Studies in Nonlinear Sciences 2 (1): 13-18, 2011
ISSN 2221-3910

Modelling Nonlinear Bivariate Dependence Using the Boubaker Polynomials Copula: Application to Infiltration Rainfall Patterns in Saddine-1 (Makthar, Northern Tunisia)

[1]*Emna Gargouri-Ellouze*, [2]*Noreen Sher Akbar* and [2]*Sohail Nadeem*

[1]Institut Supérieur des Etudes Technologiques de Radès,
Tunisie (High Institute of Technological Studies of Rades), 2098, Tunisia
[2]Department of Mathematics, Quaid-i-Azam University 45320, Islamabad 44000, Pakistan

Abstract: In this paper, we apply the Boubaker Polynomials copula to a set of discrete random vectors possessing uniform margins. We further suggest a pragmatic way to fit the dependence structure of multivariate data to Boubaker Polynomials copula and empirical contingency tables. Finally, we discuss an application of the relationship between infiltration index (ϕ-index) and the average intensity of rainfall event in northen Tunisia (Saddine 1, Makthar). We study the modeling of dependence patterns of infiltration index (ϕ-index) and the average intensity (I_{moy}) of rainfall event via a multi-parametric linear regression framework equipped with the Boubaker polynomial expansion to the multi-dimensional copula and estimate it with empirical data of infiltration index (ϕ-index) and the average intensity of rainfall event (hydrodynamical data). Our result shows that it is a feasible approach to describe the dependence patterns with the Boubaker Polynomials copulas approximation.

Key words: Boubaker polynomial copula . infiltration index (ϕ-index) . dependence patterns . intensity rainfall

INTRODUCTION

We find in Schweizer [1], Joe [2]; Frees *et al*. [3] Nelsen [4]; Roncalli [5], Denuit *et al*. [6], Genest *et al*. [7] and references herein, Copula theory.

A copula is a multivariate joint cumulative distribution defined on the p-dimensional unit cube $[0, 1]^p$ such that every marginal distribution is uniform on the interval [0, 1]. Consider a vector $X = (X_1, X_2, \ldots X_p)$ of $p \geq 2$ continuous random variables, we call the function C a copula [8]. if:

$$\begin{aligned} &\Pr\left(X_1 \le x_1, X_2 \le x_2, \ldots\ldots X_p \le x_p\right) \\ &= C\left(F_1(x_1), F_2(x_2), \ldots\ldots F_p(x_p)\right) \end{aligned} \quad (1)$$

such as: $F_k(x_k) = \Pr(X_k \leq x_k)$ is the marginal distribution of X_k.

C is the cumulative distribution function of a vector $(U_1, U_2, \ldots U_p)$ of dependent uniform random variables on the interval [0; 1]:

$$\begin{aligned} C\left(F_1(x_1), F_2(x_2), \ldots F_p(x_p)\right) &= C\left(U_1, U_2, \ldots\ldots U_p\right) \\ &= F\left(x_1, x_2, \ldots x_p\right) \end{aligned} \quad (2)$$

The associated copula to $X = (X_1, X_2, \ldots\ldots X_p)$ is, for all $(u_1, u_2, \ldots\ldots u_p) \in [0,1]^p$:

$$\begin{aligned} C(u_1, u_2, \ldots\ldots u_p) &= F\left(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots F_p^{-1}(u_p)\right) \\ &= P\left(U_1 \le u_1, U_2 \le u_2, \ldots U_n \le u_n\right) \end{aligned} \quad (3)$$

Specifically, C: $[0,1]^p \to [0,1]$ is an p-dimensional copula if:

$$C\left(u_1, u_2, \ldots, u_i = 0, \ldots, u_p\right) = 0$$

and

$$C\left(1,1,\ldots,u_i,\ldots,1\right) = u_i;\ i = 1\ldots p \quad (4)$$

C is bounded and p-increasing.
The density function f of F is expressed by:

$$\begin{aligned} f(x_1, x_2, \ldots\ldots x_p) &= c\left(F_1(x_1), F_2(x_2), \ldots\ldots F_p(x_p)\right) \\ &\times f_1(x_1) \times f_2(x_2) \times \ldots\ldots \times f_p(x_p) \end{aligned} \quad (5)$$

où f_k is density function of X_K and $c(u_1, u_2, \ldots\ldots u_p)$ is copula density definied by:

Corresponding Author: Emna Gargouri-Ellouze, Institut Supérieur des Etudes Technologiques de Radès, Tunisie (High Institute of Technological Studies of Rades), 2098 Tunisia

$$c(u_1, u_2, \ldots\ldots\ldots u_p) = \frac{\partial^p C(u_1, u_2, \ldots\ldots\ldots u_p)}{\partial u_1 \partial u_2 \ldots\ldots\ldots \partial u_p} \quad (6)$$

In addition, it is worth notice that each copula is bounded by the Frechet-Hoeffding bounds [9], so that:

$$\begin{aligned} &\max\left(F_1(x_1) + F_2(x_2) + \ldots + F_p(x_p) - 1, 0\right) \le F(x_1, x_2, \ldots x_p) \\ &\le \min\left(F_1(x_1), F_2(x_2), \ldots\ldots, F_p(x_p)\right) \end{aligned} \quad (7)$$

There are several types of copula [4, 7, 10, 11]: non-parametric copula and parametric one (one or multi-parameters). Among copula belong to the first category, empiric copula, independence copula, periodic copula introduced by Alfonsi *et al*. [4] and Fréchet copula (« upper bound » maximum copula and « lower bound » minimum copula). In the second copula category we distinguish several families: Archimedean copula [1-8], meta-elliptical copula, in particular Gaussian copula, Cauchy copula and Student copula, exponential copula and contamination copula [10], survival copula [12], extreme value copula: Joe BB5 [7] and Galambos [13].

In hydrology, one often uses families of multivariate distributions which are extensions of univariate and suffer from several limitations and constraints, such as the marginal distributions which may belong to the same probability family [14, 15]. Thus to avoid these limitations, copulas models are used [7]. Indeed, copulas allow the description of the dependence structure between random variables without information on the marginal distributions and also the description of the multivariate distributions with any kind of marginal distributions [11]. It is worth notice that Archimedean copula and meta-elliptical copula families remain the most used.

In this paper we apply the Boubaker Polynomials copulas to a set of discrete random vectors retrieved from data of infiltration index (ϕ-index) and the average intensity (I_{moy}) of rainfall event.

BOUBAKER POLYNOMIALS COPULA

For an integer d, let :U = (U_1,..., U_p) be random vector whose marginal component $U_i|_{i=1.d}$ follows a discrete uniform distribution over $T_i|_{i=1.d}$ = {0,1,2,…,m_i-1} with m_i an non-null integer and:

$$p(U_1, \cdots, U_d) := P\left(\bigcap_{i=1}^{d} \{U_i = k_i\}\right)$$

for all

$$k_i|_{i=1..d} \in \mathop{\times}_{i=1}^{d} T_i \quad (8)$$

Parallel to this definition, let:

$$\tilde{B}_m(x) = \frac{B_{4m}(\beta_i x)}{\int_0^1 B_{4m}(\beta_i x)dx} \quad (9)$$

where B_{4m} are the Boubaker polynomials [16-25] and β_i is the i[th] orderd Boubaker polynomial positive root [19-22].
Then:

$$c_{BP} = (u_1, \cdots, u_d) = \sum_{k_1=0}^{m_1-1} \sum_{k_d=0}^{m_d-1} p(k_1, \cdots, k_d) \prod_{i=1}^{d} m_i B_{m_i-1}(\beta_i u_i)$$

with:

$$(u_1, \cdots, u_d) \in [0,1]^d \quad (10)$$

defines the density of a d-dimensional Boubaker Polynomials copula. Hence, c_{BP} is the Boubaker Polynomials copula density induced by U.

The Boubaker Polynomials copula, as a function: $[0,1]^d \to [0,1]$, verifies the main properties:

- c_{BP} is grounded, i. e. for every U = (U_1,..., U_p)$\in [0,1]^d$, c_{BP} (U) = 0 if the condition: $\exists i/U_i = 0$ is ensured.
- c_{BP} is d-increasing, i. e. for every U = (U_1,..., U_p)$\in [0,1]^d$ and V = (V_1,..., V_p)$\in [0,1]^d$, the Copula-volume of the box [U,V] is strictly positive
- c_{BP} (1,..,1, U_k,1,..1) = U_k.

Hydrometric data: In hydrology, the most difficult problem is how to determine the amount of effective rainfall to route. It is a nonlinear problem that involves a variety of hydrological processes, the heterogeneity of rainfall intensities, soil characteristics and antecedent conditions. One of the approximation infiltration process models is the Horton process supposing that runoff is generated by rainfall intensities that are greater than the soil infiltration capacity. The index infiltration method (ϕ-index) represents the average value of infiltration capacity through the duration of rainfall. This method is still largely used for estimating effective rainfall and deducing flood volume for specific rainfall events.

Furthermore, the infiltration capacity and infiltration index are generally a function of the antecedent soil moisture condition. The method of Antecedent Precipitation Index (API) allows the actualization of infiltration index for the runoffs progress during the hydrological year.

Since the ϕ-index is the average value of infiltration capacity f through the duration (D_R) of rainfall (Eq. (11)), we propose to explore the relationship identified between ϕ-index and rainfall intensity.

Table 1: Characteristics of studied events

Event	P (mm)	D (min)	I_{max} (5mn) (mm/h)	I_{moy} (5mn) (mm/h)	V (m^3)	Qp (m^3/s)	tp (min)	tb (min)	ϕ (5mn) (mm/h)
20/5/92	106.0	116	260.0	55	33059.0	34.70	25	60	166.00
24/5/92	36.0	299	36.0	7	1509.0	0.60	25	85	26.20
14/9/93	26.0	27	84.0	58	10657.0	3.00	60	120	58.00
31/7/94	35.5	42	120.0	51	20843.0	19.80	30	60	73.00
8/1/95	12.0	138	10.0	5	400.0	0.10	60	190	8.98
8/6/95	14.0	28	58.8	30	1767.6	0.04	-	-	53.00
24/6/95	11.5	13	101.0	53	3980.0	0.60	50	470	87.80
24/8/95	12.5	12	102.0	63	41940.0	26.70	20	95	5.50
4/9/95	39.5	13	324.0	182	67200.0	85.60	15	29	162.00
4/9/95 is	8.5	30	33.6	17	15164.0	2.10	50	200	10.60
16/9/95	7.5	14	56.0	32	16055.0	0.10	60	-	13.10
7/2/96	8.0	73	31.2	7	3152.0	0.40	5	360	21.30
15/8/96	39.5	44	115.0	54	1476.0	2.00	10	-	103.20
9/9/96	12.0	74	56.6	10	15573.0	10.40	25	35	13.10
9/9/96bis	13.0	53	28.8	15	15030.0	0.10	40	-	13.10
18/8/97	10.5	26	68.4	24	6338.0	2.60	45	225	48.40
21/9/97	17.5	21	118.8	50	26393.0	22.30	25	85	38.80
4/11/97	5.5	16	32.4	20	383.0	0.20	95	240	31.20
6/12/97	10.5	243	4.8	3	2336.0	0.20	135	335	3.75
5/8/99	27.0	36	99.6	45	35093.0	7.90	60	390	47.30

P: rainfall depth; D: rainfall duration; Imax: rainfall maximum intensity; Imoy: rainfall average intensity; V: runoff volume

Qp: peak discharge; tp: peak time; tb: base time; ϕ: infiltration index

$$\phi = \frac{1}{D_R}\int f(t)dt \qquad (11)$$

Instead of considering the entire hyetograph, we rather focus on modeling the relationship between infiltration index and average rainfall intensity (I_{moy}).

As mentioned above, infiltration is a function of a set of variables, therefore needs the modeling of their joint behavior. On the basis of works of Ellouze-Gargouri *et al*. [26], we suggest prospecting the nonlinear dependence between ϕ-index and I_{moy}, using a nonlinear multivariate approach.

The mesearuments come from a small catchment: Saddine 1. It is located adjacent to Makthar in Tunisia (Northern latitude 35°48'06'' and Eastern longitude 9°04'09'') in a mountainous zone, monitored by the DGACTA (Direction Générale des Aménagements et Conservation des Terres Agricoles) and the IRD (Institut de Recherche et Développement) from 1992 to 2000, within the framework of the HYDROMED project (Programme de recherche sur les lacs collinaires dans les zones semi-arides du pourtour méditerranéen). This catchment is controlled by a headwater dam whose filling was realized in 1992. We select a sample of 51 (among 55, the four remaining are for validation) observed rainfall hyetographs (I) with a reference time increment Δt equal to 5 minutes; thus a set (I_{moy}) has been constituted through the observation period. On other hand, the confrontation of rainfall hyetograph-runoff volume series allows establishing ϕ-index set.

Besides the marginal distributions of each variable ϕ-index and I_{moy} are fitted; the former is exponential distribution with a mean $\mu\phi$ = 29.2 mm/h and a position parameter of ϕ_0= 3.2 mm/h; and the latter I_{moy} is Lognormal distribution with mean μ_{Imoy} = 2.8 mm/h and standard deviation σ_{Imoy}= 0.98 mm/h.

Data and application of the Boubaker Polynomials copula: In this work, the rank correlation coefficient Kendall's tau (τ) [2] is used for the characterization of dependence to measure the association between ϕ-index and I_{moy} obtained from raw maesurements (Table 1). The Kendall coefficient of rank correlation can be used for revealing dependence of two qualitative characteristics, provided that the elements of the sample can be ordered with respect to these characteristics. This coefficient, which measures non linear dependence, integrates the rank of observations rather than their value. Therefore the value itself is not so important than its rank among other values. In summary the more the Kendall's τ is high the more the dependence is important.

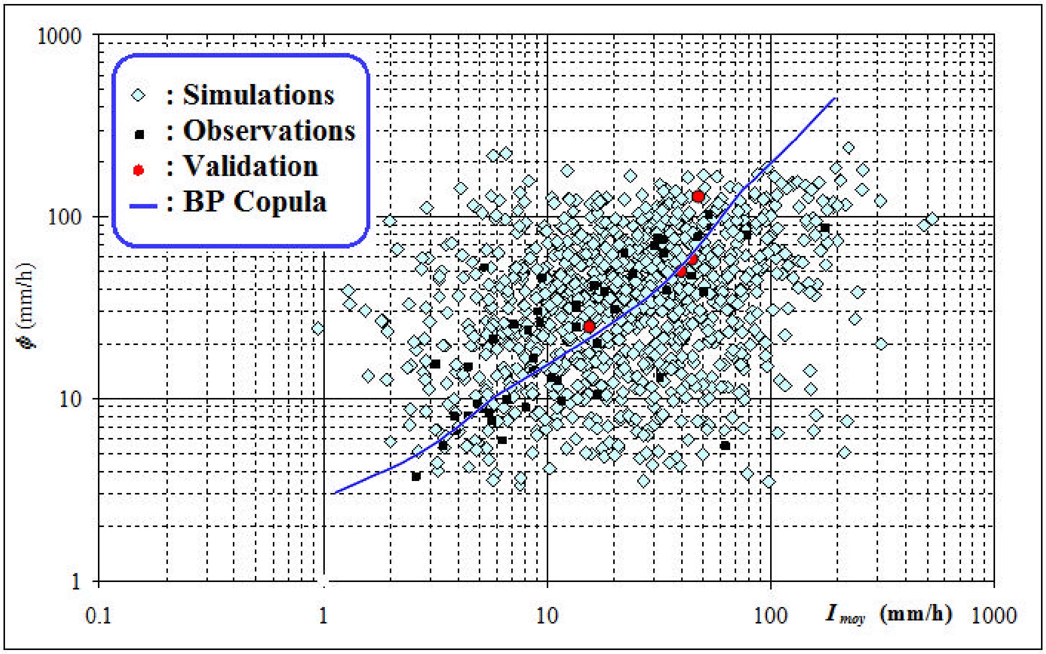

Fig. 1: Simulation of ϕ-index knowing I_{moy}

Table 2: Kendall's τ value and z* statistic

Couple	Sample size	τ	z*	H_0
(ϕ, I_{moy})	51	0.57	0.19	rejection

I_{moy}: Average rainfall intensity; τ: Kendall's tau; z*: test statistic; H_0: null hypothesis

A test of independence can be adopted for Kendall's τ, since under the null-hypothesis H_0, this statistic is close to normal with zero mean and variance (n size of sample). As a result H_0 would be rejected at approximate level α if:

$$|\tau| > z_{\alpha/2}\sqrt{\left[2(2n+5)\right]/\left[9n(n-1)\right]} \qquad (12)$$

For α=5%, $z_{\alpha/2}$ = 1.96. Let z* represent the quantity:

$$z^* = z_{\alpha/2}\sqrt{\left[2(2n+5)\right]/\left[9n(n-1)\right]} \qquad (13)$$

Table 2 deals with τ's value and the corresponding statistic for the couple (ϕ, I_{moy}). The analysis of Table shows that H_0 independence hypothesis is rejected. Consequently ϕ depends on I_{moy}.

On the basis of the methodology precedently performed by Gargouri-Ellouze *et al*. [26], Gumbel copula has been adopted for (ϕ, I_{moy}) with a parameter a = 3. Figure 1 shows the simulated and observed data. The Boubaker Polynomials copulas are applied through the conjoint distribution (ϕ, I_{moy}) according to Eq. (8-10). The observations are reconstituted; in addition the four excluded hydrological events (20/5/92, 24/5/92, 14/9/93 and 31/7/94) used for the validation have been also reconstituted. Consequently we may assume that this modeling has been validated.

ANALYSIS AND DISCUSSION

In the related exmple, the raw data ('Ovservations' in Fig. 1) exhibited a relative dispersion (Pearson coefficient ≈0.61) in the binary system (ϕ, I_{moy}). For relevant catchments located in Tunisia, an eventual dependence between these two variables has been investigated and the Boubaker polynomials copula has been used in order to smooth the empirical data using a test of goodness-of-fit.

In fact, the average infiltration rate during a rainfall event, given by the ϕ-index, is supposed to be primarily a function of soil characteristics and antecedent moisture content. Only for comparatively small rainfall amounts, i.e., having intensities lesser than infiltration capacity, the ϕ-index would also depend on rainfall characteristics. In view of this, it is meaningless and very difficult to obtain a sensible estimate of ϕ-index by conditioning it only on the average rainfall intensity. For example, for the actual soil type, ϕ-index will vary tremendously for the same average rainfall intensity if the antecedent moisture content is very low or very high. Conversely, for very low or very high average rainfall intensity, the ϕ-index may come out to be of equal magnitude if the antecedent moisture was very low or very high respectively. This indicates that there is little value of conditioning ϕ-index on rainfall.

In this context, a primal accurate sampling process has appropriately yielded a subset of validated measurements (‘Validation’ in Fig. 1). This latter was consequently subjected to the Boubaker Polynomials copula analysis (Continuous line in Fig. 1). The validated profile was in good agreement with the Boubaker Polynomials copula results along with those using the *-product defined by Darsow *et al*. [27].

CONCLUSION

Copulas are actually a major tool in modelling the dependence structure in hydrologycal, mass transfer and several other nonlinear systems applications. In this work, we studied a new object in multivariate analysis called the Boubaker Polynomials copula. We showed that, subject to regularity, any copula can be efficiently approximated by some Boubaker Polynomials copula. This copula representation should allow us to take advantage of the properties of the copula function whenever multivariate normality is not a good assumption and application to a standrd problem of dependence patterns of infiltration index (ϕ-index) and the average intensity (I_{moy}) of rainfall event in a given geographic zone could be successfully carried out. We also showed that modeling joint distributions using copulas relaxes the restriction of selecting marginals for hydrodynamic similar variables.

It is clear that a lot has been left out from this paper. The results can be extended to the right censored data using any centred unbaised Kaplan-Meier-like estimator along with i. e. Akaike information criterion and Bayesian information criteria [28]. A bandwidth choice in practice remains an open question and existing relevant methods like cross-validation or radom sampling can be investigated.

ACKNOWLEDGMENT

The authors wish to thank the eminent Professor K. Boubaker for helpful comments and correction.

REFERENCES

1. Schweizer, B., 1991. Thirty years of copulas. In: Advances in Probability Distributions with Given Marginals: Beyond the Copulas. Mathematics and Its Applications. Dall’Aglio, G., S. Kotz and G. Salinetti (Eds.), Kluwer Academic Publishers, Dordrecht, The Netherlands, pp: 13-50.
2. Joe, H., 1997. Multivariate Models and Dependence Concepts. Monographs on Statistics and Applied Probability, Chapman and Hall, London, UK.
3. Frees, E.W. and E.A. Valdez, 1998. Understanding relationships using copulas. N. Am. Actuarial J., 2 (1): 1-25.
4. Alfonsi, A. and D. Brigo, 2005. New families of Copulas based on periodic functions. Communications in Statistics: Theory and Methods, 34 (7): 1437-1447.
5. Roncalli, T., 2002. Théorie des valeurs extrêmes. DESS, Université Paris IX Dauphine, France.
6. Denuit, M. and A. Charpentier, 2004. Mathématiques de l’Assurance Non-Vie-Concepts fondamentaux de la théorie du risque. Collection «Économie et Statistiques Avancées», tome 1. Economica, Paris, France.
7. Genest, C. and A.C. Favre, 2007. Everything you always wanted to know about copula modeling but were afraid to ask. J. Hydrol. Engng ASCE, 12: 347-368.
8. Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. Publ. Inst. Statist. Université Paris VIII, France, 8: 229-231.
9. Nelsen, R.B., 1999. An Introduction to Copulas. Springer, New York, USA.
10. Kallenberg, W.C.M., 2008. Modelling dependence. Insurance: Mathematics and Economics, 42: 127-146.
11. Bárdossy, A., 2006. Copula-based geostatistical models for groundwater quality parameters. Water Resour. Res. 42, W11416, doi:10.1029/2005WR004754.
12. Venter, G., 2002. Tails of copulas. Proc. Casualty Actuarial Society, 89: 68-113.
13. Ghoudi, K., A. Khoudraji and L.P. Rivest, 1998. Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles. Rev Can. Stat., 26 (1): 187-197.
14. Salvadori, G. and C. De Michele, 2004a. Analytical calculation of storm volume statistics involving Pareto-like intensity-duration marginals. Geophys. Res Lett. 31, L04502, doi:10.1029/2003GL018767.
15. Salvadori, G. and C. De Michele, 2004b. Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. Water. Resour. Res. 40, W12511, doi:10.1029/2004WR003133.
16. Barry, P. and A. Hennessy, 2010. Meixner-Type results for Riordan arrays and associated integer sequences. section 6: The Boubaker polynomials. Journal of Integer Sequences, 13: 1-34.

17. Agida, M. and A.S. Kumar, 2010. A Boubaker Polynomials Expansion Scheme solution to random Love equation in the case of a rational kernel. El. Journal of Theoretical Physics, 7: 319-326.
18. Yildirim, A., S.T. Mohyud-Din and D.H. Zhang, 2010. Analytical solutions to the pulsed Klein-Gordon equation using Modified Variational Iteration Method (MVIM) and Boubaker Polynomials Expansion Scheme (BPES). Computers and Mathematics with Applications; DOI:10.1016/j.camwa.2009.12.026.
19. Slama, S., J. Bessrour, M. Bouhafs and K.B. Ben Mahmoud, 2009. Num. Heat Transf. Part A 55: 401-404.
20. Slama, S., M. Bouhafs, K.B. Ben Mahmoud and A Boubaker, 2008. Polynomials Solution to Heat Equation for Monitoring A3 Point Evolution During Resistance Spot Welding. International Journal of Heat and Technology, 26 (2): 141-146.
21. Tabatabaei, S., T. Zhao, O. Awojoyogbe and F. Moses, 2009. Cut-off cooling velocity profiling inside a keyhole model using the Boubaker polynomials expansion scheme. Int. J. Heat Mass Transfer, 45: 1247-1255.
22. Fridjine, S. and M. Amlouk, 2009. A new parameter: An ABACUS for optimizig functional materials using the Boubaker polynomials expansion scheme. Modern Phys. Lett., B 23: 2179-2182.
23. Belhadj, A., J. Bessrour, M. Bouhafs and L. Barrallier, 2009. Experimental and theoretical cooling velocity profile inside laser welded metals using keyhole approximation and Boubaker polynomials expansion. J. of Thermal Analysis and Calorimetry, 97: 911-920.
24. Kumar, A.S., 2010. An analytical solution to applied mathematics-related Love's equation using the Boubaker Polynomials Expansion Scheme. Journal of the Franklin Institute, 347: 1755-1761.
25. Milgram, A., 2011. The stability of the Boubaker polynomials expansion scheme (BPES)-based solution to Lotka-Volterra problem. J. of Theoretical Biology, 271: 157-158.
26. Gargouri-Ellouze, E. and A. Chebchoub, 2008. Modélisation de la structure de dépendance hauteur-durée d'événements pluvieux par la copule de Gumbel. Hydrol. Sci. J., 53 (4): 802-817.
27. Darsow, W.F., B. Nguyen and E.T. Olsen, 1992. Copulas and Markov Processes. Illinois.
28. Karmakar, S. and S.P. Simonovic, 2009. Bivariate flood frequency analysis. Part 2: A copula-based approach with mixed marginal distributions. J. Flood Risk Management, 2: 32-44.