Multivariate VaRs for Operational Risk Capital Computation : a Vine Structure Approach

Dominique Guegan, Bertrand Hassani

Documents de Travail du
Centre d'Economie de la Sorbonne

CES Working Papers

Multivariate VaRs for Operational Risk Capital Computation : a Vine Structure Approach

Dominique GUEGAN, Bertrand HASSANI

2011.17

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

Multivariate VaRs for Operational Risk Capital Computation: a Vine Structure Approach

April 4, 2011

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

Abstract

The Basel Advanced Measurement Approach requires financial institutions to compute capital requirements on internal data sets. In this paper we introduce a new methodology permitting capital requirements to be linked with operational risks. The data are arranged in a matrix of 56 cells. Constructing a vine architecture, which is a bivariate decomposition of a n-dimensional structure ($n > 2$), we present a novel approach to compute multivariate operational risk VaRs. We discuss multivariate results regarding the impact of the dependence structure on the one hand, and of LDF modeling on the other. Our method is simple to carry out, easy to interpret and complies with the new Basel Committee requirements.

Keywords: Operational risks - Vine Copula - Loss Distribution Function - Nested Structure - VaR

1 Introduction

One proposal of the Basel Committee obliges banks to carry out their own models on internal data sets to evaluate the amounts of capital needed to meet their risks (BCBS (2004)). In terms of operational risks, the proposals are still in their infancy, and most banks' internal strategies are based on traditional modelings, whose assumptions are not always realistic, to follow the Basel Advanced Measurement Approach (BCBS (2004)). It may often be observed that risk managers' teams use very simple or naive methodologies although the operational risks are intrinsically complex. Indeed, to compute capital allocations pertaining to operational risks, two sets of information which cannot be modeled in the same way need to be taken into account. Much is said about the severities and the frequencies of these risk data sets (Frachot et al. (2001)). Besides, dependencies may exist between severities or/and frequencies. As the impact of the severities of capital allocations is more important compared to the frequencies (Guégan et al. (2010)), the present paper focuses on the dependencies between severities and uses them to derive capital allocation. Thus, we propose a new methodology to associate a risk measure with different classes of operational risks, and thus a corresponding level of capital. In order to comply with the Basel committee requirements, we focus on the VaR measure (Artzner et al. (1999)) even if it is a non-coherent measure (our work can be easily extended for the Expected shortfall (ES) measure). To achieve this objective we need to know the distributions characterizing the Loss Distribution Functions, and the distribution characterizing their dependence. The originality of the present work is to provide a manner to obtain the VaR associated with a large set of operational risk categories (> 2). Therefore, to obtain the multivariate distribution associated with this large data set, our work is based on the copula methodology, and as we work with numerous risks categories, we focus on the vine approach (Aas et al. (2009), Guégan and Maugis (2010)).

In order to illustrate the methodology we propose to use the Basel Matrix (BCBS (2001)) given in Table 1. This Table is composed of 56 cases - 8 business lines ("b") $\times$ 7 event types ("e")[1].

[1]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.

This matrix provides loss amounts corresponding to specific events (fraud, clients' practice, etc.) which are associated with banks' business units. For the whole set of cells, the bank needs to provide a risk measure, and then the capital required to face these risks. In the following, we compute these risks by business unites (lines), by event types (columns) and for the global matrix. The specific loss data set we use for applications has been provided by the French *Caisse d'Epargne*. We use 99280 data corresponding to events which occurred between 2006 and 2008. The statistics of the data set are given in Appendix A.

A robust way to measure the dependence between data sets is to compute the joint distribution function associated to these sets, and naturally the concept of copula emerged, as soon as we know that a copula corresponds to a multivariate distribution function associated with a large set of data, together with their standard uniform marginal distributions (Sklar (1959), Nelsen (2006)). In the literature, it has often been mentioned that the use of copulas is difficult in high dimensions apart from when we use the Gaussian and the Student ones (Di Clemente and Romano (2004)). In this paper we drop restrictions by considering recent developments for multivariate copulas, called nested copulas (Morillas (2005), Savu and Trede (2006)) and vine copulas (Aas et al. (2009), Berg and Aas (2009) and Guégan and Maugis (2010)).

When we use n-dimensional copulas to measure risks, we first need to compute the margins, (which correspond to the distribution function of each cell in the Basel matrix in our application). These margins are important because they are used to build the LDF (Loss Distribution Function) from which we estimate the capital allocation associated with each kind of risk. We recall that the LDF is a mix of the distribution characterizing the severities and the distribution of the frequencies. There is now a huge literature for the choice of these two distributions and how to get this LDF (Chernobai et al. (2007)). In this paper we use the Poisson distribution for the frequency distribution and a panel of distributions for the severities including the empirical distribution, the lognormal one, and the Gumbel one. We will see in the application that the calibration of the severity function plays an important role in the computation of the risks, whatever the method used. This fact has already been mentioned in different academic papers (Gourier et al. (2009), Guégan et al. (2010), etc.). It is the reason why we decided to use distributions which capture information in the tails, comparing them with the lognormal distribution which decreases quicker towards zero and which avoids taking into account the probability of important losses occurring.

When the estimations of the severity distributions have been made (for each cell in Table 1), in a second step we focus on the choice of the copulas to measure the dependence between the cells. As soon as we want to work in high dimension ($n > 2$), we face different problems. Apart the multivariate Gaussian distributions and the multivariate Student distributions, Archimedean copulas (Joe (1997)) have attracted particular interest in practice due to their numerous properties which make them simple to analyze. Nevertheless, as soon as we want to measure a dependence between more than two sets, the use of multivariate Archimedean copulas becomes limited as restrictions on the parameters are imposed in order to comply with distribution function characteristics. Therefore, a large number of multivariate Archimedean structures have been developed, called fully nested structures, partially nested copulas and hierarchical ones. Nevertheless, all these architectures imply restrictions on the parameters and impose the choice of copulas at each node making their use limited in practice. Inside the class of Archimedean copulas, the Archimax copulas (Capéraà et al. (2000)), including some extreme-value copulas (Gudendorf and Segers (2010)), could be interesting to use because they provide non-symmetric copulas with positive dependence. Nevertheless in practice we face the same limitations, and we have not been able to use them in our application.

To bypass the restrictions imposed by the previous nested strategy, we can use an intuitive approach proposed by Joe (1996), based on a pair-copula decomposition. This approach rewrites the n-density function associated with the n-copula, as a product of conditional densities. We build step by step all the pair conditioning densities to get the final one. The approach is simple, and it has no restriction for the choice of the functions and their parameters. Its only limitation is the number of decompositions we have to consider. Nevertheless, good algorithms and graph theory for instance may provide an "optimal" strategy and may also limit the choice of the pair decomposition. We do not discuss this problem here, referring to the actual debate on this subject (Antoch and Hanousek (2000), Bedford and Cooke (2002) and Guégan and Maugis (2011)).

We now come back to the core of our paper and we specify how we are going to work to associate a risk measure to the whole set of cells provided in Table 1. One of our objectives is to show that it is possible to use multivariate copulas to measure operational risks capital requirements, beyond the set of Elliptical copulas (Gaussian and Student ones) which do not always correspond to the reality. The second objective is to show the influence of the margin and copula choices on capital computations. We will see that the differences between amounts can be enormous, and thus the regulator must be informed releasing the rules. We summarize our approach considering the following three steps:

- In a first step, we estimate the parameters of the severity distributions (corresponding to the 10 cells of Table 1) chosen among empirical, lognormal and Gumbel distributions.
- In a second step, we estimate the dependence structures between these severities using multivariate Gaussian distributions, nested copulas and vine architectures. We use a pseudo maximum likelihood method to estimate the copulas parameters (Mendes et al. (2007), Weiss (2010)) providing the value of the AIC (Akaike (1974)). We focus on Gumbel, Clayton and Frank copulas (Nelsen (2006)). We also use the Galambos copula (Galambos (1975)) and the Husler-Reiss one (Hüsler and Reiss (1989)) which are interesting because they enable non-linear and tail dependences to be taken into account.
- In a third step, after building the LDF which is a convolution of the distribution of the previous severities with a Poisson distribution (Guégan and Hassani (2009)) (the estimated parameters are provided in Appendix B), we apply the previous dependence structures estimated on the severities to the LDFs[2], and we compute the 99.9% quantile for this n-dimensional LDF which gives the final capital requirement.

The paper is organized as follows. In Section 2 we introduce and detail the methodology and the computation of the dependence structure for the Basel Matrix given in Table 1. Section 3 provides the corresponding capital requirement by lines, columns and for the whole matrix. Section 4 discusses some practical points. Section 5 concludes. All the tables are placed at the end of the paper.

[2]It should be recalled that we assume that the frequency data sets do not change the structure of dependence obtained from the severities.

BUSINESS UNITS	BUSINESS LINES LEVEL 1	(1) Internal Fraud	(2) External Fraud	(3) Employment Practices & Workplace Safety	(4) Clients, Products & Business Practices	(5) Damage to Physical Assets	(6) Business Disruption & System Failures	(7)Execution, Delivery & Process Management
INVESTMENT BANKING	(1) Corporate Finance (including Municipal/Gov.t Finance & merchant banking)							
	(2) Trading & Sales				F_1			
BANKING	(3) Retail Banking		F_8		F_2	F_5	F_9	
	(4) Commercial Banking							
	(5) Payment & Settlement		F_{10}		F_3		F_6	F_7
	(6) Agency Services & Custody							
OTHERS	(7) Asset Management							
	(8) Retail Brokerage				F_4			

Table 1: Basel Matrix for operational risks: each cell represents Business line/event types classification. F_i, $i \in [1, 10]$ denotes the distribution associated to each cell for which data is available.

2 A High Dimensional Structure for Operational Risk

In order to leave the elliptical domain to obtain a dependence measure for large data sets, new n-dimensional dependence structures have been proposed: the nested structures and the pair-copula structures called vine copulas. Nested and pair-copula architectures are based on successive bivariate copula decompositions. Nested copulas originally introduced by Joe (1997), have been studied by Morillas (2005), Savu and Trede (2006), Aas et al. (2009) and Berg and Aas (2009), are based on Archimedean nodes, and can be divided in four types: exchangeable nested copulas, fully nested copulas, partially nested copulas and hierarchically nested copulas. The first class imposes very restrictive dependence structures on the copulas as soon as the number of nodes increases: the m-dimensional marginal distributions ($m < n$) and the generators need to be identical. The three other classes which are extensions of the first structure are more flexible. Nevertheless they are all constrained by the choice of the copula parameter whose value has to decrease as soon as the level of nodes increases.

In Figure 1, we represent the algorithm for the fully nested copulas in a 4-dimensional case. The initial level is composed of univariate distributions, the second level presents the copula considering one-dimensional marginals, the third level is built with a copula considering bivariate marginals and a univariate distribution, and so on. In Figure 2 we illustrate the partially nested copula. The initial level is composed of univariate distributions representing each Basel category for instance, the second level represents the bivariate copulas linking these margins and the third level corresponds to the copula linking the two previous ones. The nested architecture which is a very interesting concept in theory is based on assumptions that are too limited to be used in practice. Pair-copula architectures (Figure 3), suggested by Joe (1996), developed successively by Bedford and Cooke (2001; 2002), Kurowicka and Cooke (2004) and Guégan and Maugis (2010), are more flexible: any class of copula can be used (elliptical, Archimedean or extreme-value for instance), and no restriction is required on the parameters. Nevertheless, compared to nested structures, vine copulas require testing and estimating a large quantity of copulas.

To be more precise the formal representations of nested and vines copulas are defined in the following way. Let $X = [X_1, X_2, ..., X_n]$ be a vector of random variables, with joint distribution F and marginal distributions $F_1, F_2, ..., F_n$: Sklar (1959) theorem insures the existence of a function mapping the individual distributions to the joint one:

$$F(x) = C(F_1(x_1), F_2(x_2), ..., F_n(x_n)),$$

where $x = (x_1, x_2, ..., x_n)$. From any multivariate distribution, F, we can extract the marginal distributions, F_i, and the copula, C, and given any set of marginal distributions $(F_1, F_2, ..., F_n)$ and any copula C, we can use the formula above to compute the joint distribution with the specified marginals and copula. The nested type is the most intuitive way to build n-variate copulas with bivariate copulas, and consists in composing copulas together, yielding formulas of the following type for $n = 3$:

$$\begin{aligned} F(x_1, x_2, x_3) &= C_{\theta_1,\theta_2}(F(x_1), F(x_2), F(x_3)) \\ &= C_{\theta_1}(C_{\theta_2}(x_1, x_2), F(x_3)) \end{aligned}$$

where $\theta_i, i = 1, 2$ is the parameter of the copula. This decomposition can be done several times, allowing to build copulas of any dimension.

If f denotes the density function associated with the distribution F, then another way to compute the joint n-variate distribution function is to decompose it into several increasing conditional distributions. For $n = 3$, we have the following steps:

$$f(x_1, x_2, x_3) = f(x_1).f(x_2|x_1).f(x_3|x_1, x_2),$$

where

$$f(x_2|x_1) = c_{1,2}(F(x_1), F(x_2)).f(x_2),$$

and $c_{1,2}(F(x_1), F(x_2))$ is the density copula associated with the copula C which links the two margins $F(x_1)$ and $F(x_2)$. With the same notations we have:

$$\begin{aligned} f(x_3|x_1, x_2) &= c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)).f(x_3|x_1) \\ &= c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)).c_{1,3}(F(x_1), F(x_3)).f(x_3). \end{aligned}$$

That last formula is a vine decomposition. Many other decompositions are possible using different permutations. For instance, we can consider the following one:

$$\begin{aligned} f(x_1, x_2, x_3) =& f(x_1).f(x_2).f(x_3) \\ & .c_{1,2}(F(x_1), F(x_2)).c_{1,3}(F(x_1), F(x_3)) \\ & .c_{2,3|1}(F(x_2|x_1), F(x_3|x_1)). \end{aligned}$$

In the applications below, we focus on these vine copulas. To our knowledge such a methodology has never been used to associate a risk measure with operational risks. Thus, we extend the work of Gourier et al. (2009), etc. developed in the bivariate case, or the work of Di Clemente and Romano (2004) based on elliptical copulas.

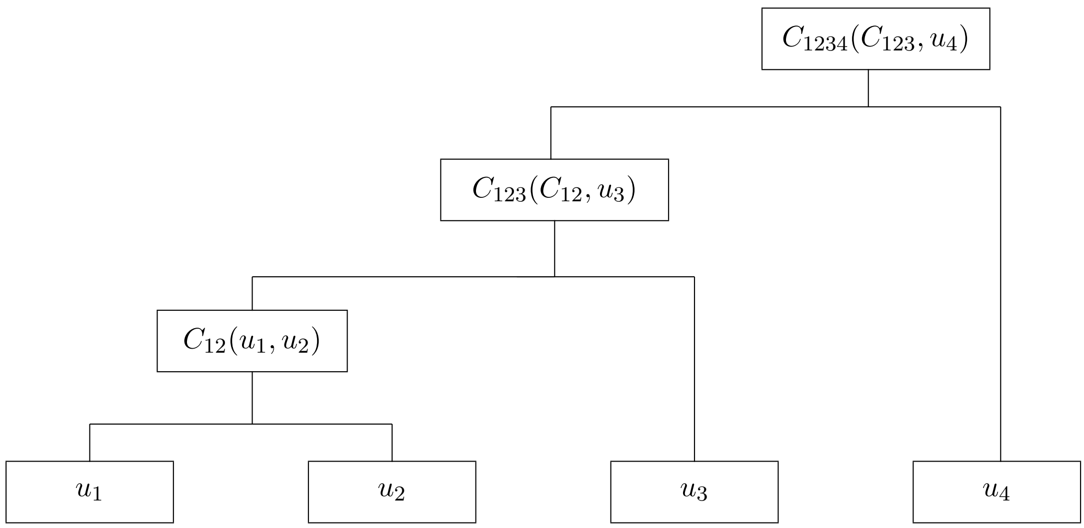

Figure 1: Fully Nested Copula illustration

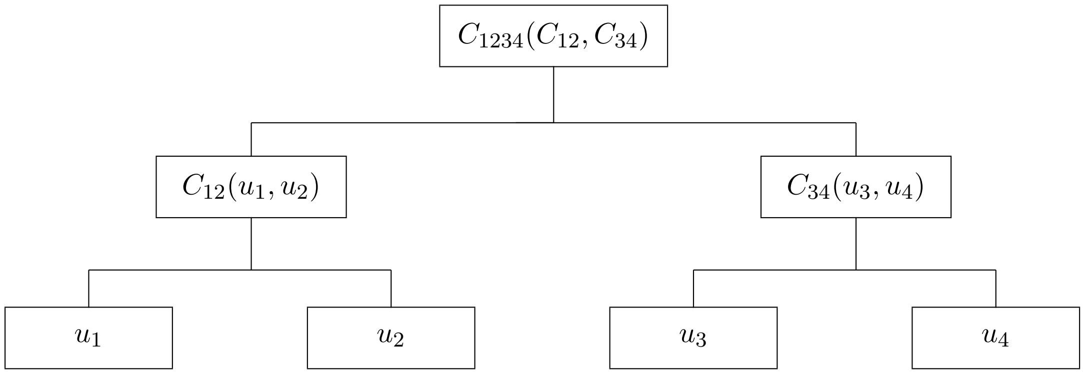

Figure 2: Partially Nested Copula illustration

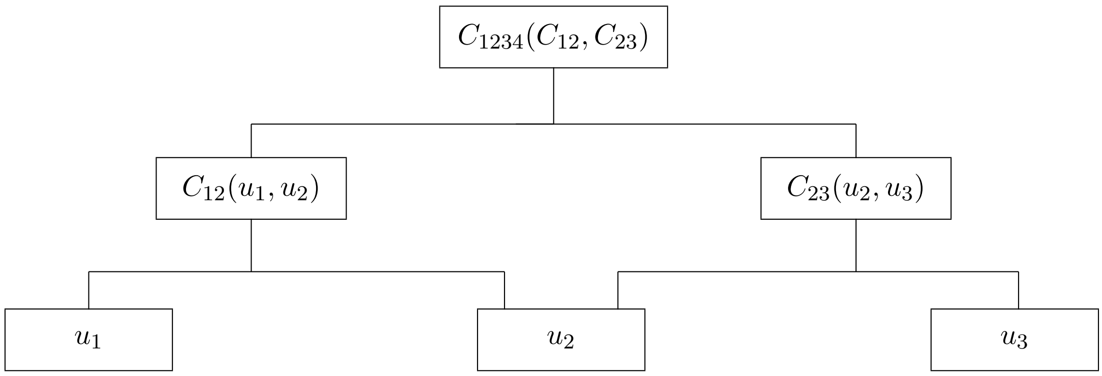

Figure 3: 3-dimensional Vine illustration

If we denote F_i, $i = 1, ..., 10$, the distributions of the severities corresponding to each cell of Table 1, using the previous methodology we proceed as follows.

After estimating empirically each margin F_i, $i = 1, ..., 10$, we focus on specific dependence structures between these margins. We measure the dependence for different sets of severities, say:

1. F_2, F_5, F_8 and F_9 which enable modeling event type dependences, considering the "Retail Banking" business unit;

2. F_3, F_6, F_7 and F_{10} which enable modeling event type dependences, considering the "Pay-

ment and Settlement" business unit;

3. F_1, F_2, F_3 and F_4 which enable modeling business unit dependences, considering the "Client, Products & Business Practices" event type; and

4. F_1, F_2, F_3, F_4, F_5, F_6, F_7 ,F_8, F_9 and F_{10}, which enable modeling dependencies of the whole Basel Matrix.

To compare the different methodologies presented previously, we successively adjust these different sets to a multivariate Gaussian copula (denoted A), a partially nested copula (PNC), a fully nested copula (FNC), and finally a vine structure (D), considering Gumbel, Clayton, Frank, Galambos and Husler-Reiss copulas. In order to estimate the successive bivariate copulas parameters, we carry out pseudo-maximum likelihood methods following Mendes et al. (2007) and Weiss (2010). We provide the results for several 4-dimensional sets in Table 2. We observe that the PNC and FNC methodologies do not always permit dependence structures between the severities to be exhibited because the constraints on the copula parameters are not verified. For the Gaussian structure, we provide the ρ parameter. For the vine approach, we only provide the results for the Gumbel copula because, examining the AIC, it is the best adjustment we obtain for all data sets among the previously mentioned copulas. We provide the value of the parameter with its standard deviation in brackets. Comparing adjustments between the multivariate Gaussian copula and the Gumbel one, with respect to the AIC, the last one again provides the best adjustment. We also observe that the vine approach beats the nested one when we are able to adjust it to a data set, here F_3, F_6, F_7 and F_{10}.

Table 3 provides the results obtained using a vine strategy to model the dependence between the ten severities. We only focus on the Gumbel copula which always arrives in the first place inside the previous estimations. Figure 4 illustrates a way to obtain the dependence structure for the whole Basel Matrix. We have chosen the following vine strategy. The first line of the graph is composed of the two Gumbel copulas $C_{2589}(F_2, F_5, F_8, F_9)$ and $C_{36710}(F_3, F_6, F_7, F_{10})$, and by the two univariate distribution functions F_1 and F_4. The second line exhibits the dependence structure between these four cumulative distribution functions, and so on.

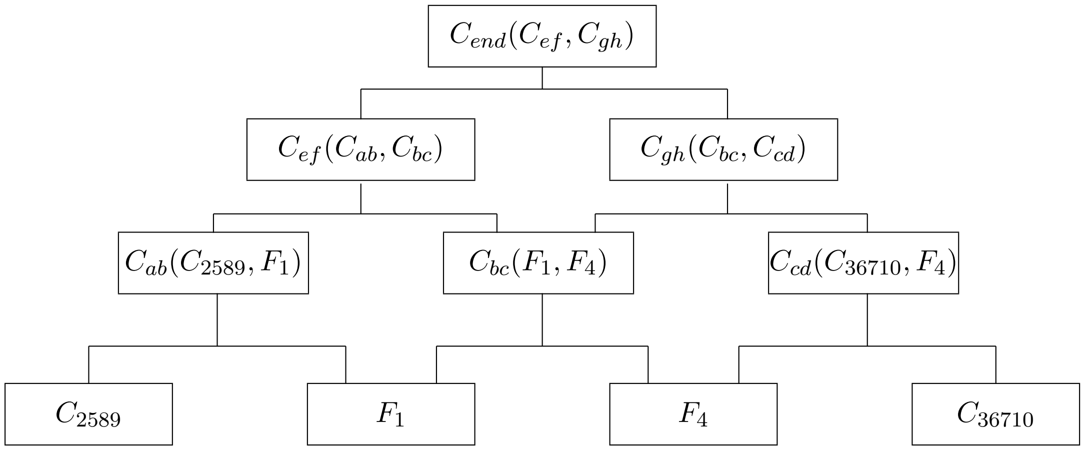

Figure 4: 4-dimensional Vine estimation to obtain the dependence structure for the whole Basel Matrix.

3 Capital Requirements

As soon as the dependence structure between the different cells of the Basel matrix is known, we use it to provide the corresponding capital requirement. An essential tool to obtain it, is the Loss Distribution Function, thus using the dependence structures estimated between the severities we build the dependence structure between the corresponding LDFs computed as a convolution of the severities with a Poisson distribution.

To obtain the capital requirement associated with the LDF - provided in Table 4 - three strategies are considered:

1. We associate a VaR measure (and then the corresponding capital) to each cell, and sum the VaR for the number of cells we decided. At each time we compute this VaR for the three severity distributions we previously retained (empirical, lognormal and Gumbel). The results of this strategy correspond to the first line of Table 4, for 4-dimensional data sets, and the first line of Table 5 for the whole matrix.

2. To take into account the dependence structures estimated previously on the severities, we link the LDFs using now a Gumbel copula which always appears as the best one (in the sense of AIC) in our previous estimations. Considering again the three distributions for the severities, we provide in Table 4 (middle line) the corresponding capital requirements

for the 4-dimensional data sets, and in Table 5 (bottom line) for the whole matrix.

3. In the last line of the Table 4 we provide the capital requirement using a multivariate Gaussian distribution as the dependence structure between the four cells. We do not provide this adjustment - in Table 5 - for the 10 cells because this adjustment cannot be retained in this case.

We explain now the results provided in Table 4:

- We observe that the most conservative value for the capital requirement is obtained when LDFs are computed as a convolution of lognormal and Poisson distributions, whatever the strategy used. On the other hand, whatever the dependence structure and the set of severities, we consider the largest risk exposure is explained when we model the LDF as a convolution of Gumbel and Poisson distributions. The results might have been different if we had considered the Gumbel distribution in Fisher and Tippett (1928) sense.
- Whatever the set of severities we consider, the multivariate Gaussian copula applied to the LDF computed as a convolution of lognormal and Poisson distributions provides the most conservative results. Note that this strategy is often used by practitioners, but in our case, we have to reject this adjustment (AIC).
- The first line of Table 4 provides the least conservative results. It corresponds to the classical strategy used by practitioners (sum of the VaRs). Thus our proposal enables "moderating" the risk-taker behavior of financial agents, taking into account the dependence structure and providing more conservative values.

Now, in order to compute the global capital value based on the whole Basel Matrix, several strategies may be carried out using the lines and the columns of Table 1:

1. We can compute the VaR for each Basel category and sum them.
2. We can use the dependence structure previously obtained between the lines and sum the corresponding aggregated VaR.
3. We can do the same considering the columns.

4. Alternatively, we can use the vine methodology developed previously on the 4-dimensional sets, and apply it to the whole structure following the scheme presented in Figure 4, using a Gumbel copula whose parameter is given in Table 3. This is the approach we have chosen.

5. We could also compute the aggregated VaR with a 10-dimensional copula, but this method is very long and we are constrained by the choice of the copulas.

The final capital requirement is given in Table 5, using the first and the fourth previous strategies with three different models for the LDFs (empirical (1), lognormal (2) and Gumbel (3)). We analyze the results provided in this Table.

- The most conservative result is obtained using the Gumbel dependence structure when the LDF is estimated as a convolution of Poisson and lognormal distributions: € 180 619 174.

- The less conservative result is obtained when we do not take into account the dependence structure summing the VaRs obtained from each LDF computed as convolution of Poisson and Gumbel distributions: € 48 871 039.

- Practically, financial institutions often compute capital requirements as the sum of the VaRs which are computed from LDFs, built as convolutions of Poisson and lognormal distributions: € 74 360 305.

We observe that our approach taking into account a dependence structure which captures extreme events with their frequency provides the most conservative result. It is important to note that the methodology behind this result is very easy to carry out by practitioners. We now analyze some points which can be interesting for application purposes.

4 To be more accurate...

In this section, we analyze several important points we have encountered, applying the previous vine approach to the operational risk data sets given in Table 1.

1. From Table 4, which provides the capital amount for a global set of 4 severities, we can derive by projection the amount corresponding to each severity. This is given in Tables 6 to 8. For example, for fourth column, "Clients, Products & Business Practices", we are able to provide the amounts pertaining independently to the "Retail Banking", the

"Trading & Sales", the "Payment & Settlement" and the "Retail Brokerage" business lines. Our approach is interesting because, even if we can directly associate with each cell the capital amount computed from univariate LDFs, here we provide these values taking into account the dependence structure between the cells.

2. Working dynamically and estimating the parameter of the dependence structure, important variations of the Gumbel parameter are observed. This is illustrated in Table 9 in which we associate a Gumbel copula to the LDFs F_9 and F_6. The parameter θ is not the same when we use as an information set the year 2006, or 2007 or 2008, or the whole sample. We notice that the upper tail dependence is larger when we use this last data set. This could have an impact for the computation of capital requirements by practitioners. Indeed, with respect to the information set we use, the capital requirement appears to be more or less conservative.

3. Now we illustrate the impact of the choice of the severity distributions associated with the choice of the dependence structure between these severities on the capital requirement. We use two specific distributions: F_9 characterizing Business Disruption & System Failure events in the Retail Banking business unit and F_6 characterizing the same events in the Payment & Settlement business unit. For the distribution F_9 we use a Gumbel distribution or a lognormal distribution, for the distribution F_6 we use a Generalized Pareto distribution (GPD) or a lognormal distribution. Table 10 provides the capital values when we link these two distributions with a Gumbel copula on one hand and with a Clayton copula on another hand. We give the projections for each cell and also the global value. Table 10 shows that depending on the way we model the margins, we may have tremendous differences between VaRs. For example, we would have a bivariate VaR equal to 117 207 402 euros if F_9 is modeled with a lognormal distribution and F_6 with a GPD distribution versus a VaR equal to 2 037 655 euros if F_9 is modeled with a lognormal distribution and F_6 with a Gumbel one. Depending on the way we model the LDFs, the aggregated VaR may be multiplied by 57.52. The same behavior is observable with the projections. For example, the multivariate VaR projection on F_9 is € 2 655 055 if F_6 is modeled using a lognormal distribution, and is equal to € 15 405 192 if F_6 is modeled using a GPD distribution. The VaR explosion value we observe is due to the fact that in that last case we capture extreme events.

4. Considering now the impact of the copula, we observe in Table 10 that the capital requirements obtained using a Gumbel copula are always bigger than the one obtained with a Clayton one. Using at the same time, copula and severity distributions which take into account information in the tail, provides very accurate results. Indeed, when we model F_6 using a GPD associated with a Gumbel copula, we provide a larger VaR than with the Clayton one. We have to compare the amounts € 105 422 356 with € 103 249 260 on the one hand, and € 117 207 402 with € 107 807 238 on the other.

Finally, all these events point to the importance of the modeling used to associate a capital allocation with any risk.

5 Conclusion

This paper proposes a new methodology to compute the capital requirement associated with operational risks in high dimensions. We focus on the vine copula architecture which permits the choice of the margins and of the copulas to be released. We can retain the following main improvements for practitioners:

- First, this methodology enables the use of numerous classes of copulas without restricting ourselves to the elliptic domain, in particular considering copulas which focus on information contained in the tails, where we can find the large losses.
- Second, our approach allows several combinations of margins to derive robust adjustments in the statistical sense.
- Third, even working in the high dimension, the procedure is easy to implement and is not too time consuming.
- Fourth, our method complies with the new Basel Committee (BCBS (2010)) requirements.

Let us note that the complete Basel matrix could contain more than 250 cells, and thus more developments will be necessary to work with a so large matrix, mainly to limit the time of computation.

The methodology presented in this paper can be applied to any computation for risk measurement and it is not limited to operational risks. The main philosophy of this paper is to provide a robust and easy tool to associate a measure for any large number of risks, bypassing the elliptical approach which is not always adapted to the reality.

Structures	$C_{2589}(F_2,F_5,F_8,F_9)$	*AIC*	$C_{36710}(F_3,F_6,F_7,F_{10})$	*AIC*	$C_{1234}(F_1,F_2,F_3,F_4)$	*AIC*
(A) *Gaussian*	$\rho = 0.665$ (0.0438)	-14.398	$\rho = 0.682$ (0.0422)	-14.367	$\rho = 0.681$ (0.032)	-13.918
(B) *PNC*	$\oslash$	-	$\oslash$	-	$\theta = 1.919$ (0.498)	-14.211
(C) *FNC*	$\oslash$	-	$\oslash$	-	$\oslash$	-
(D) *PairCopula*	$\theta = 17.353$ (0.642)	$\mathbf{-14.592}$	$\theta = 3.764$ (0.579)	$\mathbf{-14.415}$	$\theta = 5.345$ (0.889)	$\mathbf{-14.355}$

Table 2: Parameters estimation for several dependence structures applied to three sets of four severities. The corresponding standard deviations are provided in brackets. "Gaussian" denotes a multivariate Gaussian copula structure, "FNC" denotes Fully Nested copulas, "PNC" denotes Partially Nested copulas and "Pair Copula" denotes a Vine structure. A workable nested structure has been found when the dependence degree was decreasing and the level of nesting increasing. For the copulas C_{2589} and C_{36710} the AIC is better for the Pair Copula structure than for the Gaussian one. Considering the dependence between F_1, F_2, F_3 and F_4, the AIC is better for the Pair Copula structure than for the Nested one. In all cases, the robust dependence structure is provided by a Gumbel copula.

Structures	$C_{12345678910}$	AIC
Pair Copula	$\theta = 5.079$ (1.75)	-11.94

Table 3: This table provides the Gumbel copula parameter obtained modeling the dependence structure between all business lines. The corresponding standard deviation is provided in brackets.

Copula	LDF	$C_{1234}(LDF_1,LDF_2,LDF_3,LDF_4)$	$C_{2589}(LDF_2,LDF_5,LDF_8,LDF_9)$	$C_{36710}(LDF_3,LDF_6,LDF_7,LDF_{10})$
Univariate	1	29 400 454	39 287 139	28 449 066
	2	22 794 418	24 118 977	34 614 408
	3	8 968 115	25 755 128	15 719 280
Gumbel	1	30 646 550	41 456 461	33 252 197
	2	35 474 559	56 176 428	60 970 007
	3	21 778 827	25 798 375	15 762 291
Gaussian	1	31 257 604	43 621 089	37 158 832
	2	40 651 444	75 785 269	68 385 483
	3	21 381 957	25 972 225	15 881 606

Table 4: This table provides the capital allocation for 4-dimensional data sets, considering three classes of severities (1 denotes the non parametric approach of the LDF, 2 the lognormal approach and 3 the Gumbel one.) and three classes of dependence. Univariate corresponds to the VaRs sum of each LDF, Gumbel corresponds to the Gumbel copula and Gaussian to the multivariate Gaussian copula.

Copula	LDF	$C_{12345678910}(C_{2589}(LDF_2,LDF_5,LDF_8,LDF_9),LDF_1,LDF_4,C_{36710}(LDF_3,LDF_6,LDF_7,LDF_{10}))$
Univariate	1	77 018 239
	2	74 360 305
	3	48 871 039
Gumbel	1	84 220 744
	2	180 619 174
	3	49 025 722

Table 5: This table provides the capital allocation for the whole data set, considering three classes of severities (1 denotes the non parametric approach of the LDF, 2 the lognormal approach and 3 the Gumbel one.) and two classes of dependence. Univariate corresponds to the VaRs sum of each LDF, Gumbel corresponds to the Gumbel copula.

Copula	LDF	LDF_1	LDF_2	LDF_3	LDF_4
Univariate	1	278 423	19 650 986	467 434	9 003 611
	2	254 095	6 240 984	926 513	15 372 825
	3	232 591	13 590 317	212 451	7 164 041
Gumbel	1	280 423	20 556 466	524 137	9 285 523
	2	377 479	9 957 836	6 283 207	18 856 037
	3	714 190	13 602 350	260 749	7 201 537
Gaussian	1	287 179	21 110 804	603 161	9 256 460
	2	609 358	9 862 421	3 098 797	27 080 869
	3	237 529	13 632 720	213 645	7 298 063

Table 6: This table provides the VaRs associated with each LDF of the set LDF_1, LDF_2, LDF_3 and LDF_4 when we decompose the dependence structure of the 4-dimensional set C_{1234}, considering three classes of severities (1 denotes the non parametric approach of the LDF, 2 the lognormal approach and 3 the Gumbel one.).

Copula	LDF	LDF_2	LDF_5	LDF_8	LDF_9
	1	19 650 986	3 182 731	14 212 411	2 241 011
Univariate	2	6 240 984	2 151 627	11 676 534	4 049 831
	3	13 599 313	2 478 087	8 599 313	1 087 410
	1	20 578 056	3 300 471	15 191 828	2 386 106
Gumbel	2	10 732 933	2 310 337	29 525 155	13 608 000
	3	13 617 419	2 486 453	8 603 916	10 905 587
	1	22 269 192	3 301 720	15 482 009	2 568 168
Gaussian	2	14 264 778	4 157 787	30 072 894	27 259 810
	3	13 692 945	2 492 189	8 679 579	1 107 512

Table 7: This table provides the VaRs associated with each LDF of the set LDF_2, LDF_5, LDF_8 and LDF_9 when we decompose the dependence structure of the 4-dimensional set C_{2589}, considering three classes of severities (1 denotes the non parametric approach of the LDF, 2 the lognormal approach and 3 the Gumbel one.).

Copula	LDF	LDF_3	LDF_6	LDF_7	LDF_{10}
Univariate	1	467 435	2 739 085	1 295 941	23 946 606
	2	926 513	306 553	2 425 710	30 955 632
	3	212 451	772 003	829 360	13 905 464
Gumbel	1	628 263	3 220 812	1 473 900	27 929 223
	2	4 191 332	832 333	11 559 821	44 386 520
	3	213 844	775 924	831 552	13 940 970
Gaussian	1	584 053	3 572 152	1 394 149	31 608 479
	2	7 105 044	1 257 046	12 247 696	47 775 697
	3	216 849	789 476	831 840	14 053 440

Table 8: This table provides the VaRs associated with each LDF of the set LDF_3, LDF_6, LDF_7 and LDF_{10} when we decompose the dependence structure of the 4-dimensional set C_{36710}, considering three classes of severities (1 denotes the non parametric approach of the LDF, 2 the lognormal approach and 3 the Gumbel one.).

Year	θ	θ
2006	4.9202 (0.94)	
2007	3.7206 (0.75)	10.6610 (0.88)
2008	5.8490 (0.51)	

Table 9: Parameter estimation of Gumbel copulas estimated on F_9 and F_6 for each year 2006, 2007 and 2008 (second column). These parameters are compared to a Gumbel copula parameter estimated on the whole chronicle (third column). The corresponding standard deviation are provided in brackets.

Model	Gumbel Copula			Clayton Copula		
	LDF_9	LDF_6	Sum	LDF_9	LDF_6	Sum
Gumbel-GPD	2 322 782	103 099 574	105 422 356	1 154 681	102 094 579	103 249 260
Gumbel-lognormal	1 471 343	566 312	2 037 655	1 455 693	649 164	2 104 857
lognormal-GPD	15 405 192	101 802 210	117 207 402	5 631 004	102 176 234	107 807 238

Table 10: For the LDF corresponding to F_9 and F_6 we provide the VaRs computed from Gumbel and Clayton copulas for the year 2006. They are given respectively for three classes of severities. For instance, "Gumbel-GPD" means that we have chosen a Gumbel distribution to model F_9 and a mix of a lognormal and a GPD to model F_6.

A Appendix: Distributions statistics

Next table provides the four first moments of the empirical severities used in this paper.

Distributions	Mean	Variance	Skewness	Kurtosis
F_1	195.37	292732.86	7.31	71.69
F_2	1522.83	372183311.54	27.57	910.13
F_3	175.42	3804557.63	30.03	956.75
F_4	1805.81	93274002.03	18.74	457.58
F_5	1824.95	189175093.33	17.79	354.00
F_6	1200.08	438224165.80	23.69	563.48
F_7	800.14	24268504.39	10.88	139.39
F_8	1779	1602373386	19.27	435.88
F_9	1824.95	189175093.3	17.79	354.00
F_{10}	12104	519962084.2	108.03	11806.23

Table 11: Statistics of the data sets used. The distributions are right skewed and present large kurtosis.

B Appendix: LDF Parameters

This appendix provides LDF's parameters estimations regarding considered models.

Distributions	Poisson	lognormal		Gumbel	
F_1	$\lambda = 1094$	$\mu = 4.03$	$\sigma = 1.47$	$u = 149.88$	$\beta = 72.98$
F_2	$\lambda = 8448$	$\mu = 4.25$	$\sigma = 1.97$	$u = 1381.07$	$\beta = 210.43$
F_3	$\lambda = 1114$	$\mu = 2.80$	$\sigma = 2.23$	$u = 150.02$	$\beta = 40.65$
F_4	$\lambda = 3811$	$\mu = 5.72$	$\sigma = 1.99$	$u = 1443.89$	$\beta = 585.89$
F_5	$\lambda = 521$	$\mu = 4.87$	$\sigma = 2.19$	$u = 848.76$	$\beta = 471.77$
F_6	$\lambda = 575$	$\mu = 4.03$	$\sigma = 1.71$	$u = 1101.79$	$\beta = 143.73$
F_7	$\lambda = 937$	$\mu = 3.72$	$\sigma = 2.27$	$u = 721.71$	$\beta = 135.75$
F_8	$\lambda = 1178$	$\mu = 5.42$	$\sigma = 2.14$	$u = 4010.25$	$\beta = 833.45$
F_9	$\lambda = 8748$	$\mu = 4.87$	$\sigma = 2.19$	$u = 1612.31$	$\beta = 370.63$
F_{10}	$\lambda = 12103$	$\mu = 5.49$	$\sigma = 2.00$	$u = 861.38$	$\beta = 437.14$

Table 12: This table provides the parameters estimation for each LDF for the year 2008, assuming a Poisson distribution to model the frequencies, and either a lognormal or a Gumbel distribution to model the severities.

Distributions	Poisson-Gumbel	
LDF_9	$\lambda = 658$	$u = 191.5378$, $\beta = 938.9768$ (s.e. 36.70), (s.e. 35.64)
Distributions	Poisson-GPD	
LDF_6	$\lambda = 1292$	$\mu = 5.70$, $\sigma = 1.10$ $u = 1645.07$, $\beta = 932.854$, $\xi = 0.767$

Table 13: This table provides two LDFs' parameters, LDF_6 and LDF_9, for the year 2006, assuming a Poisson distribution to model the frequencies and either a Gumbel or a mix of a lognormal and a GPD to model the severities.

References

Aas, K., Czado, C., Frigessi, A. and Bakken, H. (2009), 'Pair copula constructions of multiple dependence.', *Insur. Math. Econ.* **44**, 182–198.

Akaike, H. (1974), 'A new look at the statistical model identification.', *IEEE Transactions on Automatic Control* **19**, 716–723.

Antoch, J. and Hanousek, J. (2000), 'Model selection and simplification using lattices.', *CERGE-EI Working Paper, Charles University in Prague, Czech Republic* .

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

BCBS (2001), 'Working paper on the regulatory treatment of operational risk', *Basel Committee on Banking Supervision, Basel, Switzerland* .

BCBS (2004), 'Basel 2: International convergence of capital measurement and capital standards: a revised framework', *Basel Committee on Banking Supervision, Basel, Switzerland* .

BCBS (2010), 'Basel committee: Operational risk - supervisory guidelines for the advanced measurement approaches - consultative document', *Basel Committee on Banking Supervision, Basel, Switzerland* .

Bedford, T. and Cooke, R. (2002), 'Vines: A new graphical model for dependent random variables.', *The Annals of Statistics* **30(4)**, 1031–1068.

Bedford, T. and Cooke, R. M. (2001), 'Probability density decomposition for conditionally dependent random variables modeled by vines.', *Annals of Mathematics and Artificial Intelligence* **32**, 245–268.

Berg, D. and Aas, K. (2009), 'Models for construction of multivariate dependence - a comparison study', *The European Journal of Finance* **15**, 639–659.

Capéraà, P., Fougères, A. and Genest, C. (2000), 'Bivariate distributions with given extreme value attractor.', *Journal of Multivariate Analysis* **72**, 30–49.

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

Di Clemente, A. and Romano, C. (2004), 'A copula-extreme value theory approach for modeling operational risk', *Operational Risk Modelling and Analysis* pp. 189–208.

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

Frachot, A., Georges, P. and Roncalli, T. (2001), 'Loss distribution approach for operational risk', *Working Paper, GRO, Crédit Lyonnais, Paris, France* .

Galambos, J. (1975), 'Order statistics of samples from multivariate distributions.', *Amer. Statist. Assoc.* **10**, 674–680.

Gourier, E., Farkas, W. and Abbate, D. (2009), 'Operational risk quantification using extreme value theory and copulas: from theory to practice.', *The Journal of Operational Risk* **4**.

Gudendorf, G. and Segers, J. (2010), *Extreme-value copulas. In: Proceedings of the Workshop on Copula Theory and Its Applications held in Warsaw (editor: P. Jaworski)*, Springer Media, New York.

Guégan, D. and Hassani, B. K. (2009), 'A modified panjer algorithm for operational risk capital calculations', *The Journal of Operational Risk* **4**, 53 – 72.

Guégan, D., Hassani, B. K. and Naud, C. (2010), 'An efficient peak-over-threshold implementation for operational risk capital computation', *WP-2010-96, Publications du CES Paris1 Panthéon - Sorbonne, France.* .

Guégan, D. and Maugis, P.-A. (2010), 'New prospects on vines.', *Insurance Markets and Companies: Analyses and Actuarial Computations* **1**, 4–11.

Guégan, D. and Maugis, P.-A. (2011), 'An econometric study for vine copulas.', *to appear in: International Journal of Economics and Finance* .

Hüsler, J. and Reiss, R. (1989), 'Maxima of normal random vectors: Between independence and complete dependence.', *Statistics & Probability Letters* **7**, 283–286.

Joe, H. (1996), 'Families of m-variate distributions with given margins and $\frac{m(m-1)}{2}$ bivariate dependence parameters.', *Distributions with Fixed Marginals and Related Topics, Lecture Notes-Monograph Series* **28**, 120–141.

Joe, H. (1997), *Multivariate models and dependence concepts.*, Monographs on Statistics and Applied Probability, Chapman and Hall, London.

Kurowicka, D. and Cooke, R. M. (2004), 'Distribution - free continuous bayesian belief nets.', *In Fourth International Conference on Mathematical Methods in Reliability Methodology and Practice, Santa Fe, New Mexico.* .

Mendes, B., de Melo, E. and Nelsen, R. (2007), 'Robust fits for copula models', *Comm. in Stat.: Sim. Comp* **36**, 997–1017.

Morillas, P. (2005), 'A method to obtain new copulas from a given one.', *Metrika* **61**, 169–184.

Nelsen, R. (2006), *An Introduction to Copulas.*, Springer Series in Statistics, Berlin, Germany.

Savu, C. and Trede, M. (2006), 'Hierarchical archimedean copulas.', *Working Paper, University of Münster, Germany* .

Sklar, A. (1959), 'Fonctions de répartition à n dimensions et leurs marges', *Publ. Inst. Stat.* **8**, 229–231.

Weiss, G. (2010), 'Copula parameter estimation - numerical considerations and implications for risk management', *Journal of Risk* **13**, 17–53.