Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

An efficient Peak-over-Threshold implementation for operational risk capital computation

Dominique GUEGAN, Bertrand HASSANI, Cédric NAUD

2010.96

Maison des Sciences Économiques, 106-112 boulevard de L'Hôpital, 75647 Paris Cedex 13
http://centredeconomiesorbonne.univ-paris1.fr/bandeau-haut/documents-de-travail/
ISSN : 1955-611X

An efficient Peak-over-Threshold implementation for operational risk capital computation.

November 24, 2010

Authors

- Dominique Guégan: Paris School of Economics, CES-MSE, Université Paris 1 Panthéon-Sorbonne, 106 boulevard de l'Hopital 75647 Paris Cedex 13, France, e-mail: dguegan@univ-paris1.fr
- Bertrand K. Hassani: BPCE and CES-MSE, Université Paris 1 Panthéon-Sorbonne, 106 boulevard de l'Hopital 75647 Paris Cedex 13, France, phone: +33 (0) 158403397, e-mail: bertrand.hassani@bpce.fr
- Cédric Naud: BPCE, 50 avenue Pierre Mendès-France 75013 Paris, France, phone: + 33 (0) 158404749, e-mail: cedric.naud@bpce.fr

Abstract

Operational risk quantification requires dealing with data sets which often present extreme values which have a tremendous impact on capital computations (VaR). In order to take into account these effects we use extreme value distributions to model the tail of the loss distribution function. We focus on the Generalized Pareto Distribution (GPD) and use an extension of the Peak-over-threshold method to estimate the threshold above which the GPD is fitted. This one will be approximated using a Bootstrap method and the EM algorithm is used to estimate the parameters of the distribution fitted below the threshold. We show the impact of the estimation procedure on the computation of the capital requirement - through the VaR - considering other estimation methods used in extreme value theory. Our work points also the importance of the building's choice of the information set by the regulators to compute the capital requirement and we exhibit some incoherence with the actual rules.

Keywords: Operational risk - Generalized Pareto distribution - Picklands estimate - Hill estimate - Expectation Maximization algorithm - Monte Carlo simulations - VaR.

1 Introduction

The purpose of this paper is operational risks quantification following the Basel Advanced Measurement Approach (BIS (2004)). One proposal of the Basel committee binds banks to carry out their own models on internal data sets to evaluate amounts of capital necessary to face these risks.

In this paper, we focus on the computation of the loss distribution function (LDF) (Frachot et al. (2001)), (Guegan and Hassani (2009)) which is currently used to evaluate this kind of risks (Cruz (2004), Chernobai et al. (2007)). We compute it as a convolution of two distribution functions, one modelling the frequencies of the losses and the other one their severity. We restrict to a Poisson distribution for the frequencies, and we focus on the estimation of the severity distribution. As soon as all the fittings of classical distributions fail (when we use a Kolmogorov - Smirnov test), we decide to model it as a mixing of two distributions: one fitted on the most important losses using the Generalized Pareto Distribution (GPD) for which we need to estimate a threshold above which we fit the distribution, (Pickands (1975), Davis and Resnick (1988) and Embrechts et al. (1997)), and another one on the remaining data considering a lognormal distribution. In this paper, we deeply discuss the threshold estimation, and we show that the methods used to estimate all the parameters of GDP impact drastically the computation of the VaR measure and then the corresponding capital requirement. This last point leads us to analyse in details the building of the information data sets we used, provided by the bank[1] following Basel Committee requirements. We point that the structure of the information set can be a source of errors for the computation risks associated to the operational risks.

The data set we use is organized into the Basel Matrix (BIS (2001)). In its first level of granularity, this matrix is made up of 56 cases - 8 business lines ("b") $\times$ 7 event types ("e")[2]. Nevertheless, each event type might be decomposed in several elements. For example, the "external fraud" event may be shared in two items - "Theft and Fraud" and "Systems Security"

[1] We used data from the French Caisse d'Epargne perimeter.

[2] The business lines are corporate finance, trading & sales, retail banking, commercial banking, payment and settlement, agency services, asset management and retail brokerage. The event types are internal fraud, external fraud, employment practices & workplace safety, clients, products & business practices, damage to physical assets, business disruption & system failures and execution, delivery & process management.

(second level of granularity). In a third level, the element "Theft and Fraud" may be splitted in several components: "Theft/Robbery", "Forgery" and "Check kiting". After a deep analysis, we observe that the kind of losses expected from a fraud with a credit card does not correspond to losses caused by someone hacking the system for instance; nevertheless they are in the same case. Therefore, considering the largest level of granularity, we could face multimodal empirical distributions. Consequently, the methods used to model the losses depend on the granularity level choice. This choice might have a tremendous impact on capital requirement computations. Besides, we face a trade-off between quantity of data and robustness of the estimations: indeed, if the quantity of data is not sufficient, we cannot go lower in the granularity; on another hand the ensuing empirical distribution is therefore an aggregate of various natures of data and the estimation of this last empirical distribution can be source of unusable results.

The methodology used to build the Loss Distribution Function (LDF) is very simple and efficient. We recall its representation. The LDF $G_{b,e}$ (whose density is denoted $g_{b,e}$) is a mixture of two distributions, the frequency distribution $p_{b,e}$ and the severity distribution $F_{b,e}$ (whose density is $f_{b,e}$):

$$G_{b,e} = \sum_{\gamma=0}^{\infty} p_{b,e}(\gamma;\bullet) F_{b,e}^{\otimes\gamma}(x;\bullet),\ x > 0, \tag{1.1}$$

with

$$G_{b,e} = 0,\ x = 0,$$

where $\otimes\gamma$ is the γ-order operator of convolution, and

$$g_{b,e} = \sum_{\gamma=0}^{\infty} p_{b,e}(\gamma;\bullet) f_{b,e}^{\otimes\gamma}(x;\bullet),\ x > 0. \tag{1.2}$$

The severity distribution will be defined as a combination of a lognormal distribution on the center, and a GPD on the right tail whose density $f(x;u,\beta,\xi)$ is :

$$f_{b,e}^{GPD}(x;u,\beta,\xi)) = \begin{cases} \frac{1}{\beta}\left(1+\xi\frac{x-u}{\beta}\right)^{-1-(\frac{1}{\xi})},\ \text{if } x \geq u,\ 1+\xi\left(\frac{(x-u)}{\beta}\right) > 0,\ \beta > 0 \\ \frac{1}{\beta}\left(1-\frac{x-u}{\beta}\right),\ \text{if } x \geq u,\ \xi = 0 \end{cases} \tag{1.3}$$

where $u \in \mathbb{R}^{*+}$ is the threshold. $\beta \in \mathbb{R}^{*}$ is the scale parameter and $\xi \in \mathbb{R}$ the shape parameter. Thus, the density of the severity distribution is:

$$f_{b,e}(x;u,\beta,\xi,\mu,\sigma) = \begin{cases} f_{b,e}^{(center)}(x;\mu,\sigma),\ \text{if } x < u \\ f_{b,e}^{(tail)} = \frac{1}{1-\int_0^u f_{b,e}^{(center)}(x;\mu,\sigma)dx} \times f_{b,e}^{GPD}(x;u,\beta,\xi)),\ \text{if } x \geq u \end{cases}, \tag{1.4}$$

where, μ and σ are the lognormal distribution parameters.

In order to estimate the parameters of the distribution (1.4), we begin with the threshold u carrying out a bootstrap method (Hall (1990), Danielsson et al. (2001)) for which we give practical solutions. Once u is found, ξ and β have to be estimated, therefore we implement a method based on the Anderson-Darling statistics (2.11). Then, assuming a lognormal distribution to model the central part of the severity distribution, we implement an Expectation-Maximization algorithm (Dempster et al. (1977), McLachlan and Krishnan (1997)) to estimate the parameters of the distribution. Finally, the λ parameter of the Poisson distribution $p_{b,e}(\gamma;\lambda)$ is estimated using maximum likelihood method. As soon as we have estimated the whole set of parameters, we can build the loss distribution function (1.1) using a convolution method based on a modified Monte Carlo Algorithm (Fishman (1996)).

As soon as the target of this paper is the computation of capital requirements, we compute the Value-at-Risks (Riskmetrics (1993), Artzner et al. (1999)) enforced at 99.9%, based on the computation of the loss distribution function that we apply on real operational risks data sets.

The paper is organized as follows: in Section two, we describe the estimation of the LDF; Section three is devoted to an application and Section four concludes.

2 Estimation of the LDF

We assume that we observe a data set $X = (X_1, ..., X_n)$, and we denote $\underline{X} = (X_{(1)}, ..., X_{(n)})$ its order statistics, We first estimate the parameter ξ which appears in the GPD (1.3) using the traditional Hill estimator (Hill (1975)):

$$\hat{\xi}_n(k) = \frac{1}{k}\sum_{i=1}^{k} \log X_{(n-i+1)} - \log X_{(n-k)}, \tag{2.1}$$

k being the number of losses above the threshold. We notice that as soon as k is known, the threshold u will be known and conversely. Thus, in the following, we concentrate on the best way to find u. We determine k using the Hill estimate (2.1) and implement a bootstrap method as follows:

1. We plot $\hat{\xi}_n(k)$ with respect to k, as in Figure 1.

2. If we detect a break on the Figure 1. We restrict our information set to the subset X_1 (size n_1) containing this break and the values above it. Using the values in X_1 we determine the parameter k_1 minimizing the following criterion:

$$AMSE(k_1) := E(\hat{\xi}_{n_1}(k_1) - \xi)^2 \tag{2.2}$$

3. In the relation (2.2), the parameter ξ is unknown and we estimate it bootstrapping $\hat{\xi}_{n_1}(k_1)$ as follows. From the set X_1, we draw J ($J \in \mathbb{N}$) subsamples with replacement of size n_2 ($< n_1$), $(X_1^{(j)}, ..., X_{n_2}^{(j)})$, $j = 1, ..., J$. Their corresponding order statistics are $(X_{(1)}^{(j)}, ..., X_{(n_2)}^{(j)})$ and the corresponding bootstrapped Hill estimate of $\hat{\xi}_{n_1}(k_1)$ in X_1 is equal to,

$$\hat{\xi}_{n_2}(k_2) = \frac{1}{k_2}\sum_{i=1}^{k_2} \log \underline{X}_{(n_2-i+1)} - \log \underline{X}_{(n_2-k_2)},\ k_2 = 1, ..., n_2, \tag{2.3}$$

where

$$\underline{X}_l = \frac{1}{J}\sum_{j=1}^{J} X_{(l)}^{(j)},\ l = 1, ..., n_2. \tag{2.4}$$

4. Now, replacing ξ by $\hat{\xi}_{n_2}(k_2)$ in (2.2), we minimize the following criterion with respect to k_2.

$$\widehat{AMSE}(k_2) := E(\hat{\xi}_{n_1}(k_1) - \hat{\xi}_{n_2}(k_2))^2. \tag{2.5}$$

5. Then, given (k_1, k_2), and following Hall (1982; 1990) we estimate[3] u by,

$$\hat{u} = \left\lceil k_2 * \left(\frac{n_1}{n_2}\right)^{\frac{2}{3}} \right\rceil. \tag{2.6}$$

In order to implement the previous methodology, we need to pay attention to the following points:

1. If an obvious break is observed in the representation of $\hat{\xi}_n(k)$ with respect to k (Step 1), we directly have $\hat{u}$.

2. In the step 4, k_1 is fixed. In practice, we can hesitate between several k_1-values, denoted $(k_1^1, ..., k_1^{\Lambda})$, $\Lambda \in \mathbb{N}$, which means that for each $k_1^{(j)}$, $j = 1, ..., \Lambda$, we obtain a k_2-value denoted $k_2^{(j)}$. For each couple $(k_1^{(j)}, k_2^{(j)})$, we choose the $k_2^{(j)}$-value, for which the relation

[3] We are only working on a part of the initial sample, it is necessary to adjust k_2 to the whole data set.

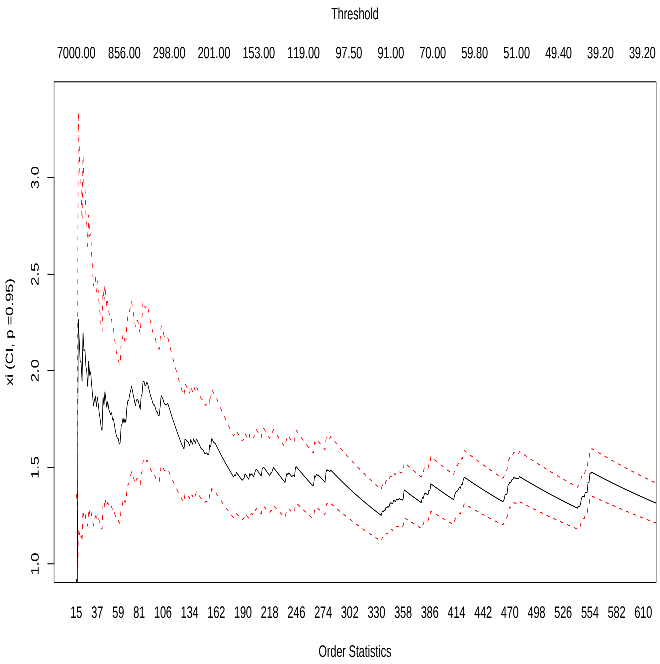

Figure 1: Using the 610 highest values of the Basel case (Payment & settlement, Delivery & Process, Management) of 2006, we represent the Hill estimates considering thresholds varying between 39.2 and 7000 euros. The graph becomes roughly stable around the value 180, but we have no clear break, thus to estimate u, we define the set X_1 from the value 100: X_1 contains 290 data.

(2.5) is the most stable, and therefore, we compute the criteria (2.5) for d successive values of k_2 taken around $k_2^{(j)}$. In that case the parameter k_2 is obtained minimizing:

$$\phi(k_2^{(j)}, d) = \sqrt{\sum_{i=1}^{d} \left(M_1^{(j)} - M_i^{*(j)}\right)^2}, \; j = 1, ..., \Lambda, \tag{2.7}$$

where, $M_1^{(j)}$ is obtained from (2.5) for each $k_2^{(j)}$, and $M_i^{*(j)}$ is computed using the i value around $k_2^{(j)}$, $i = 1, ..., d$.

3. If we cannot interpret the Hill plot (Step 1), either because the data set is to unstable or too stable, we use the whole information set X, and we follow Danielsson et al. (2001) to estimate the parameter ξ:

$$\tilde{\xi}_n(k) = \frac{\frac{1}{k}\sum_{i=1}^{k}(\log X_{(n-i+1)} - \log X_{(n-k)})^2}{2\hat{\xi}_n(k)}. \tag{2.8}$$

For a certain[4] n_2, we bootstrap $\tilde{\xi}_n(k)$ which appears in relationship (2.8) following the step 3 of the previous algorithm to create $\tilde{\xi}_{n_2}(k_2)$, where the parameter k_2 is obtained minimizing,

$$\widetilde{AMSE}(k_2) := E(\hat{\xi}_{n_2}(k_2) - \tilde{\xi}_{n_2}(k_2))^2. \quad (2.9)$$

Then for given k_2, $\hat{u}$ is equal to:

$$\hat{u} = \frac{(k_2)^2}{k_3} \left(\frac{(\log k_2)^2}{(2\log n_2 - \log k_2)^2} \right)^{\frac{\log n_2 - \log k_2}{\log n_2}}, \quad (2.10)$$

where k_3 is obtained exactly as k_2 with $n_3 = \frac{n_2^2}{n}$.

When the estimated threshold u is known, the severity distribution is constituted of two sets: the values below $\hat{u}$ representing the left tail and the center of the severity distribution, and the values above $\hat{u}$ representing the right tail. To estimate the parameters ξ and β which appear in the right tail of the severity distribution (1.4), we privilege the Anderson-Darling approach (Luceno (2006)):

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{(F^{(GPD)}(x; u, \beta, \xi) - S_n(x))^2}{F(x; u, \beta, \xi)(1 - F^{(GPD)}(x; u, \beta, \xi)))} dF^{(GPD)}(x; u, \beta, \xi), \quad (2.11)$$

where, $F^{(GPD)}(x; u, \beta, \xi)$ is the cumulative distribution function of a GPD, and $n \in \mathrm{R}^+$ the number of observations. We compare this method with other methods in the applications and we show that it provides the better result. To estimate the parameters μ and σ of the truncated lognormal distribution which appears in (1.4), we use the Expectation-Maximization method (Dempster et al. (1977)).

Once we have estimated the parameters of the severity distribution, to build the LDF given in relationship (1.1) we use Monte Carlo simulations (Fishman (1996)) using the parameters of each distribution estimated previously. We proceed as follows:

1. We simulate n realizations of the frequency distribution, $q_1, ..., q_n$, $n \geq 1000000$.
2. For each q_i, $i = 1, \cdots, n$, we simulate q_i values of the central distribution, $(s_1, ..., s_{q_i})$, and compare them to the threshold:
   (a) If $s_j < \hat{u}$, $\forall j \in [1, q_i]$ we keep them.

[4] If necessary, a short algorithm providing the optimal n_2 is given in Danielsson et al. (2001).

(b) If $s_j \geq \hat{u}$, we count the number of exceedences, $(z_1, ..., z_n)$ for each frequency q_i and for each i, we draw z_i realizations of the GPD, $(s_1^{(GPD)}, ..., s_{z_i}^{(GPD)})$.

3. Finally, we have $G_i = \sum_{j=1}^{q_i}(s_j + s_j^{(GPD)})$, a realization of the LDF, $i = 1, ..., n$.

4. The set $\hat{G} = (G_1, ..., G_n)$ provides an empirical approximation of the LDF.

We use the estimated LDF $\hat{G}$ to compute a VaR, given a confidence level $\alpha \in [0,1]$, such that,

$$P(\hat{G} > VaR_{1-\alpha}) = (1-\alpha) \tag{2.12}$$

The Figure 2 illustrates the method introduced in this section through a simulation experiment. We exhibit the histogram of the Historical LDF. The black line corresponds to a LDF mixing a Poisson distribution and a lognormal severity. The associated 99.9% VaR is pointed out by Δ. The dash line represents a LDF mixing a Poisson distribution and a multiple pattern of severity distributions (lognormal-GPD given by (1.4)) using the algorithm described in this section. The corresponding VaR is represented by a +. The right graph focuses on the right tail of the LDFs described in this paragraph. This figure highlights the fact that this method enables thickening up the right tail of the LDF.

3 Carrying out

We applied the previous methodology to real operational risks data. We use two data sets denoted $X^{(1)}$ and $X^{(2)}$.

- The set $X^{(1)}$ represents the severity of the business line "Payment & Settlement" and the event type "Delivery, Execution and Process Management" for the year 2006.
- The set $X^{(2)}$ represents the severity of the business line "Retail Banking" and the event type "Business Disruption & System Failure and Exectution" for the year 2007.

Their statistics are given in Table 1. We deduce that their corresponding empirical severity distributions are asymetric (right skewed) and leptokurtic. This statement strengthens our decision to implement a method which cares about extreme values in the distributions.

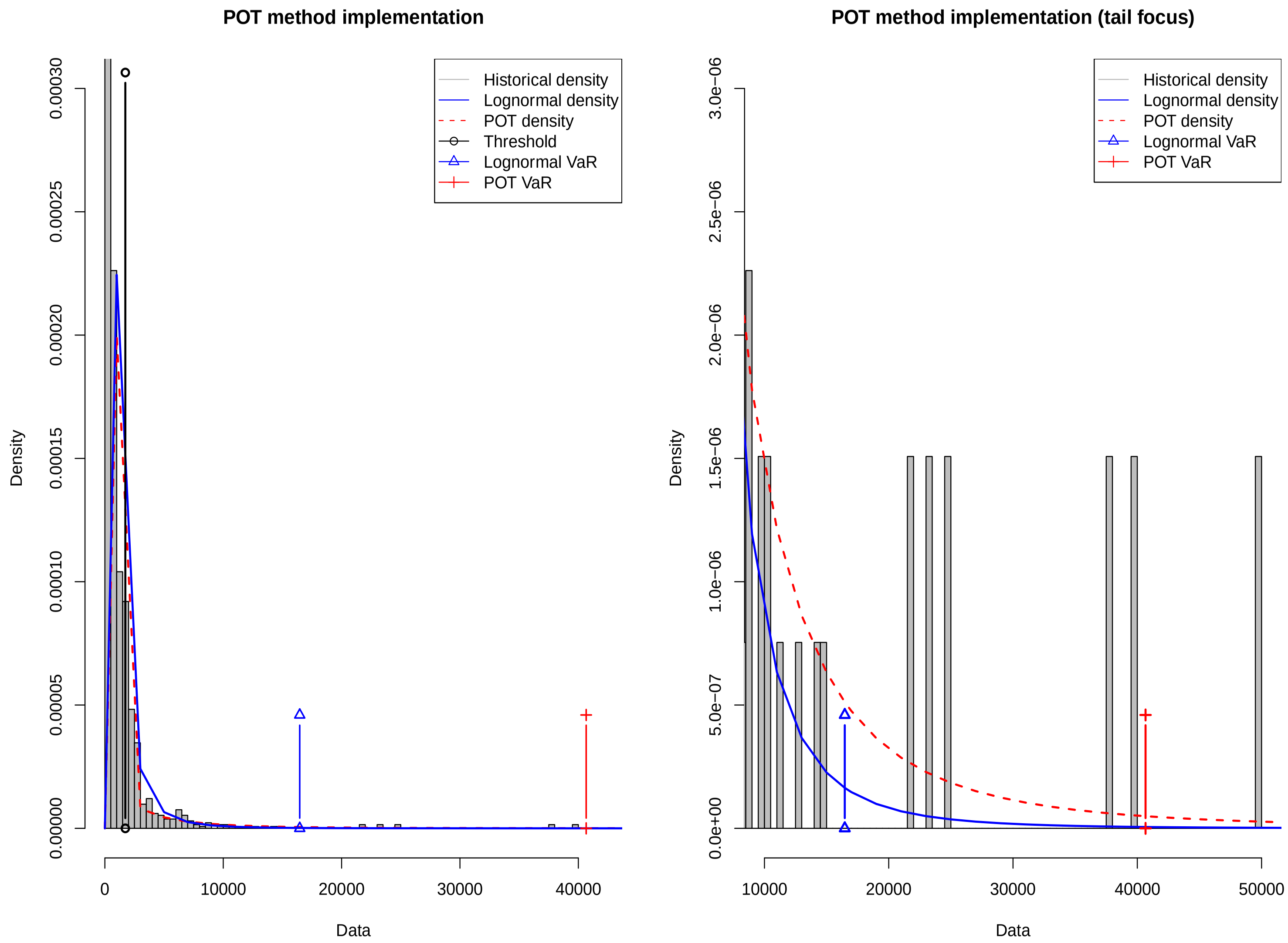

Figure 2: Method Illustration: This figure presents the histogram (in grey) of the Historical LDF. The black line corresponds to a LDF mixing a Poisson distribution and a lognormal severity. The associated 99.9% VaR is pointed out by Δ. The dash line represents a LDF mixing a Poisson distribution and a multiple pattern severity (lognormal-GPD given by (1.4)) using the algorithm discribed in the second section. The corresponding VaR is represented by a $+$. The right graph focuses on the right tail of the LDFs described above. This figure highlights the fact that this method enables thickening up the right tail of the LDF.

Moments	$X_{(1)}$	$X_{(2)}$
Mean	958.4922	1265.871
Standard Deviation	5010.661	10852.48
Skewness	29.84618	10.62595
Kurtosis	1163.478	119.5531

Table 1: First four moments of the two data sets $X^{(1)}$ and $X^{(2)}$. The third order empirical moment is non null and the fourth order emirical moment is far from 3.

In order to calibrate the appropriate loss distribution function on each data set, we use a Poisson distribution to model the frequency data. Estimating the λ parameters by maximum likelihood method, we get $\hat{\lambda}^{(1)} = 2653$ and $\hat{\lambda}^{(2)} = 1292$ respectively on the sets $X^{(1)}$ and $X^{(2)}$. To fit the severities, we tried several usual distributions suggested in the literature, but the Kolmogorov-Smirnov test rejected all of them. Therefore, we applied our methodology. To estimate the threshold u of the GPDs in the relation (1.4), we define two new sets $X_1^{(1)}$ an $X_1^{(2)}$, following the notations introduced in the previous section. Their sizes are respectively $n_1^{(1)} = 531$ and $n_1^{(2)} = 194$. When we bootstrap the data, we used respectively two sets of size $n_2^{(1)} = 266$ and $n_2^{(2)} = 97$. The resulting thresholds are $\hat{u}^{(1)} = 179$ and $\hat{u}^{(2)} = 1645.07$. They imply respectivelly 191 and 315 data above these thresholds. We denote $X^{(GPD_1)}$ and $X^{(GPD_2)}$ the two corresponding data sets on which we fit the GPDs. We estimate the ξ and β parameters of the GPDs defined in (1.4) using the method introduced by (Luceno (2006)) given in (2.11), denoting this method M1. We also consider three other alternative estimation methods to estimate these parameters in order to check their impact on VaR computations. These ones are the Pickands method (M2) (Pickands (1975)), the Hill method (M3) (Hill (1975)), and the Maximum Likelihood method (M4). We provide in Table 2 the estimations of GDPs parameters obtained for both sets, using these four methods. Then we estimate the parameters of the lognormal distributions for the sets $X^{(1)} - X^{(GPD_1)}$ and $X^{(2)} - X^{(GPD_2)}$ and the results are provided in Table 3.

As soon as all the parameters have been estimated, we implement a Monte Carlo algorithm with a million of iterations to build $\hat{G}$, and we use it to compute a 99.9% VaR which provides the required amount of capital. Table 4 provides the VaRs of the set $X^{(2)}$ using the four estimation

Table 2: Estimations of the GPD's parameters ξ and β given in (1.3), using four methods respectively for $\hat{u}^{(1)} = 179$ and $\hat{u}^{(2)} = 1645.07$. We provide in brackets the standard deviations computed by bootstrapping.

Method	$X^{(1)}$		$X^{(2)}$	
	β	ξ	β	ξ
M1	136.678	2.085	932.854	0.767
	(36.91)	(0.687)	(83.71)	(0.101)
M2	133.092	2.173	682.615	1.144
	(39.70)	(0.361)	(160.70)	(0.266)
M3	171.5	1.581	1007	0.66
	(26.19)	(0.668)	(214.36)	(0.228)
M4	159.535	1.948	904.087	0.827
	(27.3755)	(0.21)	(92.31)	(0.097)

$X^{(1)} - X^{(GPD_1)}$		$X^{(2)} - X^{(GPD_2)}$	
$\mu = 3.593098$	$\sigma = 1.510882$	$\mu = 5.702144$	$\sigma = 1.103373$

Table 3: Estimations of the parameters of the lognormal distributions fitted on the data sets $X^{(1)} - X^{(GPD_1)}$ and $X^{(2)} - X^{(GPD_2)}$.

methods to estimate the threshold. We observe miscellanous amounts, varying from 5 725 341 euros to 538 480 990 euros. Thus, as it has a great impact on VaR computation, the choice of the estimation precedure is very important. The M1 method has the advantage that it maximizes the goodness-of-fit of the GPD and insures that the distribution cannot be statistically rejected. Therefore, one can argue that the VaR at 15 700 112 euros could be a reasonnable value considering the information set we have, and the assumptions done.

We do not provide the VaRs for the set $X^{(1)}$ as the results we obtained imply a VaR multiplied by more than one hundred compared to the historical one. Analysing the results and controling every step of the method, we concluded that the problem comes from the structure of the data set. Considering that different natures of data are mixed in a single set as we noticed in the introduction, it appears difficult to fit an accurate distribution on the severities. In order to analyze the impact of the granularity strategy, it would be interesting to split the data in multiple subsets whether the information sets sizes are sufficient, then the computed VaRs would reflect risks created by specific kinds of risk, and finally aggregating them we could compare the results to the largest level of granularity. We cannot provide here this exercise because we have not the source of the data sets.

Method	VaR
M1	538 480 990 euros
M2	5 725 341 euros
M3	21 584 346 euros
M4	27 944 558 euros
M5	15 700 112 euros

Table 4: VaR estimations for $u = 1645.07$, $\mu = 5.681191$ and $\sigma = 1.081609$ for the data set $X^{(2)}$. Estimations have a tremendous impact on the VaRs. In this table, these ones may differ by 9405%.

Finally we want to notice the influence of the threshold on the estimation of the GPD's parameters β and ξ. This effect is due to the fact that the information set changed. We illustrate this point providing the estimates $\hat{\beta}$ and $\hat{\xi}$ for five threshold values, on the set $X^{(1)}$ (Table 5). All the estimated values are relevant despite a great instability.

Threshold u	β	Standard error β	ξ	Standard error ξ
1162.12	1110.8538	87.77	0.6026	0.06877
1608.27	946.6787	94.68	0.7924	0.09329
1645.07	904.087	92.31	0.827	0.09713
2021.48	1188.2956	155.48	0.9149	0.1397
2177.58	1385.5295	199.56	0.8879	0.14

Table 5: This table presents GPD's parameters estimates and their standard errors for different thresholds $\hat{u}$ for the data set $X^{(2)}$. Parameters are very sensitive to the threshold value.

Conclusion

In this paper, we focus on the estimation of the LDF used to estimate operational risks. We highlight that the VaR computation depends on the method implemented. Therefore, a sharp analysis of the LDF is essential, and even if the LDF is a convolution of two distinct distributions, the capital requirements seem particularly sensitive to the severity distribution. We present a new method to compute the severity distribution, spliting it in two parts in order to better take into account the large losses. We use a GPD on the right tail for which we provide innovative theoretical and practical solutions, and fit a lognormal distribution on the remaining data. Then, to build the final LDF, we apply a new adapted Monte Carlo algorithm.

Second, we underline the fact that once the threshold of the GPD has been found, the method chosen to estimate the GPD's parameters tremendously impacts the VaR: it seems that this fact has never been discussed before.

In our applications, we have been confronted to an important problem arising from the data bases construction. Following the regulator, who demands to take into account the right tail of the severity distribution to estimate the LDF, we sometimes obtained unrealistic VaRs (i.e. equal to half the bank capitalization). On the other hand, if we do not care about the severity tail, we have negligible VaRs. Then, even if it seems reasonable to follow regulator requirements, we point out two important facts: either the Bank are overexposed to these risks or the data sets are badly built. We privilege the last point and ask operational risks managers to pay attention to the large diversity of data origins. The correct question is to propose another way for building the data sets used to compute operational risks taking into account the difference which exists inside the different risk categories. It seems that it has no sense to mix certain risks together. In our knowledge, this problem has not yet been discussed and the regulators need to reexamine risk typology and data sets granularity on which are computed operational risk capital requirements.

The method we provide enables to compute capital requirements with respect to Basel accords, so, we stricly stick to them. In this paper we focused on the effectiveness of the proposed method of an operational standpoint, however, other research tracks can extend the results developed in this paper. Even if in this paper we focus on the VaR computation to obtain the corresponding capital requirement, it will be interesting to use other risks measures, for instance the Expected Shortfall (ES) measure, and analyse their impact on the computation of the associated capital requirement. The use of ES, for instance, is not required by the regulator but it would be interesting because it is a coherent risk measure (Artzner et al. (1999)). In another hand there are alternatives to the Monte Carlo method to calculate the LDF (Luo and Shevchenko (2009), Guegan and Hassani (2009)) permitting to improve the speed of the computation that could be investigated. Finally, to take into account the effect of the tail behavior, we focus in this paper on the generalized Pareto distribution, other alternatives could be considered, for example the extreme value distributions described by the Fisher-Tippett theorem, (Fisher and Tippett (1928)), or multivariate methods through the extreme value copulas (Gudendorf and Segers (2010)). All these subjects will be discussed in companon papers.

References

BIS. International convergence of capital measurement and capital standards. *Bank for International Settlements*, 2004.

A. Frachot, P. Georges, and T. Roncalli. Loss distribution approach for operational risk. *Working Paper, GRO, Crédit Lyonnais, Paris*, 2001.

D. Guegan and B. K. Hassani. A modified panjer algorithm for operational risk capital computation. *The Journal of Operational Risk*, 4:53 – 72, 2009.

M. Cruz. *Operational Risk Modelling and Analysis.* Risk Books, London, 2004.

A. Chernobai, S. T. Rachev, and F. J. Fabozzi. *Operational Risk: A Guide to Basel II Capital Requirements, Models, and Analysis.* John Wiley & Sons, 2007.

J. Pickands. Statistical inference using extreme order statistics. *annals of Statistics*, 3:119–131, 1975.

R. Davis and S. Resnick. Tail estimates motivated by extreme value theory. *Annals of Stat.*, 12: 1467 – 1468, 1988.

P. Embrechts, C. Klüppelberg, and T. Mikosh. *Modelling Extremal Events: for Insurance and Finance*. Springer, Berlin, 1997.

BIS. Working paper on the regulatory treatment of operational risk. *Bank for International Settlements*, 2001.

P. Hall. Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of multivariate analysis*, 32:177–203, 1990.

J. Danielsson, L. de Haan, L. Peng, and C.G. de Vries. Using a bootstrap method to choose the sample fraction in tail index estimation. *Journal of Multivariate Analysis*, 76:226–248, 2001.

A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B*, 39:1 – 38, 1977.

G. McLachlan and T. Krishnan. *The EM algorithm and extensions*. John Wiley & Sons, London, 1997.

G. S. Fishman. *Monte Carlo*. Springer-Verlag, Berlin, 1996.

Riskmetrics. Var. *JP Morgan*, 1993.

P. Artzner, F. Delbaen, J-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance 9*, 3:203–228, 1999.

B. M. Hill. A simple general approach to inference about the tail of a distribution. *Ann. Statist.*, 3:1163 – 1174, 1975.

P. Hall. On some simple estimates of an exponent of regular variation. *J.R.S.S.B.*, 44:37 – 42, 1982.

A Luceno. Fitting the generalized pareto distribution to data using maximum goodness-of-fit estimators. *Computational statistics and data analysis*, 51:904 – 917, 2006.

X. Luo and P. Shevchenko. Computing tails of compound distributions using direct numerical integration. *Journal of Computational Finance*, 13, 2009.

R. A. Fisher and L. H. C. Tippett. Limiting forms of frequency distributions of the largest or smallest member of a sample. *Proc. Cambridge Philosophical Soc.*, 24:180 – 190, 1928.

G. Gudendorf and J. Segers. *Extreme-value copulas. In: Proceedings of the Workshop on Copula Theory and Its Applications (editor: P. Jaworski)*. 2010.