FX-rates, that is why we could object to the fact that we take randomly the training set and the test set. That is why we also tested the robustness of the BestFit function by fitting a distribution with the data as of July 2015 and test that the data from 2016 still fits the model. (cf. figure 5.9 and 5.10)

```
> BestFit(Log1[1:4212])
Akaike Information Criterion :
         norm      student t        doublex        cauchy Gen Hyperbolic
     -30516.30      -30751.92      -30672.36      -29739.23      -30752.15

Bayesian Information Criterion :
         norm      student t        doublex        cauchy Gen Hyperbolic
     -30503.61      -30732.88      -30659.66      -29726.54      -30726.77

Anderson Darling Test P-values:
         norm      student t        doublex        cauchy Gen Hyperbolic
  1.428015e-07   2.708682e-01   1.565382e-03   1.424501e-07   6.806283e-01

Cramèr Von Mises Test P-values:
         norm      student t        doublex        cauchy Gen Hyperbolic
  2.600234e-06   2.399668e-01   5.688518e-03   2.705265e-07   6.979721e-01

Kolmogorov Smirnov Test P-values:
         norm      student t        doublex        cauchy Gen Hyperbolic
  6.104165e-06   3.599648e-01   6.265719e-03   2.431388e-13   5.320538e-01

Chosen model:

Data:      x
Parameter estimates:
        mu       delta       alpha        beta
 2.652e-04   4.619e-03   2.644e+02  -6.819e+00
Likelihood:          15380.07
criterion :          MLE
Method:              Nelder-Mead
Convergence code:    0
Iterations:          347
```

Figure 4.14: BestFit as of 2015

```
> ad.test(Log1[4213:4462],null="pghyp",2.652e-04,4.619e-03,2.644e+02,-6.819)

 Anderson-Darling test of goodness-of-fit
 Null hypothesis: distribution 'pghyp'

data:  Log1[4213:4462]
An = 1.3848, p-value = 0.2065

> ks.test(Log1[4213:4462],"pghyp",2.652e-04,4.619e-03,2.644e+02,-6.819)

 One-sample Kolmogorov-Smirnov test

data:  Log1[4213:4462]
D = 0.0665, p-value = 0.2188
alternative hypothesis: two-sided

> cvm.test(Log1[4213:4462],null="pghyp",2.652e-04,4.619e-03,2.644e+02,-6.819)

 Cramer-von Mises test of goodness-of-fit
 Null hypothesis: distribution 'pghyp'

data:  Log1[4213:4462]
omega2 = 0.2093, p-value = 0.2503
```

Figure 4.15: Goodness-of-fit on 2016

The P-values on the test set are around 0.68, hence the test set can be modeled thanks to the BestFit function applied to the training set.

The P-values on the test set are around 0.22, hence the 2016 data set can be modeled thanks to the BestFit function applied to the 1999 to 2015 data set.

4.3 Comparison of the results

We will now compare the extreme value theory methods with one of the 3 Monte Carlo methods we described earlier (one using the empirical distribution, one using a body-tail hybrid distribution and one using the BestFit function). In order to choose one we will compare the stress factors generated by each method for the EUR/USD and the EUR/BRL exchange rates. Choosing USD and BRL would allow us to see how the model reacts to a quite stable currency (USD) or to a very volatile currency (BRL). On one

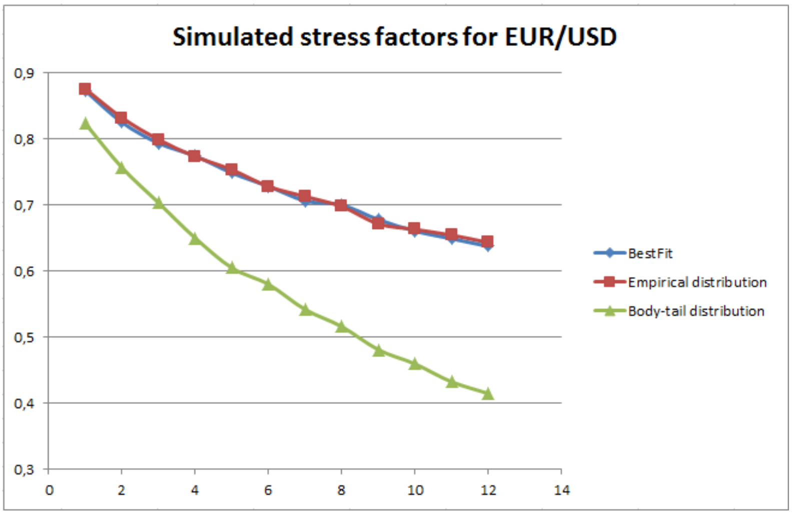

Figure 4.16: Comparision of the simulation methods for EUR/USD

hand, can see that the BestFit method and simulating directly out of the Empirical distribution gives very comparable stress factors. On the other hand the body-tail model is way too strict, it would be impossible to consider that the EUR/USD exchange rate would be divided by 2 in only 8 months. This would lead to over-conservativeness that would be rejected by both the Front Office and the regulatory auditing committees. Thus we can eliminate the body-tail model.

We can see here that the BestFit gives stricter stress factor, but we have to keep in mind the fact that BRL was a highly unstable currency (the EUR/BRL exchange rate once dropped of 10% in only one day). The Empirical distribution gives stress factors not far from the EUR/USD stress factors. It would be unthinkable to consider that the EUR/USD and the EUR/BRL have very close risk profiles. Any regulator would consider that using the Empirical distribution would lead to an underestimation of the exchange rate risk on any volatile currency. That is why we will chose the BestFit model over the two others. We will refer to it as the "Simulations" for all that follows.

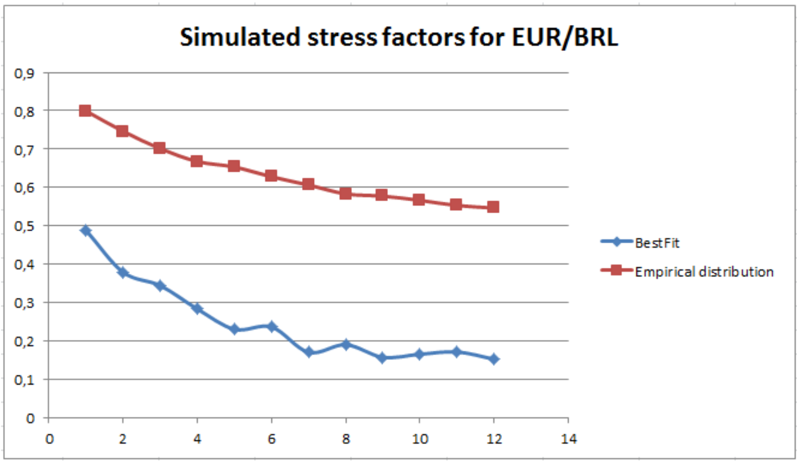

Figure 4.17: Comparision of the simulation methods for EUR/BRL

We have computed the stress factor given by each method for each maturity from 1 to 12 months.The stress factors are necessarily between 0 and 1. The smaller it is the more conservative we are. We will now display the stress factors given by the different method for the following exchange rates : EUR/USD, EUR/SEK and EUR/BRL (Figure 4.18 to 4.20). The dots in red represent the stress factors that were denied by the statistical tests. The simulations give results that are quite intuitive. In fact the longer the period is the higher the variations can be. We can conclude not only that the Monte Carlo method is the most stable and the one that shows the fewer flaws but on top of that it is the method that provides the most conservative results. The validity of the stress factor only depends on the fitting of the historical data whereas in the other method the choice of the threshold and of the block sizes matter a lot and it is not always possible to fit the extreme values. Moreover the dependence of the monthly variations is a major problem in these methods.

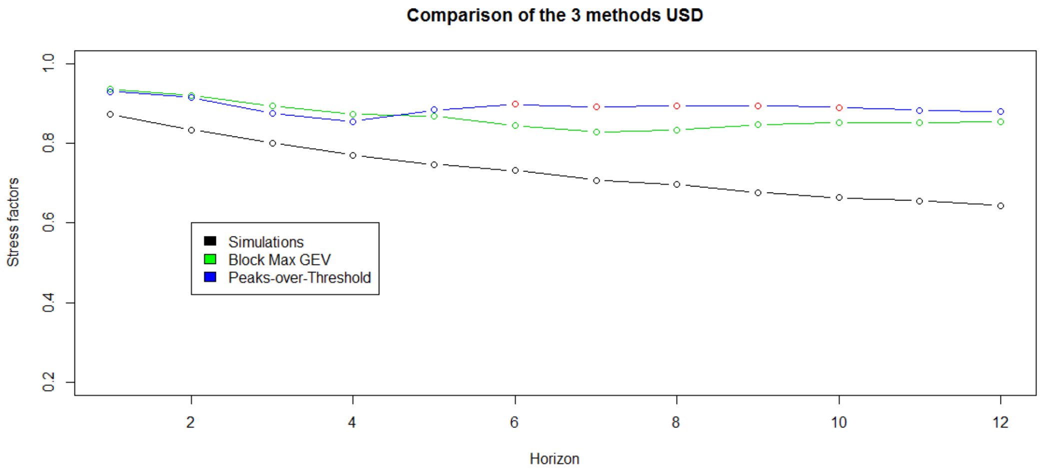

Figure 4.18: Comparision of the methods for EUR/USD

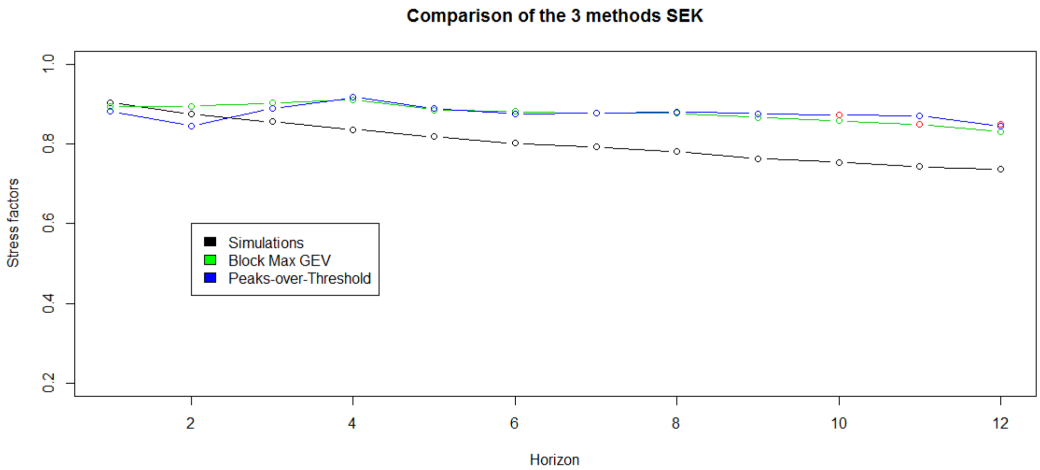

Figure 4.19: Comparision of the methods for EUR/SEK

4.4 Focus on the simulations

4.4.1 Sufficient amount of data

We also had to check if taking a historical data set from 1999 to 2016 was enough or if we had to dig deeper data and try to extrapolate the EUR data from previous European currencies such as the French Franc or the Deutsche Mark. In order to extrapolate the value of the EUR before 1999

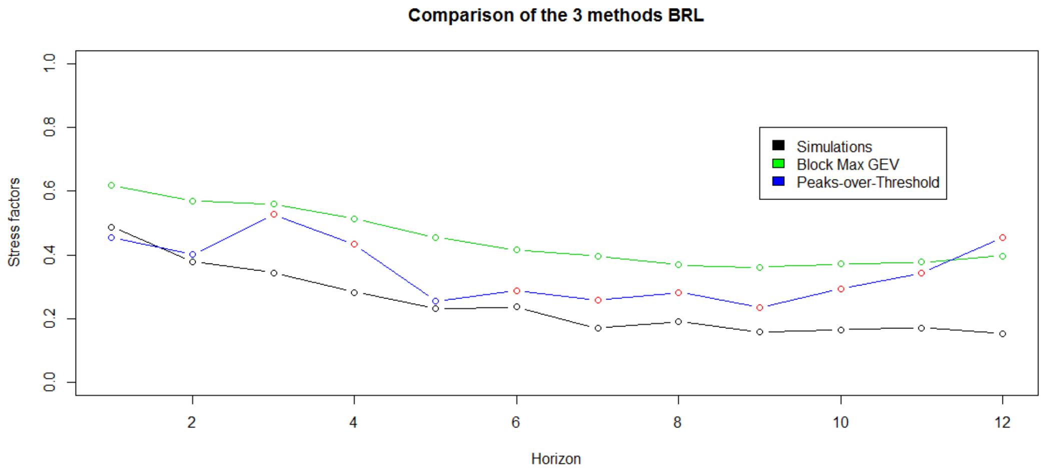

Figure 4.20: Comparision of the methods for EUR/BRL

we take the data from the Website of the Bank of England. The method they used is detailed in the following link [1] This allows us to have EUR/USD daily exchange rates from 1975 to 2016. We compare the outcome of the simulations with this new data set in Figure 4.21. We can see that the stress factors are very close obtained with the two different data history are very close and that the one obtained with the FX-rates since 1999 is slightly more conservative. This shows us that the results are consistent with an estimation of a deeper data history.

4.4.2 Comparing the outcome for several currencies

We computed the stress factors for each currency for horizons from 1 to 12 months. We display it in the Figure 4.22. We can see that all simulations have the same behavior. We also notice that this model allows very high stress factors for extremely volatile FX-rates such as BRL or RUB.

[1] [1], http://www.bankofengland.co.uk/statistics/pages/iadb/notesiadb/Spot_rates.aspx

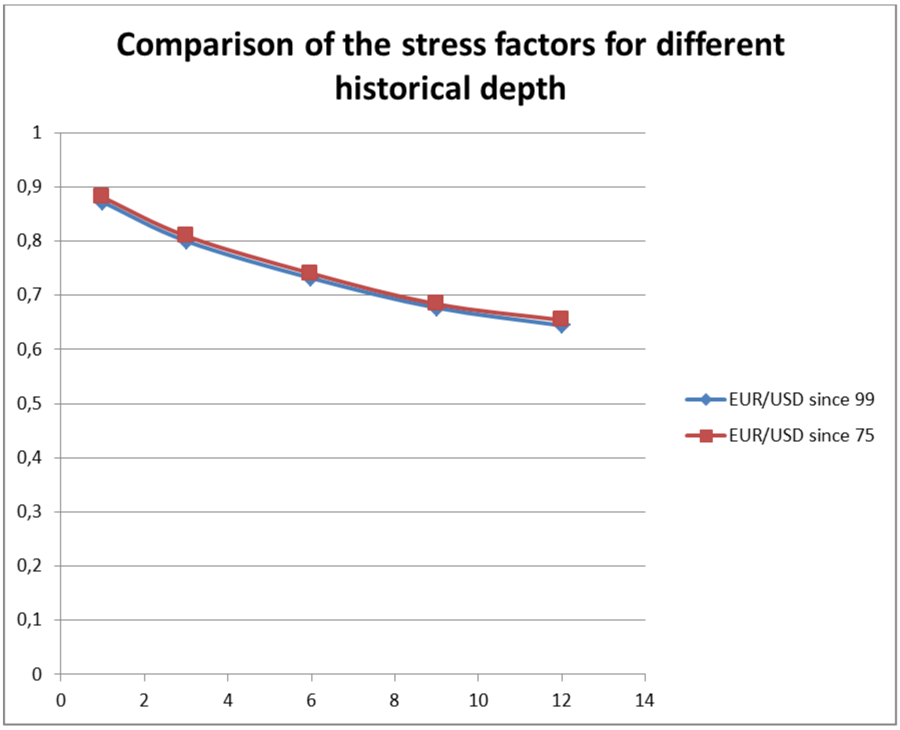

Figure 4.21: Comparision with a deeper history dataset

4.5 Potential methodology improvements

The issue tackled in this document is the problem of univariate modeling of extreme exchange rates risk scenarios. Nonetheless, several currencies depend on each other. If we take the example of the EUR/DKK exchange rate, it is almost constant since the creation of the EUR. This means that the EUR and the DKK values are very strongly correlated. The DKK case is only one example among several others. So the greatest complement to this study would be searching for a Multivariate model taking into account all the correlation effect between all the currencies in the portfolio. Exploring copula-based methods could lead to very promising results. Taking a time series approach could also be worth trying, in the cas of EUR related exchange rate the historical data is not deep enough to observe seasonality on monthly variations, and as far as daily variations are concerned we showed previously that it could be considered as a stationary noise.

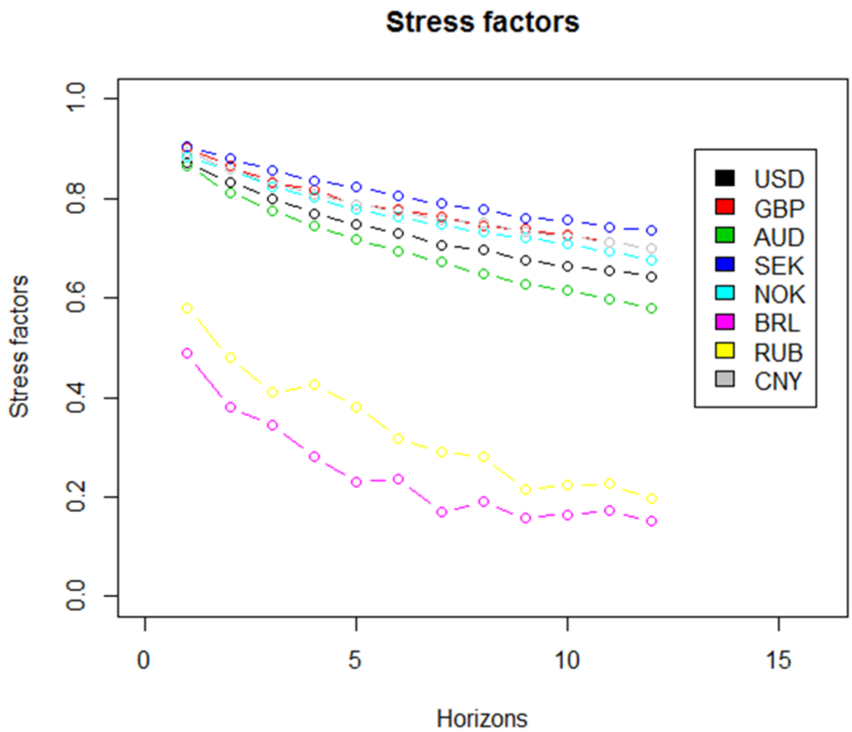

Figure 4.22: Comparision with several currencies

Conclusion

To conclude, we can see that Extreme Value Theory methods require a lot of data since they are based on suppressing a very big part of the initial dataset. For the POT method we only keep 15% of the data and for the Block Maxima method the dataset size is divided by 30. This can explain the difficulties to obtain statistical models that are accepted by the statistical tests for these methods. The Monte-Carlo method is more reliable since it only relies on the fitting of the largest dataset that we have (the daily log-returns of the FX-rate). The random simulations give more conservative stress factors than the EVT methods. That is why in order not to underestimate the risk the Bank will rather use the Monte-Carlo method that is also more consistent.
We can conclude that EVT is a very powerful tool in the modern statistics and could be very important in quantitative risk analysis, but they require a huge amount of data history and their performance to the statistical tests is not guaranteed.

Bibliography

[1] Bank of england's website. for EUR extrapolation until 1970 URL : http://www.bankofengland.co.uk/statistics/pages/iadb/notesiadb/Spot_rates.aspx.

[2] Quandl. https://www.quandl.com/ for the exchange rates values.

[3] Dictionary of financial terms, 2008.

[4] Kenneth P. Burnham and David R. Anderson. Multimodel inference: Understanding aic and bic in model selection, 2004.

[5] Ana C. Cebrian, Michel Denuit, and Philippe Lambert. Generalized pareto fit to the society of actuaries large claims database, 2003.

[6] DBRS. Methodology u.s. trade receivables, 2009.

[7] E. Eberlein, U. Keller, and Bernoulli. Hyperbolic distributions in finance, 1995.

[8] M. Gilli and E. Kellezi. An application of extreme value theory for measuring financial risk, 2006.

[9] Samuel Kotz, Tomasz J. Kuzubowski, and Krysztof Podgrorski. The laplace distribution and generalizations a revisit with new applications, 2001.

[10] Yannick Malevergne and Didier Sornette. Extreme financial risks from dependence to risk management, 2006.

[11] Alexander J. Mc Neil and Thomas Saladin. The peaks over thresholds method for estimating high quantiles of loss distributions, 1997.

[12] Standard & Poor's. Trade receivable criteria, 2010.

[13] Fitch Ratings. Global rating criteria for trade receivables securitisations, 2012.

[14] M. A. Stephens. Edf statistics for goodness of fit and some comparisons. journal of the american statistical association, vol. 69, no. 347 pp.730-737, 1974. URL : www.jstor.org/stable/2286009.