

Vine copula: the new approach for modeling high dimensional dependencies -application to financial data-

الرابطة العنقودية: النهج الجديد لنمذجة الارتباطات عالية الأبعاد- تطبيق على البيانات المالية-

Hamadouche Bekhta ¹, Ben Anaya Djelloul ²

¹ Assistant Professor, Djilali Bounama University-Khemis Miliana(Algeria), b.hamadouche@univ-dbk.m.dz

² Professor, Djilali Bounama University-Khemis Miliana (Algeria), d.ben-anaya@univ-dbk.m.dz

Received: 2021/2/23

Accepted: 2021/3/26

Published: 2021/7/1

Abstract:

A popular approach to build a High-dimensional dependence is based on copulas and vine copulas. The aim of this work is: the construction of several structures of dependence of a set of stock market indices via vine copulas for risk management. For this, we took as data four stock market indices, namely the cac40, dawJones, Nikkei225, FTSE100. Each series was modeled separately using the ARMA-GARCH models then we switched to multiple modeling via vine copula. The structure of the vine was identified using Kendall's τ .

Keywords: Copulas ; vine copula ; dependency structure ; stock market indices.

Jel Classification Codes: C02 , G11..

ملخص:

من بين النماذج الشائعة لنمذجة الارتباطات عالية الأبعاد (أي بين عدد كبير من المتغيرات) هي نماذج الرابطة copula و الرابطة العنقودية vine copulas. الهدف من هذا العمل : نمذجة الارتباطات لمجموعة من مؤشرات البورصة باستعمال نماذج الرابطة و الرابطة العنقودية ، لهذا الغرض استعملنا أربع مؤشرات لبورصات عالمية هي: cac40 ، dowJones ، Nikkei225 ، FTSE100. تمت نمذجة كل متغير على حدى باستعمال نماذج ARMA-GARCH ثم انتقلنا إلى النمذجة المتعددة باستعمال نماذج الرابطة العنقودية ، تم تحديد هيكل الرابطة العنقودية باستعمال معامل ارتباط الرتبة لكندل Kendall's τ .

كلمات مفتاحية: الرابطة ، الرابطة العنقودية، هيكل الارتباط، مؤشرات البورصة .
تصنيف JEL : XN1 ، XN2 .

Corresponding author: HAMADOUCHE Bekhta, e-mail: b.hamadouche@univ-dbk.m.dz.

1.INTRODUCTION:

Within the last decades, not just within the last financial crisis, risk management has become more and more important. Key points are the aggregation of risks and their joint evaluation. For risk assessment risk measures were introduced. Up to now the most widely used one is « value-at-risk » (VaR). For the aggregation of different risks it is highly desirable to model these risks jointly.

At first multivariate distributions like Gaussian or Student's t were used for this task. But as the dependence structure among different risks is not homogeneous the need for more flexible models was obvious. Copula models started to bridge this gap. The copula theory, introduced in 1959 by Abe Sklar and applied for the first time in finance at the end of the twentieth century, is a very powerful and flexible statistical tool for modeling.

Copulas are mathematical objects that fully capture the dependence structure among random variables and hence, offer a great flexibility in building multivariate stochastic models.

These functions serve as a link between the multivariate distribution function of several variables and its one-dimensional margins (Sklar [1959]) and, hence, present a modern approach allowing for the separate modeling of the underlying dependence structure.

Although bivariate copulas were soon well explored, multivariate ones were harder to construct and still suffered a lack of flexibility. This was solved with so-called vine copula. It is a scheme to construct high-dimensional copulas out of bivariate building blocks.

We will try in this paper, using vine copula, to build a flexible multivariate model which takes into account the non-homogeneous dependencies between the different financial markets.

2. Copulas theory

This part is a brief introduction to copulas. It does not dwell on their mathematical foundations or the full range of their properties, but the topics covered are intended for the proper understanding of the next steps. Note that we are introducing the concept of two-dimensional copulas for the sake of simplicity; all the results presented can be generalized to a larger dimension. In addition, it takes up the order of presentation and several examples from the work of Nelsen (Nelson, 2007).

Copulas are used, among other things, to describe the joint distribution of random variables. However before you can perform this task, it is important to understand what a copula. A very simple definition could be: a copula is a function C having certain characteristics and whose domain is $[0,1] \times [0,1]$ and the image is in $[0,1]$. The next section focuses on describing the characteristics that make a function a copula, and then the link between copulas and random variables is described. For the continuation I will represent the interval $[0,1]$.

2.1. Copula definitions:

Definition 1: A d -dimensional copula, $C : [0; 1]^d \rightarrow [0; 1]$ is a cumulative distribution function (CDF) with uniform marginals

We write $\mathbf{C}(\mathbf{u}) = \mathbf{C}(\mathbf{u}_1, \dots, \mathbf{u}_d)$ for a generic copula and immediately have the following properties :

1. $C(u_1, \dots, u_d)$ is non-decreasing in each component, u_i .
2. The i^{th} marginal distribution is obtained by setting $u_j = 1$ for $j \neq i$ and since it is uniformly distributed $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$
3. For $a_i \leq b_i$, $P(U_1 \in [a_1, b_1], \dots, U_d \in [a_d, b_d])$ must be non-negative.

This implies the rectangle inequality

$$\forall (u_1^{(1)}, \dots, u_d^{(1)}) \in [0,1]^d \sum_{(i_1, \dots, i_d) \in \{1,2\}^d} (-1)^{\sum_{j=1}^n i_j} C(u_1^{(i_1)}, u_2^{(i_2)}, \dots, u_d^{(i_d)}) \geq 0$$

$$\forall (u_1^{(2)}, \dots, u_d^{(2)}) \in [0,1]^d$$

The reverse is also true in that any function that satisfies properties 1 to 3 is a copula. It is also easy to confirm that $C(1, u_1, \dots, u_{d-1})$ is a $(d-1)$ -dimensional copula and, more generally, that all k -dimensional marginals with $2 \leq k \leq d$ are copulas.

We now recall the definition of the quantile function or generalized inverse :

for a CDF, F , the generalized inverse, $F^\leftarrow$, is defined as

$$F^{\leftarrow} := \inf\{v : F(v) \geq x\}.$$

We then have the following well-known result:

Proposition 1: if $U \sim U[0, 1]$ and F_X is a CDF, then

$$P(F^{\leftarrow}(U) \leq x) = F_X(x).$$

In the opposite direction, if X has a continuous CDF, F_X , then

$$F_X(X) \sim U[0, 1].$$

Now let $X = (X_1, \dots, X_d)$ be a multivariate random vector with CDF F_X and with continuous and increasing marginals. Then by Proposition 1 it follows that the joint distribution of $F_{X_1}(X_1), \dots, F_{X_d}(X_d)$ is a copula, C_X say. We can find an expression for C_X by noting that

$$\begin{aligned} C_X(u_1, \dots, u_d) &= P(F_{X_1}(X_1) \leq u_1, \dots, F_{X_d}(X_d) \leq u_d) \\ &= P(X_1 \leq F_{X_1}^{-1}(u_1), \dots, X_d \leq F_{X_d}^{-1}(u_d)) \\ &= F_X(F_{X_1}^{-1}(u_1), \dots, F_{X_d}^{-1}(u_d)) \end{aligned} \quad (1)$$

If we now let $u_j := F_{X_j}(x_j)$ then (1) yields

$$F_X(x_1, \dots, x_d) = C_X(F_{X_1}(x_1), \dots, F_{X_d}(x_d))$$

This is one side of the famous Sklar's Theorem which is the fundamental tool of copula theory. It allows the application of copulas in statistical modeling.

Skalar theorem (1959) :

Consider a d-dimensional cumulative distribution function CDF, F , with marginals $F_1, \dots, F_d$. Then there exists a copula, C , such that:

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (2)$$

for all $x_i \in]-\infty, \infty[$ and $i = 1, \dots, d$

If F_i is continuous for all $i = 1, \dots, d$, then C is unique; otherwise C is uniquely determined only

$\text{Ran}(F_1) \times \dots \times \text{Ran}(F_d)$ where $\text{Ran}(F_i)$ denotes the range of the CDF, F_i

In the opposite direction, consider a copula, C , and univariate CDF's, $F_1, \dots, F_d$. Then F as defined in (2) is a multivariate CDF with marginals $F_1, \dots, F_d$

Example Let Y and Z be two IID random variables each with CDF, $F(\cdot)$. Let $X_1 := \min(Y, Z)$ and

$X := \max(Y, Z)$ with marginals $F_1(\cdot)$ and $F_2(\cdot)$, respectively. We then have

$$P(X_1 \leq x_1, X_2 \leq x_2) = 2F(\min\{x_1, x_2\})F(x_2) - F(\min\{x_1, x_2\})^2 \quad (3)$$

We can derive (3) by considering separately the two cases (i) $x_2 \leq x_1$ and (ii) $x_2 > x_1$.

We would like to compute the copula, $C(u_1, u_2)$, of (X_1, X_2) .

Towards this end we first note the two marginals satisfy

$$\begin{aligned} F_1(x) &= 2F(x) - F(x)^2 \\ F_2(x) &= F(x)^2. \end{aligned}$$

But Sklar's Theorem states that $C(\cdot, \cdot)$ satisfies

$$C(F_1(x_1), F_2(x_2)) = F(x_1, x_2)$$

so if we connect the pieces we will obtain

$$C(u_1, u_2) = 2\min\{1 - \sqrt{1 - u_1}, \sqrt{u_2}\}\sqrt{u_2} - \min\{1 - \sqrt{1 - u_1}, \sqrt{u_2}\}^2.$$

The Density and Conditional Distribution of a Copula

If the copula has a density, i.e. a PDF, then it is obtained in the usual manner as

$$c(u) = \frac{\partial^d C(u_1, \dots, u_d)}{\partial u_1 \dots \partial u_d}$$

It is also easy to obtain the conditional distribution of a copula. In particular, we have:

$$P(U_2 \leq u_2 / U_1 = u_1) = \frac{\partial}{\partial u_1} C(u_1, u_2)$$

This implies that the conditional CDF may be derived directly from the copula itself.

The following result is one of the most important in the theory of copulas. It essentially states that if we apply a monotonic transformation to each component in $\mathbf{X} = (X_1, \dots, X_d)$ then the copula of the resulting multivariate distribution remains the same.

2.2. Some popular copulas

Some of the most widely used bivariate copula families are: Gauss, Student's t, and the Archimedean copula. Although those copula families have multivariate expansions, but these are harder to handle and restrict the joint dependence structure of a multivariate distribution to one type. In the second section (vine

copula theory) we will build up multivariate copulas out of bivariate ones, which solve the disadvantage of lacking flexibility.

The Gaussian copula $C_\rho(u, v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$

Where Φ is the standard univariate normal cumulative distribution function(CDF) and Φ_p denotes the joint CDF.

The t-copula $C_\rho(u, v) = t_{\rho, \nu}(t_\nu^{-1}(u), t_\nu^{-1}(v))$

where ρ is a correlation matrix and t_ν^{-1} is the inverse CDF of the standard Student with ν degrees of freedom.

Archimedean copulas are built with a continuous function $\Phi : [0, 1] \rightarrow [0, \infty]$,

The Archimedean copula has the form : $C(u, v) = \Phi^{-1}[\Phi(u) + \Phi(v)]$ where $\Phi : [0, 1] \rightarrow [0, +\infty[$ is a decreasing convex function such that $\Phi(1) = 0$ and Φ^{-1} the pseudo-inverse of Φ . we put

$$\Phi(0) = +\infty \text{ if } \lim_{t \rightarrow +\infty} \Phi(t) = +\infty \text{ and } \Phi^{-1}(t) = 0 \text{ if } t \geq \Phi(0).$$

The function Φ is the Archimedean generator of C . if $\Phi(0) = +\infty$ then the copula C is strictly Archimedean in which case Φ^{-1} coincides with the inverse function of Φ . The theory of Archimedean copulas was introduced and developed by Professor Christian Genest of the University of Laval in Quebec (Genest, C. and R.J. MacKay, 1986).

The following table gives some classic families of Archimedean copulas:

table 1. the main Archimedean copula families

Families	Generator	Copula
Indépendant e	$\varphi_\theta(t) = -\ln t$	$C(u, v) = \Pi(u, v) = uv$
Clayton	$\varphi_\theta(t) = t^{-\theta} - 1$	$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}; \theta \geq 0$
Gumbel	$\varphi_\theta(t) = (-\ln t)^\theta; \theta \geq 1$	$C_\theta(u, v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right\}$
Frank	$\varphi_\theta(t) = -\log\left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}\right)$	$C_\theta(u, v) = \frac{-1}{\theta} \log\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right); \theta \neq 0$

Source : Nelson(2007)

2.3. Measures of Dependence

Understanding the dependence structure of copulas is vital to understanding their properties. There are three principal measures of dependence (Trivedi & Zimmer, 2005):

- a- The usual Pearson, i.e. linear, correlation coefficient is only defined if second moments exist. It is invariant under positive linear transformations, but not under general strictly increasing transformations. Moreover, there are many fallacies associated with the Pearson correlation.
- b- Rank correlations only depend on the unique copula of the joint distribution and are therefore invariant to strictly increasing transformations. Rank correlations can also be very useful for calibrating copulas to data.
- c- Coefficients of tail dependence are a measure of dependence in the extremes of the distributions.

The Pearson's correlation coefficient is a parametric statistic; it requires normality of the data. Another one drawback of r is that it can capture only a linear relationship. In practice, however, sampled data do not fulfill these criteria. An alternative in such situations is to use non-parametric techniques based on ranks. There are two commonly used rank based methods: Spearman's ρ , and Kendall's τ . The most important feature of the rank based methods is their invariance under monotonic transformations. The Spearman's ρ is exactly the same as the Pearson correlation r calculated on the ranks of the observations.

For a definition of Kendall's τ , we need such notions as concordant and discordant pairs.

Two pairs (X_1, Y_1) and (X_2, Y_2) are said to be concordant when

$(X_1 - X_2)(Y_1 - Y_2) > 0$, and discordant when $(X_1 - X_2)(Y_1 - Y_2) < 0$.

Definition 2 : (Kendall's τ) Let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables with joint distribution F and marginal distributions F_X and F_Y . The Kendall's τ is given by (Drouet-Mari & Kotz, 2001):

$$\tau(X, Y) = P((X_1 - X_2)(Y_1 - Y_2) > 0) - P((X_1 - X_2)(Y_1 - Y_2) < 0) \in [-1, 1]$$

And finally we can show that Kendall's tau for the pair of random variables (X, Y) is written as a function of their copula.

3. VINE COPULAS

There are a large number of bivariate copula families that allow modeling of different dependence structures. An extension of the bivariate copulas to higher dimensions is in general not trivial. It requires additional parameter restrictions that cause flexibility reduction. Besides, the extension up to the higher dimension of some copula families is even not possible

Based on the work of Joe (Joe, 1996) , Bedford and Cooke (Bedford & Cooke, Probability density decomposition for conditionally, 2001) introduced the so called pair-copula approach (vine copula) which provides very flexible tool for constructing multivariate dependency. The idea of this concept is to decompose the joint density function into a cascade of building blocks of the bivariate copulas and their conditional and unconditional distribution functions. The building blocks are called pair-copulas.

The most important advantage of this methodology lies in the ability to comprise miscellaneous dependency structures, i.e varying copula families are allowed for different marginal pairs. We present the most important topics of the pair-copula theory based on the article of (Aas, K., Czado, C., Frigessi, A., and Bakken, H , 2009)

This method uses the factorization of a joint density function as starting point,

$$f(x_1, \dots, x_n) = f(x_1) \cdot f(x_2/x_1) \cdot f(x_3/x_1, x_2) \cdot \dots \cdot f(x_n/x_1, \dots, x_{n-1}) \dots \dots \dots (1)$$
 which is a simple consequence from the definition of conditional densities

$$f(x_j/x_1, \dots, x_{j-1}) = \frac{f(x_1, \dots, x_{j-1}, x_j)}{f(x_1, \dots, x_{j-1})} \dots \dots \dots (2)$$

Where $j = 2, \dots N$.

By the equation:

$$f(x_1, \dots, x_n) = c_{12 \dots n}(F_1(x_1), \dots, F_n(x_n)) \cdot f_1(x_1) \cdot \dots \cdot f_n(x_n), \dots \dots \dots (3)$$

the second factor from (1) is :

$$f(x_2/x_1) = c_{12}(F_1(x_1), F_2(x_2)) \cdot f(x_2) \dots \dots \dots (4)$$

The third factor can be expanded by $f(x_1)$, and can be treated as a bivariate density again :

$$\begin{aligned} f(x_3/x_1, x_2) &= \frac{f(x_1, x_2, x_3)}{f(x_1, x_2)} \\ &= \frac{f(x_2, x_3/x_1)}{f(x_2/x_1)} \\ &= c_{23/1}(F_{2/1}(x_2/x_1), F_{3/1}(x_3/x_1)) \cdot f(x_3/x_1) \dots (5) \end{aligned}$$

Recall that $c_{23/1}$ is the copula density for $f(x_2, x_3/x_1)$ with marginals $F_{2/1}$ et $F_{3/1}$. Similar to (4), we can now decompose $f(x_3/x_1)$ leading to :

$$f(x_3/x_1, x_2) = c_{23/1}(F_{2/1}(x_2/x_1), F_{3/1}(x_3/x_1)) \cdot c_{31}(F_3(x_3), F_1(x_1)) \cdot f(x_3)$$

Notice that we could have expanded by $f(x_2)$ in (5), too. This would lead to a different decomposition:

$$f(x_3/x_1, x_2) = c_{13/2}(F_{1/2}(x_1/x_2), F_{3/2}(x_3/x_2)) \cdot c_{32}(F_3(x_3), F_2(x_2)) \cdot f(x_2)$$

Hence, it is obvious that there is much different decomposition for the same factorization.

In general, the copula $c_{23/1}$ depends on the conditioning value, here X_1 , but we will only consider copulae where the copula density is independent from the conditioning value of the distribution. Nevertheless, the arguments of the copula still depend on the conditioning variable (e.g $F_{1|2}(x_1|x_2)$ depends on x_2). We can generalize the factorization from above to :

$$f(x_i/v) = c_{x_i x_j / v_{-j}}(F(x_i/v_{-j}), F(x_j/v_{-j})) \cdot f(x_j/v_{-j}) \dots \dots (8)$$

For $i, j = 1, \dots, n$ and where V stands for an arbitrary set of $X_1, \dots, X_n$ with X_j in it but

not X_i . then V_{-j} denotes all the elements from V without X_j .

Iteratively applying (8) leads to a factorization of $f(x_i/v)$ in some bivariate copulae and the marginal densities $f(x_i)$, $i = 1, \dots, n$.

Together with the factorization of $f(x_1, \dots, x_n)$ from (3) we have a way to express the univariate density through a combination of the marginal densities and bivariate copulae.

Recall that if we have $f(x_i/v)$, with V representing $d > 1$ conditional variables, we need d steps and therefore d bivariate copulae in our representation. All

together, if we have n variables, we need $1 + 2 + \dots (n - 1) = n(n - 1)/2$ copulas.

Combining all this, one possible decomposition may be :

$$f(x_1, \dots, x_n) = f(x_1) \prod_{i=2}^n f(x_i/x_1, \dots, x_{i-1})$$

$$= \prod_{i=1}^n f(x_i) \cdot \prod_{i=2}^n \prod_{j=1}^{i-1} c_{ij/(j+1)\dots(i-1)}(F(x_i/x_{j+1}, \dots, x_{i-1}), F(x_j/x_{j+1}, \dots, x_{i-1}))$$

This is an example for a Pair Copula Construction. Similar to (6) and (7) there are different ways of decomposing, and hence we need to specify different copulas. This is an example of a D-Vine specification which we will explain in more detail later.

3.1. D and C vine :

Bedford and Cooke in (Bedford & Cooke, Probability density decomposition for conditionally, 2001) and (Bedford & Cooke, 2002) noticed that they can represent these pair-copula decompositions graphical with a sequence of nested trees with undirected edges, which they call a vine tree. Edges in the trees denote the indices used for the conditional copula densities. A *regular vine tree* on d variables consists of connected trees $T_1, \dots, T_{d-1}$ with nodes N_i and edges E_i for $i = 1, \dots, d - 1$, which satisfy the following

1. T_1 has nodes $N_1 = \{1, \dots, d\}$ and edges E_1 .
2. For $i = 2, \dots, d-1$ the tree T_i has nodes $N_i = E_{i-1}$.
3. Two edges in tree T_i are joined in tree T_{i+1} if they share a common node in tree T_i

Two special vine tree specifications were identified by Bedford and Cooke, one they called *drawable vine trees* or short *D-vine trees*, while the other one is called *canonical vine trees* or short *C-vine trees*. They are defined as follows

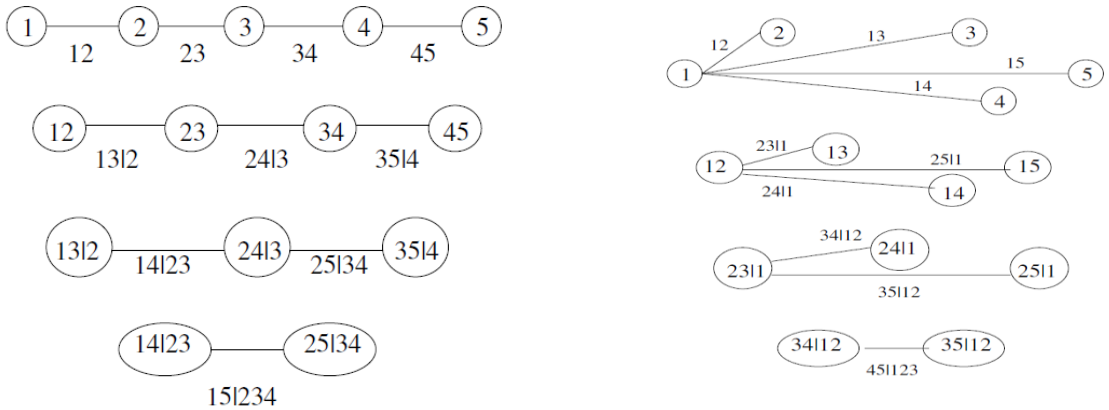
A regular vine tree is called

- *D-vine tree* if each node in $T - 1$ has at most 2 edges.
- *C-vine tree* if each tree T_i has a unique node with $d - i$ edges. The node with $d - 1$ edges in tree T_1 is called the *root*.

To build a C-vine, at the first tree, we link all the variables to one variable (Dissmann, J F; Brechmann, E C; Czado, C; Kurowika, D, 2013). Then at the second tree, a pair of variables at all others, and so on. For a D- vine , we join each

variable with its neighbor. And so on for subsequent trees, however using the segments of the previous tree as a variable. For the C-vine, decisions have to be made at each tree, but D-vine is fully defined at the first tree.

Fig. 1. example of a canonical vine (on the right) and a D vine (on the left) on 5 variables..



Source : Brechmann, E. C. and U. Schepsmeier (2013)

4. The construction of dependence structure between many financial markets:

We will try in this part to construct a model which resumes the structure of dependence between four financial markets for knowing the American market, London market, French market, and Japanese market (Cherubini, Luciano, & Vecchiato, 2004).

To accomplish this part we use four historical daily series of four stock market indices:

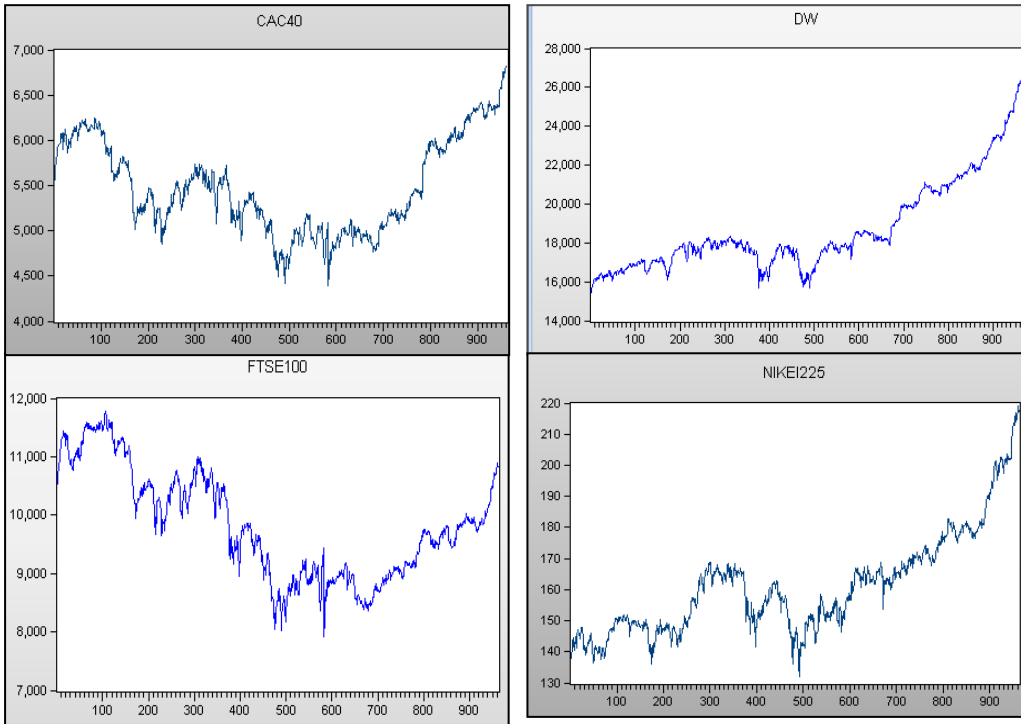
- **The CAC 40:** is the benchmark French stock market index.
- **The Dow Jones Industrial Average (DJIA),** Dow Jones, or simply the Dow, is a stock market index that measures the stock performance of 30 large companies listed on stock exchanges in the United States
- **FTSE 100 :** Financial Times Stock Exchange 100 Index, also called "Footsie" is a share index of the 100 companies listed on the London Stock Exchange with the highest market capitalization.

- **The NIKKEI 225:** or the Nikkei Stock Average , is the stock market index for the Tokyo Stock Exchange (TSE)

the dataset consists of daily time series of the four stock market indices and contain a total of 962 observations for each index. The sample covers the period from 02/02/2014 to 30/01/2018 (we noticed 81 missing observations for each series):

Their correspondent graphics are represented in the following figure:

Fig.2. time series of the four indices between 02/02/2014 and 30/01/2018



Source : produced by the authors using Eviews

The work methodology is as follows:

4.1. Margins modeling

As already mentioned before, we later want to fit copula and vine copula models to our data to explore the dependence structure among the different financial institutions. We have learned that copulae are multidimensional distribution functions with uniformly distributed marginals that are used to

describe the dependencies among several random variables with independent and identically distributed observations in $[0,1]$. As we see later, the observations in the data set of daily log returns that we obtained are neither independent (they exhibit serial autocorrelation as shown by the results of the Ljung-Box test), nor uniformly distributed, nor are they observations in $[0,1]$. Hence, we need to prepare the data accordingly to be able to later use them as the input for our copula models.

so we use the ARMA (p, q) -GARCH (p, q) models to make them iid
we selected among the ARMA-GARCH models estimated the appropriate for each series taking into account the model choice criteria.

Table 2. Model ARMA(p,q) – GARCH(p,q)specification, estimated parameters and corresponding p-values of the t -test of the fitted time series models for all indices

	Moyenne			Variance		coefficient	z-statistique
dcac	ARMA(1,1)	AR	-0,419	EGARCH(1,1)	C(3)	-0,629	-8,61
					C(4)	0,24	10,00
					C(5)	-0,147	-8,34
		MA	0,573		C(6)	0,951	142,26
ddw	ARMA(1,1)	AR	0,814	EGARCH(1,1)	C(3)	-1,00	-6,615
					C(4)	0,166	5,32
					C(5)	-0,213	-8,77
		MA	-0,826		C(6)	0,911	66,82
dftse	ARMA(1,1)	AR	0,852	TGARCH(1,1)	C	4,55E-6	65,59
					A	0,026	2,08
					B	0,239	6,94
		MA	-0,870		GARCH(-1)	0,810	37,74

Dnikei	MA(1)	MA(1)	EGARCH(1,1)	C(3)	-0,511	-6,08
				C(4)	0,124	5,24
				C(5)	-0,132	-8,52
				C(6)	0,953	115,9

Source: produced by the authors from the results obtained under EViews

4.2. Data Transformation

Now that we have obtained the standardized residuals of the time series models that fulfill the desired characteristics, we perform the second step of the data preparation. Since we need the data to be uniformly distributed on $[0,1]$, we transform the obtained residuals using the probability integral transformation (Hendrich, Copula-based Analysis of Interdependence among Companies in the Banking and Insurance sector, 2012).

In statistics, the probability integral transformation is used to convert observations of a given random variable with a specified continuous distribution to random variables having a uniform distribution. More specifically, let Z be a random variable with cumulative distribution function $F(Z)$. Then, the random variable U defined as:

$$U = F_Z(Z)$$

follows the standard uniform distribution $U[0, 1]$. This relationship holds exactly if the distribution used is the true distribution of the random variable. If, on the contrary, the distribution has been fitted to the data, the result holds approximately in large samples of data. In addition, because all cumulative distribution functions are of the form

$F: \mathbb{R} \rightarrow [0,1]$, the data are automatically transformed to the unit interval.

In our case, the obtained standardized residuals of the time series models,

$$(Z_i, t) \quad t = 1, \dots, 962, i = 1, \dots, 4,$$

are observations of random variables with either the Student-t distribution, the Normal Inverse Gaussian distribution, the Generalized Error distribution, or the Generalized Hyperbolic distribution. Hence, we can use the cumulative

distribution functions of these distributions to apply the probability integral transformation on the standardized residuals of each of the 4 variables . For $i = 1, \dots, 4$, this yields:

$$U_{i,t} = F(z_{i,t}) \forall t = 1, \dots, 962 \quad \Rightarrow \quad (U_{i,t})_{t=1, \dots, 962} \sim u[0,1]$$

We, thus, receive the desired copula data, since for e.g. the bivariate copula C the following holds:

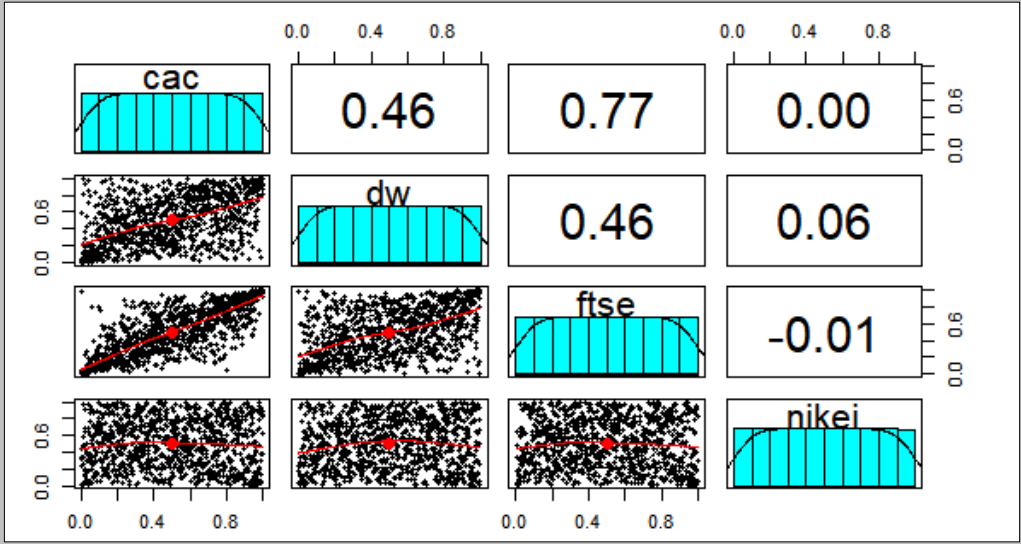
$$H(z_1, z_2) = C(F_1(z_1), F_2(z_2)) = C(u_1, u_2),$$

for H being the joint distribution function of the two random variables Z_1 and Z_2 , F_1 and F_2 being their cumulative distribution functions, respectively, and $u_i := F_i(Z_i), i = 1, 2$.

We include the histograms of the obtained copula data in Figure B.4 to show that these are in fact uniformly distributed. Peaks around the value of 0.5 for several of the companies again result from the high number of zeros in these time series.

$Z_{(i,t)}, i = 1, \dots, 4; t = 1, \dots, 962$, are the series of residuals resulting from the estimation of ARMA(p,q)-GARCH(p,q) models. Using R we can transform the data

Fig. 3 .Visualization of transformed data using scatter plot (bottom) and Kendall's tau calculated between pairs of variables (top)



Source : produced by the authors using R software

Now that we have transformed the data into pseudo-data, we take the next step to multiple modeling via copulas.

4.3. Method of adjusting a regular vine (r-vine) :

After getting transformed residue of the estimated series (ARMA-GARCH) to obtain uniform margins in the interval $[0, 1]$, we propose the following methodology to fit a regular "R-Vine":

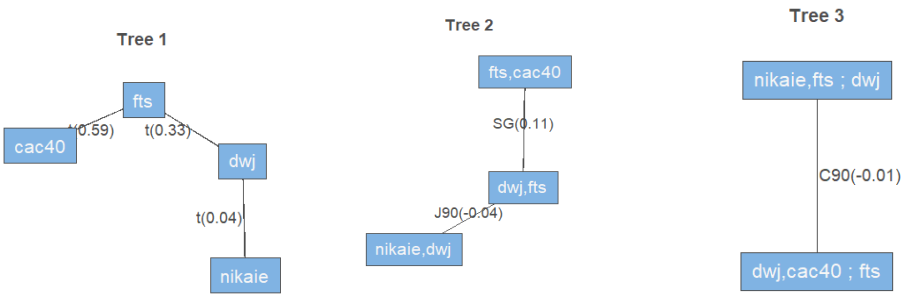
- ✓ Our matrix of residue is composed of four columns (corresponding to the number of indices) and N rows (with $N = 962$ in our study).
- ✓ We choose a set of families of bivariate copula. We have chosen as families of copulas, the elliptical copulas: the Gaussian (1) and the Student (2), the Archimedean copulas: Clayton (3), Gumbel (4), Frank (5) and Joe's copula (6) as well as their rotations at 90, 180 and 270 degrees. The numbers indicated in parenthesis in front of each copula represent its identifier in the "package: VineCopula" of the software R.
- ✓ Next, we use two functions from the "package: VineCopula" in our program to estimate, respectively, the "R-Vine" model and the tree graph. These functions were developed by Dissmann et al. (2013) and Brechmann (2013). A brief description of these functions is given below:
 - 1- « *RVineStructureSelect ()* » : allows the selection of the structure of the pairs of copula according to the predefined families of copula. For example,
library (VineCopula):
MODEL <- RVineStructureSelect (data, familyset, type=0, selectioncrit="AIC", indeptest=FALSE, level=0.05, trunclevel, progress=TRUE, weights)}.
 - 2- *RVineTreePlot* produces plots of trees and desired parameters of copulas. For example,
library (VineCopula): {RVineTreePlot (data, MODEL=MODEL, method="mle", tree="ALL", edge.labels)}

4.4. R-vine estimation

The copula families chosen for the application in this part are: Gaussian, Student's copula, Clayton's copula, Frank's copula, Gumbel's copula and Joe's.

Our data matrix contains the pseudo-data of four stock market indices. The dimension determines the total number of trees to be built. The trees are illustrated in Figures 4. We also find the names of the copulas and the empirical Kendall coefficients associated with the copulas on the arcs of the trees.

Fig. 4. the trees obtained from the three estimation steps



Source: developed by the authors using R software

The copulas of the first tree are the functions that best fit the graphs between (CAC40 and FTSE100), (FTSE100 and DW J) and (DW J and NIKEIE 225). The AIC criterion is used to determine which copulas are the most appropriate between each pair of variables. You can get an intuition by looking at the graph (scatter-plot). If the graph shows an asymmetry from the diagonal, then an unchangeable copula would be appropriate. Also, if the graphs suggest a greater dependence between big values than between small values; elliptical copulas are therefore to be avoided, since they are radially symmetrical. To choose the copulas of subsequent trees, we must calculate the conditional observations, then carry out the same procedure. Fortunately, this method is already directly implemented in the *RVineStructureSelect* function.

The results of the vine estimate are presented below :

Table 3. The results of the vine estimate

Pair	Copula	Parameter (s)
tree 1		
CAC40 – FTSE100	Student	$\rho = 0,800 ; \vartheta = 3,660$
FTSE100 – DWJ	Student	$\rho = 0,493 ; \vartheta = 3,120$
DWJ – NIKEI 225	Student	$\rho = 0,058 ; \vartheta = 4,628$
tree 2		
CAC40,DWJ/FTSE100	Survival Gumbel	$\theta = 1,921$
FTSE100,NIKEI225/DWJ	Rotated Joe (90 degrees)	$\theta = -1,081$
Arbre 3		
CAC40, NIKEI225/DWJ, FTSE100	Rotated Clayton Copula	$\theta = -0,0195$

Source: produced by the authors from the results obtained.

Note that the best model according to the criterion of AIC is d-vine decomposition.

In the world of stock markets, the concept of dependence between different assets constitute the foundation of modern portfolio management (asset allocation, risk management, etc.). Understanding the dependency structure between assets is important not only for potential reconciliations of various asset returns, but also for risk diversification and portfolio construction.

However, multivariate financial data generally exhibit a complex pattern of dependency.

5. CONCLUSION

Over the past decades, and not just during the last financial crisis, risk management has become increasingly important. Several risk measures have been introduced. So far, the most widely used is “value at risk” (VaR).

For the aggregation of different risks, it is highly desirable to model these risks conjointly.

At first, multivariate distributions like Gaussian or Student were used for this task. But since the structure of dependency between different risks is not homogeneous, the need for more flexible models was obvious. Copula models have bridged this gap. Although bivariate copulas were soon well explored, multivariates were more

difficult to construct and still left a lack of flexibility. This has been solved with so-called pair-copula constructions.

It's about building large-dimension copulas from two-variable building blocks. One of the approaches is the regular vines (R-vines). They were introduced by Bedford & Cooke (2001) and discussed at length over the past decade (see Kurowicka & Cooke (2006) or Aas et al. (2009)). There are also implementations for the R software package (Schepsmeier et al. 2012).

In this paper we have tried to apply these models to four historical stock market indices: the CaC40, the Nikkei225, FTSE100, DWJ; to achieve a structure that reflects the dependencies between these markets in order to be able to quantify the risks.

6. Bibliography List :

1. Aas, K., Czado, C., Frigessi, A., and Bakken, H . (2009). *Pair-copula constructions of multiple dependence*. Insurance Mathematics and Economics, 44(2):182–198.
2. Bedford, T., & Cooke, R. (2001). *Probability density decomposition for conditionally*.
3. Bedford, T., & Cooke, R. (2002). *Vines - a new graphical model for dependent random variables*. Annals of Statistics 30, 1031-1068.
4. Brechmann, E C; Czado, C. (2011). *Risk management with high-dimensional vine copulas: An analysis of the Euro Stoxx 50*. Statistics & Risk Modeling, 30 (4), 307-342.
5. Brechmann, E C; Schepsmeier , U. (2013). *Modeling Dependence with C- and D-Vine Copulas: The R Package CDVine*. Journal of Statistical Software, 52 (3), 1-27. <http://www.jstatsoft.org/v52/i03/>.
6. Cherubini, U., Luciano, E., & Vecchiato, W. (2004). *Copula Methods in Finance*. New York: John Wiley.
7. Cont, R. (2001). *Empirical proeprties of asset returns: Stylized facts and statistical issues*. Quantitative Finance 1.
8. Dissmann, J F; Brechmann, E C; Czado, C; Kurowika, D. (2013). *Selecting and estimating regular vine copulae and application to financial returns*. Computational Statistics & Data Analysis, 59 (1), 52-69.
9. Drouet-Mari, D., & Kotz, S. (2001). *Correlation and Dependence*. London: Imperial College Press.

10. Genest, C. and R.J . MacKay. (1986). *Copules Archimédiennes et familles de Lois bidimensionnelles dont les marges sont données*. The Canadian Journal Of Statistics,14,145-159.
11. Hendrich, K. (2012). *Copula-based Analysis of Interdependence among Companies in the Banking and Insurance sector*. munich: technical university of munich.
12. Joe, H. (1996). *Families of m-variate distributions with given margins and $m(m - 1)/2$ bivariate dependence parameters*. In L. Rüschendorf, B. Schweizer, and M. D. Taylor.
13. Joe, H. (1997). *Multivariate Models and Dependence Concepts*. London: Chapman & Hall.
14. Kurowicka, D., & Cooke, R. M. (2006). *Uncertainty Analysis with High Dimensional Dependence Modelling*. Chichester: John Wiley.
15. LONGIN, F. (1996). *The asymptotic distribution of extreme stock market returns*. Journal of Business, volume 69, N°3.
16. Nelson, R. (2007). *un introduction to copulas*. New York: Springer.
17. Schepsmeier, U; Stoeber, J. (2012). *Derivatives and Fisher information of bivariate copulas*. Statistical Papers, 55 (2), 525-542.
<http://link.springer.com/article/10.1007/s00362-013-0498-x>.
18. Trivedi, P. K., & Zimmer, D. M. (2005). *Copula Modeling: An Introduction for Practitioners* (Vols. vol 1, no 1). Foundations and Trends R in Econometrics .

