Operational Risk Quantification using Extreme Value Theory and Copulas: *From Theory to Practice*

Elise Gourier, Walter Farkas and Donato Abbate

Industrial-Academic Forum on Operational Risk
Fields Institute - Toronto

March 26, 2010

Our work in short

Focus:

Illustrate some possible limitations of the standard techniques in operational risk quantification.

Motivation:

- Scarce literature with practical applications.
- Usually focus either on the modeling of operational risk using Extreme Value Theory (EVT) or on the aggregation of with copulas, not on both.

Contribution:

- Make the link between the existing theory on EVT <u>and</u> on Copulas using real heavy-tailed data.
- Evaluate and compare the impact of the different parameters involved in the process.

Operational risk: an acute challenge

Several examples:

- December 2008: Fraud at Bernard Madoff's firm → 50 billion dollar loss.
- January 2008: Fraud at the Société Générale → 4.9 billion euro loss.

⇒ 2004: The Basel II Accord gives the first response to the need of a regulatory framework for operational risk, defined as:

> *"the risk of direct or indirect loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk."*

New features in Basel II: Capital charges allocated to operational risk.

⇒ How to determine these capital charges?

The Advanced Measurement Approach (Basel II)

- Losses classified in 8 business lines and 7 event types.
- Collection of internal and external data.
- Relies on internal simulations of potential loss distributions.
- Aggregation of losses.
- Stress scenarios.

We will examine the following points, based on real data:

1. How does Classical analysis perform to model OR losses?
2. Extreme Value Theory: a solution to fit the tail?
3. Is the VaR a reliable measure of risk for calculating capital charges?
4. How do copulas perform to aggregate the capital charges of different business lines?

Description of the data

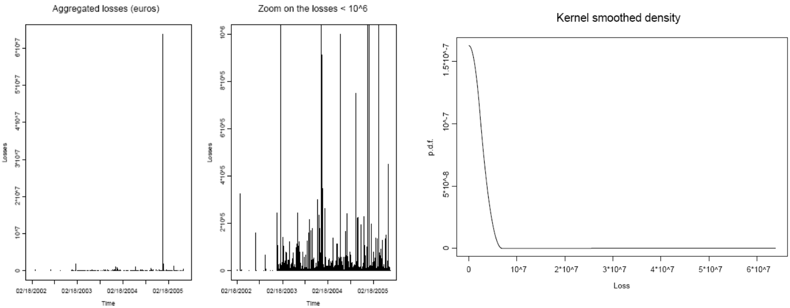
Main features:

- 7514 data points collected from an individual bank database between 2002 and 2006.
- 9 business lines (BL), 7 event types (ET).

Business lines
BL1: Agency Services
BL2: Asset Management
BL3: Commercial Banking
BL4: Corporate Finance
BL5: Payment & Settlement
BL6: Private Banking
BL7: Retail Banking
BL8: Retail Brokerage
BL9: Trading & Sales

Event types
ET1: Business Disruption and System Failures
ET2: Clients, Products and Business Practices
ET3: Damage to Physical Assets
ET4: Employment Practices and Workplace Safety
ET5: Execution, Delivery and Process Management
ET6: External Fraud
ET7: Internal Fraud

Graphical overview of the losses



1. Some statistics after bootstrapping:
   skewness=64.17; kurtosis=4630.

2. Very heavy-tailed data, due to large impact events.

Conventional inference

- Goodness-of-fit tests: Kolmogorov-Smirnov / Anderson-Darling.
- Calculated for classical distribution functions.

	K-S	A-D	$c.v_{5\%}$ K-S	$c.v_{5\%}$ A-D
Exponential	0.43	$+\infty$	0.007	0.20
Weibull	0.07	$+\infty$	0.007	0.23
Gamma	0.07	$+\infty$	0.007	0.21
Lognormal	0.015	3.78	0.007	0.23

$\Rightarrow$ Bad fitting of the tail.

Extreme Value Theory

- Objective: Fit the tail of the distribution: rare, extreme events.
- Provides a model for exceedances over a given threshold u.

Generalized Pareto Distribution

$$GPD_{\xi,\beta}(x) = \begin{cases} 1-(1+\frac{\xi x}{\beta})^{\frac{-1}{\xi}}, & \xi \neq 0 \\ 1-exp(-\frac{x}{\beta}), & \xi = 0. \end{cases}$$

where $\beta > 0$, $x \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq \frac{-\beta}{\xi}$ when $\xi < 0$. The parameters ξ and β are referred to, respectively, as the shape and scale parameters.

- $\xi < 0$ indicates a **short-tail** Pareto type II distribution.
- $\xi = 0$ yields for **exponential** distributions.
- $\xi > 0$ is satisfied for **heavy-tailed** ordinary Pareto distributions (infinite-mean if $\xi > 1$).

Extreme Value Theory

Define the excess distribution over a threshold u:

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

for $0 \leq x < x_F - u$, where $x_F \leq \infty$ is the right endpoint of F.

Balkema-De Haan (1974) and Pickands (1975)

Under certain conditions (which are satisfied for most of the classical distributions),

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

Tail estimation

Mean-excess plot

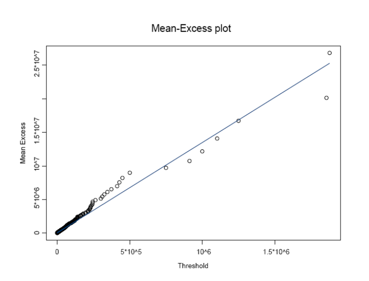
- Define the mean-excess function:

$$e(u) = E(X - u | X > u) = \int x dF_u(x)$$

- For a $GPD_{\xi,\beta}$ random variable:

$$e(u) = \frac{\beta + \xi u}{1 - \xi}$$

where $0 \leq u < \infty$ and $0 \leq u \leq -\frac{\beta}{\xi}$ if $\xi < 0$.



$\Rightarrow$ The GPD seems to be a good candidate

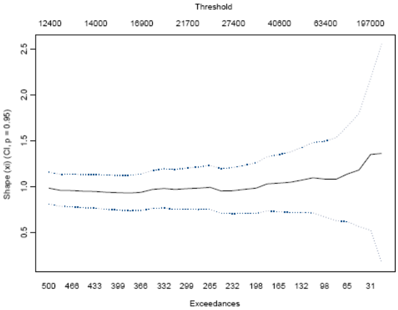
Selection of a threshold

Objective: Select the optimal threshold:

- Balkema-de Haan and Pickands theorem $\Rightarrow$ high threshold.
- If too high $\Rightarrow$ too little data for the fitting.

$\Rightarrow$ Trade-off between **bias** and **variance**!

1. Stability of the shape parameter estimate $\hat{\xi}$ against the threshold.
2. We fix the threshold at the 90% percentile (736 excesses).

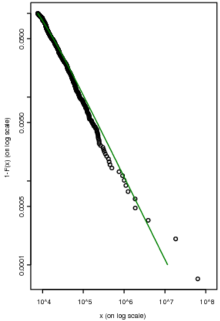


Fitting the tail

We fit the GPD to excesses:

- 0.95 confidence interval for the shape parameter (bootstrapping procedure): $\hat{\xi} \in [0.75, 1.05] \Rightarrow$ Infinite-mean model?
- Tail plot: $\{log(x), log(1 - \hat{F}(x))\}$
- QQ-Plot
- GoF tests:

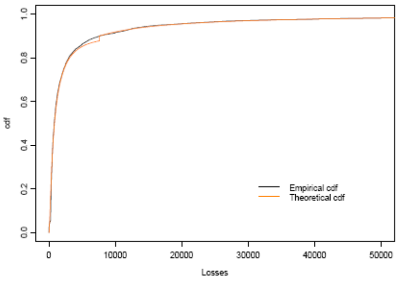
	K-S	A-D	$c.v_{5\%}$ K-S	$c.v_{5\%}$ A-D
GPD	0.003	0.026	0.005	0.13

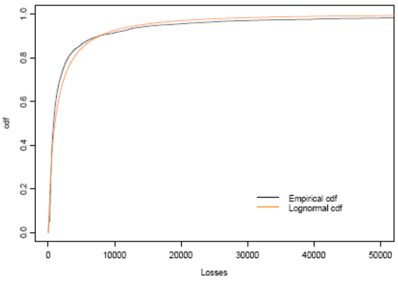


Final model

Let us define F_1 distribution before the threshold and F_2 distribution of the excesses. We have:

$$F(x) = [F_1(x) \cdot 0.90 + 0].1_{x \le u} + [1 \cdot 0.90 + F_2(x) \cdot 0.10].1_{x > u}$$





Capital charges

- Total claim amount:

$$S_N = \sum_{i \leq N} X_i$$

N: claim counting process $\equiv$ Poisson process.

- Distribution of S_N:

$$P(S_N \leq x) = \sum_{n=0}^{\infty} P(S_N \leq x, N = n) = \sum_{n=0}^{\infty} P(N = n) F^{n*}(x)$$

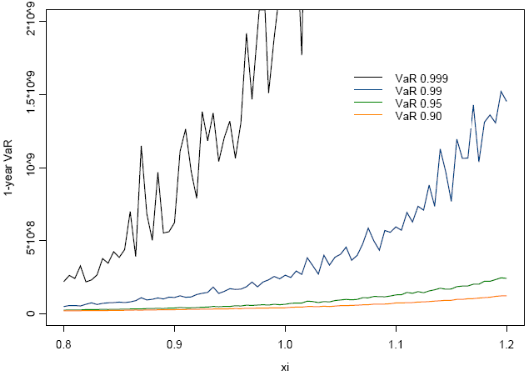
where $F^{n*}(x) = P(S_N \leq x)$ is the n-th convolution of F.
$\Rightarrow$ FFT, Panjer, Monte-Carlo.

α	0.90	0.95	0.99	0.999
VaR_α	26	36	104	639

1-year Value at Risk (in million)

Stability of the VaR

- Sensitivity with respect to the shape parameter:
- Sensitivity with respect to the confidence level:



Final models for the severities

We do the same analysis for business lines and event types when we have more than 250 data.

Objective: Identify the most dangerous units.

	Distribution	Threshold	$\hat{\xi}$	c.i.
BL1	lognormal-GPD	0.85	0.87	[0.55,1.23]
BL3	lognormal-GPD	0.65	1.70	[1.13,2.25]
BL6	GPD	0	0.85	[0.77,1.01]
BL7	lognormal-GPD	0.90	1.05	[0.85,1.20]
BL9	lognormal-GPD	0.75	0.37	[0.13,0.60]

- BL3 very heavy-tailed.
- Several infinite-mean models: encourages to prudence: one big loss can cause ruin.

Coherence of the VaR

α	$\text{VaR}_\alpha(\sum_{i=1,3,7,9} BL_i)$	$\sum_{i=1,3,7,9} (\text{VaR}_\alpha\ BL_i)$
0.90	87	74
0.95	214	187
0.99	2 461	2 331
0.999	126 007	109 052

$\Rightarrow$ VaR super-additive : NOT a coherent risk measure!

Without BL3:

α	$\text{VaR}_\alpha(\sum_{i=1,7,9} BL_i)$	$\sum_{i=1,7,9} (\text{VaR}_\alpha\ BL_i)$
0.90	30	29
0.95	47	48
0.99	196	195
0.999	1 922	2 020

$\Rightarrow$ When $\xi_i < 1$, VaR subadditive for α sufficiently large but not stable

Key points:

- Classical distribution functions do not perform well when heavy-tailed distributions. EVT is needed.
- When $\xi < 1$, VaR is subadditive for high confidence levels, but it is not stable when the quantile is too high: can lead to an overestimation of the capital charges.
- When $\xi > 1$, VaR is never coherent: can lead to underestimation of the capital charges.

Definition

Basel II: The BLs are comonotonic. $\Rightarrow$ Decrease dependence between the BLs

Copula

If the random vector $\mathbf{X}$ has joint df F with continuous marginal distributions $F_1, ..., F_d$, then the copula of F (or $\mathbf{X}$) is the df C of $F_1(X_1), ..., F_d(X_d)$.

Sklar theorem

X_1,...,X_d r.v. with marginal df F_1,...,F_d and joint df F. **There exists** a copula $C : [0,1]^d \to [0,1]$ such that, for all $\mathbf{X} = (X_1, ...X_d)' \in \mathbb{R}^d$,

$$F(X_1, ...X_d) = C(F_1(X_1), ..., F_d(X_d)). \tag{1}$$

If the margins are continuous, then C is unique.

Conversely, let C be a copula, $F_1, ..., F_d$ univariate df, then the function F defined as in (1) is a joint df with margins $F_1, ..., F_d$.

Best-known copulas

1. **Implicit** elliptical copulas:

   We know the joint distribution and the margins.

   Example: **Gaussian copula**: $C_P^{Ga} = \mathbf{\Psi}_{\mathrm{P}}(\Psi^{-1}(\mathrm{u}_1), ..., \Psi^{-1}(\mathrm{u}_\mathrm{d}))$
   Meta-copulas allow to add flexibility in the marginal df.

   Features

   - Easy to use
   - Symmetric

2. **Explicit** Archimedean copulas :

   Close form expression: $C(u_1, ...u_d) = \phi^{-1}(\phi(u_1) + ... + \phi(u_d))$
   $\phi =$ generator of the copula (has to satisfy certain conditions)

   Features

   - Not symmetric any more.
   - Only 1 parameter. $\rightarrow$ Hierarchical copulas.

Calculation of the capital charges

For yearly losses:

1. Choose a copula.
2. Fit the parameters by MLE.
3. Simulate losses by Monte-Carlo.

Resulting capital charges:

	0.90	0.95	0.99	0.999
VaR comonotonicity	65	163	2 076	98 669
VaR_P meta-Gaussian copula	72	172	2 023	103 805
VaR_{Id} meta-Gaussian copula	72	172	2 145	97 217
VaR meta-Student copula	69	171	2 133	126 914
Frank copula	76	180	2 119	87 438
Cook-Johnson copula	68	177	2 131	92 789

- Using copulas yields higher capital charges than making the sum of the VaR. (comonotonicity)!!!
- Goes against the diversification principle $\Rightarrow$ counter-intuitive.

Results

Without BL3:

	0.90	0.95	0.99	0.999
VaR comonotonicity	24	35	110	868
VaR meta-Gaussian copula	24	34	103	742

$\Rightarrow$ When data have finite mean, copulas help decrease the capital charges.

Key points:

- When data have infinite-mean, lowering dependencies between the business lines using copulas does not lower the capital charges.
- The choice of the copula does not have a huge impact on the resulting capital charges. The shape parameter has a much higher impact.

Conclusion

When $\xi < 1$

- VaR is sub-additive when the quantile is sufficiently high.
- VaR is not robust when the quantile is too high.
- Aggregating losses using copulas decreases the capital charges.

When $\xi > 1$

- VaR is not coherent.
- VaR is not robust.
- Using copulas does not lead to a reduction of capital charges.

$\Rightarrow$ **Applying the AMA yields highly uncertain capital charges when infinite-mean models are involved. Caution is required...**

Conclusion

THANK YOU

FOR YOUR ATTENTION !