

ESTIMATION DE L'ALÉA HYDROMÉTÉOROLOGIQUE PAR UNE MÉTHODE PAR SIMULATION (SHYPRE)

Philippe Cantet ¹ & Patrick Arnaud ¹

¹ *IRSTEA, 3275 route de Cézanne - CS 40061 - 13 182 AIX-EN-PROVENCE Cedex 5
philippe.cantet@irstea.fr*

Résumé. La méthode SHYPRE est une méthode de prédétermination de l'aléa hydro-météorologique, basée sur la simulation des processus. Elle met en oeuvre un générateur stochastique de pluies horaires couplé à une modélisation conceptuelle de la transformation pluie-débit. Le but de la méthode est de simuler un grand nombre de scénarios de crues afin d'estimer empiriquement leurs probabilités de dépassement. Les crues simulées diffèrent les unes des autres uniquement par la multitude d'évènements pluvieux générés par le modèle de pluies horaires. Celui-ci est basé sur une décomposition du signal pluie à l'aide de variables descriptives (durée d'une période pluvieuse/sèche, nombre d'averses, leurs volumes et leurs formes, . . .). La principale difficulté rencontrée dans l'élaboration du générateur a été la prise en compte des dépendances entre les variables descriptives afin de reproduire au mieux les pluies extrêmes de toutes durées. Celle-ci a principalement été modélisée avec les copules. Cet outil a permis de modéliser la dépendance temporelle des volumes des averses et celle entre le volume et la durée d'une averse menant à une réelle amélioration dans la prédétermination des pluies extrêmes. Dans le but de connaître l'aléa pluvial en tout point de la France, le nombre de paramètres du générateur de pluies a été limité à trois qui sont uniquement issus de l'information pluie journalière (plus abondante que l'information horaire). La régionalisation des paramètres par des méthodes d'interpolation spatiale (KED, régression multiple, . . .) permet alors d'estimer des quantiles de pluies extrêmes et de crues sur l'ensemble du territoire français.

Mots-clés. Modèle stochastique, pluies extrêmes, copules, interpolation spatiale

Abstract. The SHYPRE method is a hydro-meteorological hazard estimation method, based on the simulation process. It implements an hourly rainfall generator coupled with a conceptual rainfall-runoff model. This method aims to simulate a large number of flood scenarios in order to empirically estimate their exceedance probabilities. All generated floods are different only due to the numerous rainy events furnished by the hourly rainfall generator. This generator is based on a decomposition of the rainfall signal with descriptive variables (duration of a wet/dry period, number of rainstorms, their depths and their shapes, . . .). The main difficulty met to elaborate the rainfall model was the taking into account the dependence between the descriptive variables in order to accurately reproduce extreme rainfall of different duration. This is essentially modeled with copulas. This tool enables to model the temporal dependence between the depths of successive

storms and the one between the depth and duration of a rainstorm. To obtain a knowledge of rainfall hazard at any point in France, the number of rainfall model parameters was limited to three coming from a daily information (more widespread than hourly information). The regionalization of these parameters by spatial interpolation methods (EDK, multiple linear regression, ...) enables to estimate extreme rainfall and flood quantiles on the whole French territory.

Keywords. Stochastic model, extreme rainfall, copulas, spatial interpolation

1 Le générateur stochastique de pluies horaires : SHYPRE

SHYPRE est un modèle événementiel de simulation d'hydrogrammes basé sur un générateur de pluies horaires couplé à un modèle pluie-débit. Il est développé à l'Irstea d'Aix-en-Provence depuis une dizaine d'années. Pour plus d'information, un guide de SHYPRE a été rédigé Arnaud et Lavabre (2010). Les renseignements sur la transformation pluie-débit sont mentionnés dans Arnaud et al (2014).

Les variables descriptives du générateur :

Le générateur stochastique de pluies horaires repose au préalable sur l'analyse de la structure temporelle interne des événements pluvieux observés au pas de temps horaire, conduisant à la définition de 9 variables descriptives (voir Fig. 1). Afin de séparer les événements plutôt orageux, des événements moins intenses et de durées plus longues, on distingue deux saisons ("été" de juin à novembre et "hiver" décembre à mai). Les variables du modèle sont alors estimées indépendamment sur chaque saison.

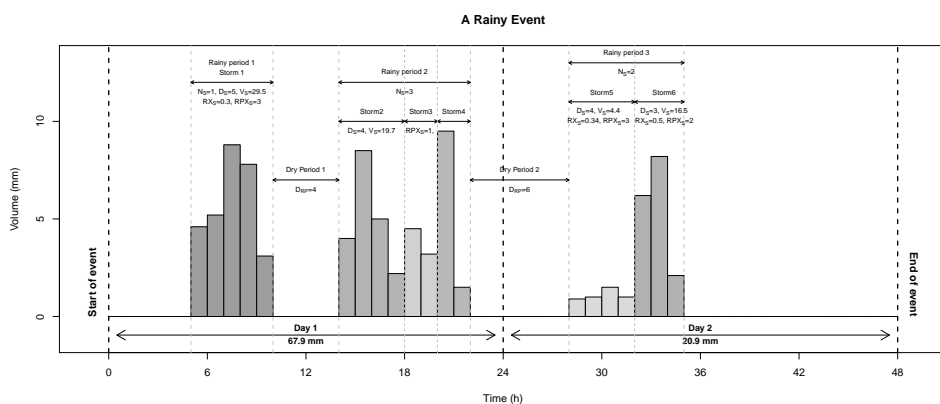


FIGURE 1 – Illustration d'un évènement pluvieux (88.8mm en 2 jours) avec 3 périodes pluvieuses et 6 averses (dont 2 majeures : Storm1 et Storm2)

Simulation et estimation des quantiles de pluies :

Chaque variable descriptive est caractérisée par une loi de probabilité commune à l'en-

semble des stations. Par exemple, la loi exponentielle a été choisie pour générer les volumes des averses, leurs durées sont générées par une loi de Poisson. Seuls les paramètres de ces lois théoriques permettent de différencier les climats et sont estimés à l'aide des pluies observés sur une station donnée. 20 paramètres permettent de caler entièrement le modèle de pluie sur chacune des deux saisons. Parmi ces 20 paramètres, certains n'ont que peu d'impacts sur l'estimation des quantiles extrêmes. Celle-ci est faite à l'aide de la simulation de longues chroniques de pluies (méthode de Monte-Carlo). Par exemple, le quantile centennal (période de retour 100 ans) est estimé en générant une chronique équivalente à 100 000 ans (il correspond à la 1000e statistique d'ordre de la série simulée).

2 Modélisation de la dépendance

Dans SHYPRE, la modélisation de la dépendance a été effectuée à l'aide des copules. Elles ont permis de modéliser, dans un premier temps, la dépendance des volumes des averses, et celle entre la durée et le volume des averses.

Mesure de dépendance :

Pour une dépendance non linéaire, l'utilisation de coefficients de corrélation basés sur les rangs est fortement recommandée comme par exemple le tau de Kendall qui est défini comme suit :

Soient X et Y deux variables aléatoires continues. On suppose disposer un n -échantillon de réalisations de X et de Y : $\mathbf{X} = (X_i)_{i=1..n}$ et $\mathbf{Y} = (Y_i)_{i=1..n}$. Soit R_i (resp. S_i) le rang de X_i (resp. Y_i) parmi les n réalisations.

$$\tau_n = \frac{P_n - Q_n}{C_n^2} = \frac{4}{n(n-1)} P_n - 1$$

où P_n (resp. Q_n) est le nombre de paires concordantes (resp. discordantes). Deux paires $(X_i, Y_i), (X_j, Y_j)$ sont dites concordantes si $(X_i - X_j)(Y_i - Y_j) > 0$ et discordantes si $(X_i - X_j)(Y_i - Y_j) < 0$.

Si X et Y sont indépendantes alors $\tau_n \approx 0$. Plus $|\tau_n|$ est proche de 1, plus la dépendance entre les deux variables est forte. Si $\tau_n > 0$ (resp. $\tau_n < 0$) alors la dépendance est dite positive (resp. négative).

On note également que sous l'hypothèse H_0 : indépendance de X et Y , la distribution de τ_n est proche d'une loi normale de moyenne 0 et de variance $2(2n+5)/(9n(n+1))$.

La notion de copules :

Pour simplifier la compréhension, on parlera seulement du cas bi-dimensionnel ($n = 2$). Soient F_{XY} une fonction de distribution jointe avec les marginales F_X et F_Y . Alors, d'après le théorème de Sklar, il existe une 2-copule C telle que :

$$F_{XY}(x, y) = C(F_X(x), F_Y(y)), \quad \forall x, y \in \mathbb{R}$$

Si F_X et F_Y sont continues alors C est unique. L'utilisation de copules est répandue car elles permettent de mettre en évidence la dépendance entre variables indépendamment

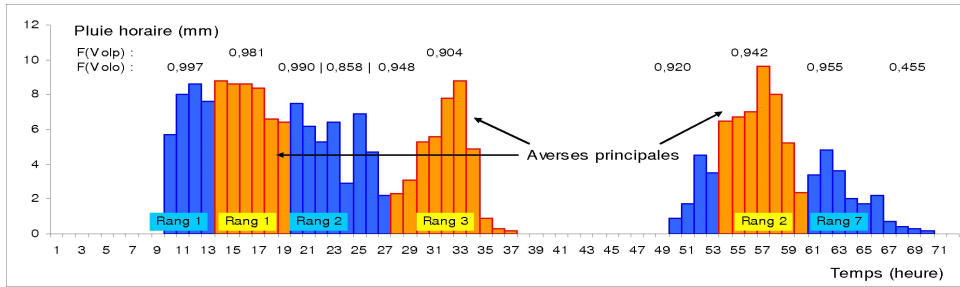
de leurs lois marginales. Ici Seules trois copules archimédiennes sont considérées : Frank, Clayton et Gumbel pour lesquels leur $C(u, v)$ est définis dans Tab. 1.

Copulas	$C_\theta(u, v)$	Paramètre θ	tau de Kendall
Indépendance	uv	No	—
Frank	$\frac{1}{\theta} \ln \left(1 + \frac{(e^{\theta u} - 1)(e^{\theta v} - 1)}{(e^\theta - 1)} \right)$	\mathbb{R}^*	$\tau(\theta) = 1 - \frac{4}{\theta} + \frac{4}{\theta^2} \int_0^\theta \frac{t}{e^t - 1} dt$
Gumbel	$\exp \left\{ - \left[(-\ln u)^\theta + (-\ln v)^\theta \right]^{1/\theta} \right\}$	$\theta \geq 1$	$\tau(\theta) = 1 - \frac{1}{\theta}$
Clayton	$(u^{-1/\theta} + v^{-1/\theta} - 1)^{-\theta}$	$\theta \geq -1$	$\tau(\theta) = \frac{\theta}{\theta + 2}$

TABLE 1 – Définition des quatre copules archimédiennes et l’espace d’état de leur paramètre (θ), et leur lien avec le tau de Kendall (τ).

Dépendance des volumes des averses :

La dépendance des volumes des averses, appelée "persistance" est illustrée par cette Figure :



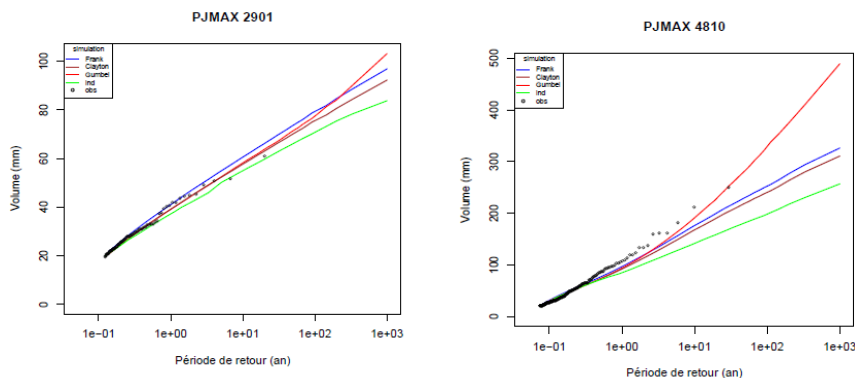
Dans cet évènement observé en février 1990, il y a les trois plus fortes averses¹ de toute la chronique. Sous hypothèse d’indépendance des volumes, reproduire un tel évènement (ou plus "fort") a environ 1 chance sur 10000 d’être généré par le modèle de pluie alors que la chronique entière ne comporte qu’une trentaine d’évènements. Il apparaît donc nécessaire de modéliser la dépendance des volumes afin de pouvoir générer de tels évènements dans des gammes de probabilités similaires aux observations.

Dans un premier temps, nous avons modélisé la dépendance chronologique des averses : le volume de l’averse $i+1$ était dépendant de celui généré pour l’averse i . Ce modèle n’était pas adéquat car le test sur le tau de Kendall conduisait à l’indépendance pour les évènements forts. Nous avons alors proposé de modéliser la persistance par la dépendance des "volumes ordonnées" -ie- une modélisation liant les volumes triés de façon décroissante. Le cadre des copules traditionnelles n’étant pas adapté à ce modèle, nous avons mis en place une méthode afin d’estimer et générer une telle dépendance à l’aide des copules qui a été validé par simulation dans Cantet (2009).

L’utilisation de la copule de Gumbel dans ce cadre montre des résultats satisfaisants au vu des observations. De plus, la paramétrisation de ce modèle permet de "produire

1. celles qui apportent le plus de volume de pluie

de la persistance” seulement sur les postes où la persistance est marquée (méditerranéen, alpins et tropicaux). Le choix de la copule de Gumbel pour modéliser la dépendance entre les volumes triés a été motivé par un test d’adéquation. Ce choix influe beaucoup sur le comportement à l’infini du générateur (voir Fig. 2). Le fait que cette copule modélise une dépendance asymptotique permet de générer des événements extrêmes qui ne seraient que très rarement obtenus avec les copules de Clayton ou de Frank.



(a) Poste de Bretagne (faible persistance) (b) Poste des Cévennes (persistance marquée)

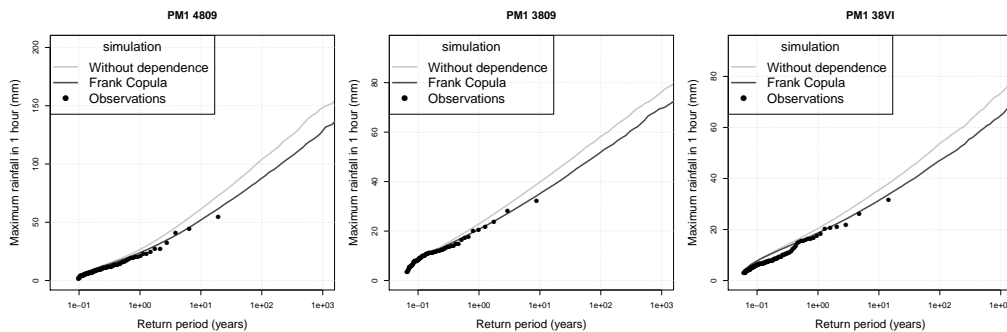
FIGURE 2 – Quantiles de la pluie journalière maximale de 2 postes issus de climat tempéré, et méditerranéen estimés par SHYPRE à partir de différents choix de copule dans la modélisation de la persistance.

Dépendance entre volume et durée des averses :

Dans un premier temps, le volume et la durée des averses était généré de façon indépendante ce qui menait à une sur-estimation des quantiles des pluies horaires via SHYPRE. En effet, un fort volume pouvait être associé à une durée courte (par exemple 1h). À l’aide des observations, nous avons pu constater que cette indépendance était remise en cause sur nombreux postes, notamment sur les postes alpins et méditerranéens².

Cette dépendance a alors été modélisée par les copules. Selon les tests d’adéquation et les résultats, la copule de Frank apparait comme la plus adéquate pour modéliser ce phénomène. Cette prise en compte de la dépendance a permis d’améliorer significativement l’estimation des quantiles de pluies horaires tout en limitant l’impact sur les pluies journalières sur l’ensemble de la France Cantet et Arnaud (2014).

2. correspondant aux régions où la sur-estimation des pluies horaires avait pu être constatée



(a) Station 4809

(b) Station 3809

(c) Station 38VI

FIGURE 3 – Impact de la modélisation de la dépendance volume/durée sur la distribution de la pluie maximale horaire sur trois postes où la dépendance est significative.

3 Performances du modèle

Dans le cadre du projet ANR Extraflo visant la comparaison d’approches de prédétermination, la méthode SHYPRE a montré de bons résultats pour les partie pluie et débit. Les critères mis au point dans le cadre du projet permettent de caractériser d’une part leur justesse, mesurant la capacité d’un modèle à donner une valeur proche de la valeur réelle et d’autre part leur robustesse, capacité pour un modèle à donner des estimations proches avec des échantillons différents. Les critères de justesse ont été utilisés dans Cantet et Arnaud (2014) pour illustrer les améliorations apportées par la modélisation de la dépendance entre le volume et la durée des averses pour la version locale de SHYPRE. Carreau et al (2013) les utilisent pour comparer la version régionalisée de SHYPRE à une loi régionale montrant la justesse et la robustesse des méthodes de prédétermination dites régionales.

Bibliographie

[1] Arnaud, P. et Lavabre, J. (2010), J., *Estimation de l’aléa pluvial en France métropolitaine*, QUAE.
 [2] Arnaud, P. ; Cantet, P. et Aubert, Y. (2014), Relevance of an at-site flood frequency analysis method for extreme events based on stochastic simulation of hourly rainfall, *Hydrological Sciences Journal*
 [3] Cantet, P. (2009), *Impacts du changement climatique sur les pluies extrêmes par l’utilisation d’un générateur stochastique de pluies*, Université de Montpellier II
 [4] Cantet, P. et Arnaud, P. (2014), Extreme rainfall analysis by a stochastic model : impact of the copula choice on the sub-daily rainfall generation, *Stochastic Environmental Research and Risk Assessment*, Springer Berlin Heidelberg, 28(6), 1479-149
 [5] Carreau, J. ; Neppel, L. ; Arnaud, P. et Cantet, P. (2013), Extreme Rainfall Frequency Analysis in the South of France : Comparisons of Three Regional Methods, *Journal de la Société Française de Statistique*, 154(2), 119-138