



## Création et utilisation d'une copule à trois paramètres pour le dimensionnement des ouvrages côtiers

Olivier ORCEL<sup>1</sup>, François ROPERT<sup>1</sup>, Philippe SERGENT<sup>1</sup>

<sup>1</sup> CEREMA, DtecEMF, 134 rue de Beauvais, 60280 Margny lès Compiègne CS 60321, France.

*Olivier.Orcel@cerema.fr ; Francois.Ropert@cerema.fr ; Philippe.Sergent@cerema.fr*

### Résumé :

Le dimensionnement des structures soumises aux actions de la houle nécessite d'estimer précisément les fréquences d'occurrences exceptionnelles. Elles sont généralement déterminées à l'aide de paramètres principaux : le niveau de la mer, la hauteur de la houle  $H$  et sa période  $T$ . Les effets combinés du niveau de la mer, de la houle, et de la période sont prises en compte par des corrélations linéaires qui ne représentent pas toute la complexité de la réalité, ce qui peut conduire à des erreurs de dimensionnement dommageables. Cet article présente une méthode innovante pour l'analyse de probabilités conjointes : l'utilisation de l'outil mathématique copule.

**Mots-clés :** Copule, Hydrodynamique côtière, Houle, Surcote, Période.

### 1. Introduction

L'enjeu de ce travail est de contribuer à remédier à l'insuffisance des approches courantes de représentation de la dépendance entre les divers paramètres caractérisant les états de la mer tels que notamment la hauteur de la houle, la période et la surcote :  $H, T$  et  $s$  (SERGENT *et al.*, 2014 ; DEFRA, 2005). Actuellement, dans certains ouvrages de référence tels que le guide enrochements (CIRIA *et al.*, 2009), c'est un simple coefficient multiplicateur à appliquer au produit des fréquences relatives à la hauteur de la houle et au niveau de la mer qui permet de représenter la dépendance entre ces deux paramètres. Or cet usage introduit forcément une erreur significative dans les estimations des périodes de retour des événements à considérer et donc dans le choix des situations de projet qui conditionnent le dimensionnement même des structures.

L'aspect novateur de la démarche repose sur le recours à la théorie des copules mis au point par le mathématicien (SKLAR, 1959). Il s'agit d'un outil mathématique dont les compagnies d'assurance et les financiers ont commencé à faire un large usage depuis la fin des années 90. Les copules sont des fonctions permettant de multiples représentations de la dépendance de plusieurs variables aléatoires corrélées à des degrés divers. Nous avons construit une copule trivariée originale et performante à partir de deux copules bivariées de Clayton, avec deux coefficients de corrélation.

Nous définissons l'outil mathématique copule et expliquons comment obtenir la meilleure copule bivariée. Pour la copule de Clayton sélectionnée, nous représentons

graphiquement l'évolution de l'erreur en fonction du taux de corrélation. Nous construisons des fonctions trivariées à partir de copules bivariées et montrons que ce sont des copules mathématiquement valides. Enfin, nous représentons les courbes d'isoprobabilités de dépassements conjoints des hauteurs de houle et des surcotes pour une période de retour de 10, 100 et 1000 ans avec les copules de Gumbel et de Clayton.

## **2. Présentation des copules**

La copule est un outil permettant de caractériser la dépendance entre plusieurs variables aléatoires là où en général les corrélations linéaires ne sont pas en mesure de les représenter fidèlement (CHARPENTIER, 2014).

### 2.1 Définition d'une copule

Une copule de deux variables  $C : [0,1]^2 \rightarrow [0,1]$  est une fonction de répartition, qui vérifie les trois conditions suivantes :

$$\begin{aligned} C(u_1, 0) &= C(0, u_2) = 0 \\ C(u_1, 1) &= u_1 \text{ et } C(1, u_2) = u_2 \quad \forall u_1 \in [0,1] \text{ et } \forall u_2 \in [0,1] \\ C &\text{ est une fonction 2 croissante, } \forall 0 \leq u_1 \leq u_2 \leq 1 \text{ et } \forall 0 \leq v_1 \leq v_2 \leq 1 \\ C(v_1, v_2) &+ C(u_1, u_2) - C(u_1, v_2) - C(v_1, u_2) > 0 \end{aligned} \quad (1)$$

La copule est un outil relativement innovant de modélisation de la structure de dépendance de plusieurs variables aléatoires. La notion de copule a été introduite en 1959 (SKLAR, 1959) motivée par les travaux de (FRECHET, 1950). Les copules sont aussi appelées "fonction de dépendance" (DEHEUVELS, 1979). La dépendance entre les variables aléatoires réelles  $X_1, \dots, X_n$  est totalement décrite par leur fonction de répartition jointe  $F$ . Considérons les variables aléatoires  $X_1, \dots, X_n$  de fonctions de répartition continue  $F_1, \dots, F_n$ , nous pouvons introduire la copule  $C$  :

$$\begin{aligned} F(x_1, \dots, x_n) &= P[X_1 < x_1, \dots, X_n < x_n] \\ &= C[F_1(x_1), \dots, F_n(x_n)] \end{aligned} \quad (2)$$

### 2.2 Obtention de la meilleure copule

Deux méthodes s'offrent à nous pour choisir une copule : la méthode de l'erreur (se reporter au paragraphe 2.4), et la méthode du Maximum de Vraisemblance (se reporter au paragraphe 2.3). La méthode, qui semble la plus simple, est la méthode de calcul Log-Vraisemblance. Elle donne pour chaque copule le coefficient de corrélation le plus approprié. On retient ensuite comme copule le plus pertinent, celle qui présente la plus grande Log Vraisemblance:

$$\text{Max Log}(\prod_{i=1}^n f(x_i; w)) \quad (3)$$

### 2.3 Méthode de Maximum de Vraisemblance

Nous avons calculé la Log-Vraisemblance pour deux échantillons de valeurs différentes, l'un concernant le site de Saint-Malo (5888 valeurs) et l'autre le site du Havre (3040 valeurs). Nous avons voulu comparer les pertinences de neuf copules différentes. Nous avons appliqué, pour chaque copule, la Méthode du Maximum de Vraisemblance pour caler le coefficient de corrélation propre à chaque copule, puis comparé l'ensemble des Log-Vraisemblance ainsi obtenus.

Tableau 1. Maximum de Vraisemblance pour 9 différentes copules.

Copules	Coefficient de Corrélation	Coefficient de Corrélation	Coefficient de Corrélation	Log- Vraisemblance	Log- Vraisemblance
Echantillon		Saint-Malo	Havre	Saint-Malo	Havre
Gumbel	$[1, \infty[$	1.09	1.29	51.53	184.7
Clayton	$]0, \infty[$	0.384	0.738	<b>200.3</b>	<b>387.2</b>
Gauss	$[-1, +1]$	0.22	0.42	149.2	236.7
Franck	$\neq 1$	1.25	2.67	123.2	270.6
Student	$[-1, +1]$	0.22	0.42	157.5	303.3
Plackett	$]0, \infty[$	1.88	3.58	127.4	277.2
Joe	$[1, \infty[$	1.033	1.208	4.14	76.02
AMH	$[-1, +1]$	0.71	0.96	196.2	375.5
Galambos	$]0, \infty)$	0.396	0.536	41.37	174.6

On retient donc la copule de Clayton, qui a la plus grande Log-Vraisemblance, dont l'expression est :

$$C(u,v) = (u^x + v^x - 1)^{-1/x} \quad (4)$$

x : coefficient de corrélation attaché à cette copule.

### 2.4 Méthode de l'erreur

Afin d'illustrer l'incidence du coefficient de corrélation sur l'erreur, nous représentons sur la figure 1 l'évolution de celle-ci pour la copule de Clayton en fonction de ce coefficient. Le coefficient de corrélation retenu correspond au minimum de l'erreur.

F<sub>mes</sub>(H,s) : fréquence de dépassement conjointe mesurée (hauteur de houle, surcote)

F<sub>cal</sub>(H,s) : fréquence de dépassement conjointe calculée avec la copule (hauteur de houle, surcote).

e : erreur entre les fréquences conjointes mesurées et calculées de la houle et de la surcote.

e<sub>min</sub> : erreur minimum.

$$e = \frac{F_{cal} - F_{mes}}{F_{cal}} \quad (5)$$

Tr : taux d'erreur

$$Tr = \exp(-e) - 1 \quad (6)$$

## Thème 1 – Hydrodynamique côtière

Nous représentons sur la figure 2 les fréquences mesurées et calculées de dépassement conjointes des hauteurs de houle et des surcotes, pour l'échantillon du Havre (3040 couples de valeurs). La corrélation idéale s'obtient lorsque le nuage de points suit parfaitement la droite  $y = x$ , ce qui est le cas avec la copule de Clayton, mais pas avec la copule de Gumbel, en particulier pour les fréquences rares.

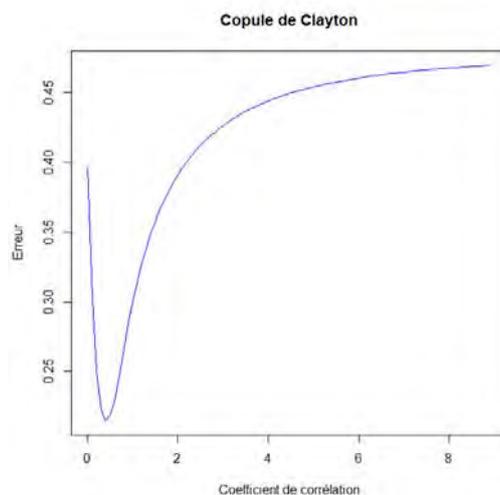


Figure 1. Evolution de l'erreur en fonction du coefficient de corrélation pour la copule de Clayton.

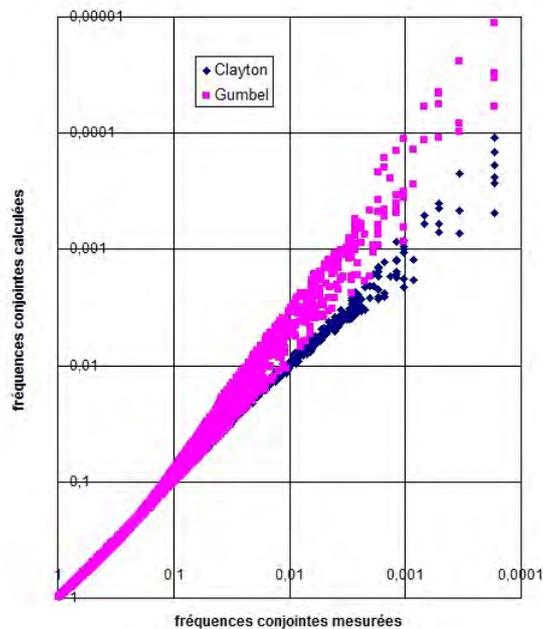


Figure 2. Comparaison des fréquences mesurées et calculées de dépassement conjointes des hauteurs de houle et de surcote pour les copules de Gumbel et de Clayton.

### 3. Construction d'une copule trivariée

Il est difficile d'obtenir des copules d'ordre n à partir de copules d'ordre n - 1 (NELSEN, 2006). Nous allons construire une fonction à trois variables, puis montrer que c'est une copule.

#### 3.1 Construction de fonctions à trois variables

Notations :

C : Copule à trois variables

C<sub>1</sub>, C<sub>2</sub> : Copule bivariée

φ : Fonction génératrice de copule (=  $\frac{1}{\theta}(t^{-\theta}-1)$  pour Clayton )

F<sub>H</sub> : Fonction de répartition de hauteur de houle

F<sub>T</sub> : Fonction de répartition de période de houle

F<sub>s</sub> : Fonction de répartition de surcote

##### 3.1.1 Construction d'une fonction à trois paramètres à partir de copules bivariées

$$C(F_H, F_s, F_T) = C_1(C_2(F_H, F_s), F_T) \quad (7)$$

##### 3.1.2 Construction de copules archimédiennes

Définition : Soit φ une fonction décroissante convexe sur [0,1]→[0,∞[ telle que φ(1)=0 et φ(0) = ∞. On appellera copule archimédienne stricte de générateur φ la copule définie par : C(u,v) = φ<sup>-1</sup>[φ(u) + φ(v)], (u,v) ∈[0,1]<sup>2</sup>. Nous pouvons généraliser cette construction pour obtenir des copules multivariées avec un paramètre unique de corrélation. Dans notre cas, nous aurons :

$$C(F_H, F_s, F_T) = \phi^{-1}[\phi(F_H) + \phi(F_s) + \phi(F_T)] \quad (8)$$

#### 3.2 Vérification des propriétés des copules

On ne connaît pas de méthode générale pour construire des copules d'ordre plus élevé à partir de copules d'ordre inférieur (DURRLEMAN, 2010). En général, C<sub>1</sub>(u,C<sub>2</sub>(v,w)) et C<sub>1</sub>(C<sub>2</sub>(u,v),C<sub>3</sub>(w,x)), (où C<sub>1</sub>, C<sub>2</sub> et C<sub>3</sub> sont des copules d'ordre 2) ne sont pas des copules. Il convient donc de vérifier les propriétés des copules. Cependant, A. Charpentier (CHARPENTIER A., 2014) explique que pour les copules archimédiennes multiples :

$$C(u_1, \dots, u_d) = \phi_1^{-1}[\phi_1[\phi_2^{-1}(\phi_2[\dots + \phi_{d-1}^{-1}(\phi_{d-1}(u_1) + \phi_{d-1}(u_2)) + \dots + \phi_2(u_{d-1})] + \phi_1(u_d)] \quad (9)$$

avec φ<sub>i</sub> des générateurs de copules, C'est une copule si φ<sub>i</sub> ∘ φ<sub>i-1</sub><sup>-1</sup> est l'inverse d'une transformée de Laplace. Il fait également remarquer que pour les copules de Clayton, une condition suffisante pour que C soit une copule est que les coefficients de corrélations relatifs aux copules soient croissants et supérieurs à 0. Ceci est le cas pour C<sub>1</sub>(C<sub>2</sub>(F<sub>H</sub>,F<sub>T</sub>),F<sub>s</sub>) et prouve que la fonction construite à partir de deux copules bivariées de Clayton est une copule.

#### 4. Choix de la meilleure copule trivariée

Des trois combinaisons  $C_1(C_2(F_H, F_s), F_T)$ ,  $C_1(C_2(F_T, F_s), F_H)$  et  $C_1(C_2(F_H, F_T), F_s)$ , la copule trivariée  $C_1(C_2(F_H, F_s), F_T)$ , qui a le taux d'erreur le plus faible (cf tableau 4), s'avère être la plus précise. Nous présentons donc son obtention.

##### 4.1 Obtention de $F_{HT}$ avec le coefficient de corrélation

Nous déterminons la Log-Vraisemblance pour  $F_H$  et  $F_s$  et obtenons la copule de Clayton comme meilleure copule avec son coefficient de corrélation de 2.37.

Tableau 2. Log-Vraisemblance pour les différentes copules bivariées entre les paramètres  $F_H$  et  $F_T$ .

<i>Copule</i>	<i>Coefficient de corrélation</i>	<i>Log-Vraisemblance</i>
<i>Gumbel</i>	1.59	1059
<i>Clayton</i>	2.37	<b>1565</b>
<i>Gauss</i>	0.77	1369
<i>Franck</i>	7.27	1333
<i>Student</i>	0.77	1404
<i>Plackett</i>	15.64	1349
<i>Joe</i>	2.058	651.5
<i>Galambos</i>	1.25	1038

##### 4.2 Obtention de $F_{HT_s}$ à partir de $(F_{HT}, F_s)$

Nous pouvons donc agréger  $F_H$  et  $F_T$  par la copule de Clayton avec un coefficient de corrélation de 2.37 et obtenir  $F_{HT}$ . On détermine la Log-Vraisemblance pour  $F_{HT}$  et  $F_s$  et obtenons la copule de Clayton comme meilleure copule, avec un coefficient de corrélation de 0.56.

Tableau 3. Log-Vraisemblance pour les différentes copules bivariées entre les paramètres  $F_{HT}$  et  $F_s$ .

<i>Copule</i>	<i>Coefficient de Corrélation</i>	<i>Log-Vraisemblance</i>
<i>Gumbel</i>	1.25	119.7
<i>Clayton</i>	0.56	<b>289.4</b>
<i>Gauss</i>	0.365	195.4
<i>Franck</i>	2.08	155.7
<i>Student</i>	0.35	214.3
<i>Plackett</i>	2.84	164.7
<i>Joe</i>	1.72	35
<i>Galambos</i>	0.5	111.3

#### 5. Résultats pour les copules trivariées

Sous réserve que les fonctions construites trivariées soient des copules, nous présentons les différents taux d'erreur pour le Havre ( $Tr = \exp(e_{\min}) - 1$ ) (cf tableaux 4, 5 et 6).

Ces résultats soulignent clairement l'apport de la construction d'une copule trivariée à l'aide de deux copules de Clayton  $C_1$  et  $C_2$  et de deux coefficients de corrélation. En effet, nous obtenons un taux d'erreur de 3.77%. Avec une copule de Gumbel et un seul coefficient de corrélation, nous obtenons un taux d'erreur très médiocre (cf tableau 6).

Tableau 4. Taux d'erreur pour les différentes combinaisons de paramètres avec deux coefficients de corrélation.

Copule de Clayton	$C_1(C_2(F_H, F_s), F_T)$	$C_1(C_2(F_T, F_s), F_H)$	$C_1(C_2(F_H, F_T), F_s)$
Taux d'erreur	6.93%	4.7%	3.77%

Tableau 5. Comparaison des taux d'erreur pour les copules de Clayton et de Gumbel avec deux coefficients de corrélation.

$C_1(C_2(F_H, F_T), F_s)$	Copule de Clayton	Copule de Gumbel
Taux d'erreur	3.77%	22.18%

Tableau 6. Comparaison des taux d'erreur pour les copules de Clayton et de Gumbel avec un coefficient de corrélation.

$C(F_H, F_s, F_T)$	Copule de Clayton	Copule de Gumbel
Taux d'erreur	8.8%	168.99%

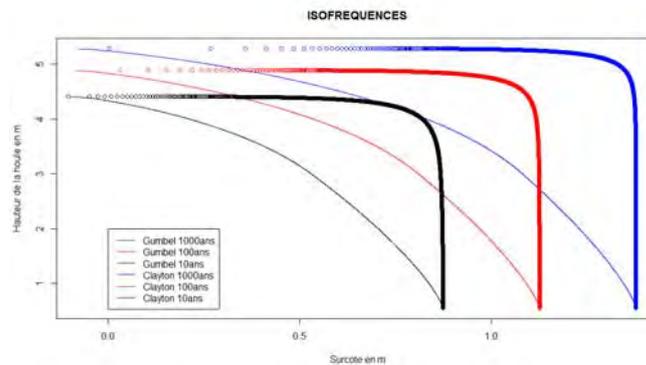


Figure 3. Courbes de probabilités conjointes de dépassement des hauteurs de houle et des surcotes pour des périodes de retour de 10, 100 et 1000 ans avec copule de Clayton pour les points épais (corrélation 0.738), et de Gumbel pour lignes (corrélation 1.29).

## 6. Probabilités conjointes de dépassement

Nous constatons que les probabilités conjointes de dépassement des hauteurs de houle et des surcotes pour différentes périodes de retour obtenues avec les copules de Clayton et de Gumbel divergent sensiblement. En effet, les résultats obtenus pour les isofréquences avec la copule de Gumbel sont imprécis, en particulier pour les fréquences rares comme le montre la figure 2. Il faut préciser que le calage est effectué sur l'ensemble de l'échantillon. En tronquant l'échantillon pour des valeurs de fréquences jointes inférieures à 0.01, nous aurions obtenu un coefficient de corrélation de fréquences

jointes de Gumbel beaucoup plus important. Ceci aurait fait converger les courbes de Gumbel et Clayton sur la figure 3. En revanche, le coefficient de corrélation de Clayton s'avère bien plus stable pour l'ensemble de l'échantillon, comme le montre la figure 2.

## **7. Conclusion**

Le manque de précision dans la prise en compte de la dépendance entre la hauteur de houle, la période de houle et la surcote conduit à des erreurs dans le dimensionnement des ouvrages côtiers. L'innovation de ce travail consiste à introduire l'outil mathématique copule pour corréler les différents paramètres hauteur de houle, période de houle et surcote. Nous avons construit une copule trivariée originale, très performante à partir de deux copules bivariées de Clayton et de deux coefficients de corrélation. Enfin, nous montrons son intérêt. Son taux d'erreur descend à 3.8%, alors qu'avec une copule de Gumbel avec un seul coefficient de corrélation, il s'établit à 168%.

Les perspectives de ce travail sont la prise en compte d'un quatrième paramètre, la direction de la houle et d'intégrer dans les calculs de statistiques la marée astronomique afin d'obtenir les fréquences de retour d'événements définis par les quatre paramètres que sont la hauteur de la houle, sa période, sa direction et le niveau de la mer.

## **8. Références bibliographiques**

- CHARPENTIER A. (2014). *Approches Statistiques des Risques*, Editions Technip, Société Française de Statistique. Copules et risques multiples chap. 6, pp 171-172.
- CIRIA, CUR, CETMEF (2009). *Guide Enrochement. L'utilisation des enrochements pour les ouvrages hydrauliques*. Version française du Rock Manual, P09-01, CETMEF, Compiègne.
- DEFRA (2005). *Use of Joint Probability Methods in Flood Management*. A Guide to Best Practice. R&D 333. Technical Report FD2308/TR2 - DEFRA/Environment Agency.
- DEHEUVELS P. (1979). *La fonction de dépendance empirique et ses propriétés - Un test non paramétrique d'indépendance*, Académie Royale de Belgique Bulletin de la Classe des Sciences 5e Série, Vol. 65, pp 274-292.
- DURRLEMAN V. (2010). *Les fonctions copules. Définition, exemples et propriétés*. Cours de l'Ecole Polytechnique. 28 p.
- FRECHET M. (1950). *Sur les tableaux de corrélation dont les marges sont données*. Ann. Univ. Lyon. Sect. A 9, pp 53-77.
- NELSEN R.B. (2006). *An Introduction to Copulas*. Springer Series in Statistics. 269 p.
- SERGENT P., PREVOT G., MATTAROLO G., BROSSARD J., MOREL G., MAR F., BENOIT M., ROPERT F., KERGADALLAN X., TRICHET J., MALLET P. (2014). *Adaptation of Coastal Structures to Mean Sea Level Rise*. La Houille Blanche, n° 6, pp 54-61. <http://dx.doi.org/10.1051/lhb/2014063>
- SKLAR A. (1959). *Fonction de répartition à n dimensions et leurs marges*. Publications de l'Institut de Statistique de l'Université Paris, pp 229-231.