DOI: 10.24850/j-tyca-2025-01-08

Articles

Trivariate flood frequencies analysis with regional dependence and Copula Functions

Análisis de frecuencias de crecientes trivariados con dependencia regional y funciones Cópula

Daniel Francisco Campos-Aranda[1], ORCID: https://ordic.org/0000-0002-9876-3967

[1]Retired professor of the Autonomous University of San Luis Potosi, Mexico, campos_aranda@hotmail.com

Corresponding author: Daniel Francisco Campos-Aranda, campos_aranda@hotmail.com

Abstract

Design floods (DF) give dimension for hydrological security to the hydraulic protection works. The most reliable estimate is obtained through the univariate *frequency analysis* (FA), which represents the maximum annual flows available, with an appropriate probability distribution function (PDF), to estimate the *predictions* sought. In this study, the FA is carried out with the *trivariate* approach, processing a base record of flows *QX* and two other auxiliaries, *QY* and *QZ*, which are correlated to the first and have the same amplitude. The verification of the *simultaneous* character of the *QX*, *QY* and *QZ* flows (that they belong to the same event analyzed) is described in detail. The *joint* trivariate PDF of flows was obtained using the Gumbel-Hougaard *Copula function*, which showed an excellent fit and reproduced the observed dependency on flows. A numerical application exposed here processed 43 annual flows and was carried out at the hydrometric stations, Tempoal as base, and El Cardón and Terrerillos as auxiliaries of the Tempoal river system of Hydrological Region No. 26 (Pánuco), Mexico. In order to obtain the ideal *marginal* PDFs, the Moment Ratios Diagram L was used and, in addition, the Kappa and Wakeby PDFs were applied to contrast predictions. Finally, conclusions are formulated, which highlight the importance of the trivariate approach, based on regional dependence, to validate the behavior in magnitudes of the DF estimated with the *univariate* approach.

Keywords: Frank and Gumbel-Hougaard *CF*, symmetric multivariate *CF*, asymmetric trivariate *CF*, Kendall's tau ratio, upper tail and observed dependences, secondary return period, design events.

Resumen

Las *crecientes de diseño* (CD) permiten dar dimensión por seguridad hidrológica a las obras hidráulicas de protección. Su estimación más confiable se obtiene con el *análisis de frecuencias* (AF) univariado, el cual representa los gastos máximos anuales disponibles, con una función de distribución de probabilidades (FDP) idónea, para estimar las *predicciones* buscadas. En este estudio, el AF se realiza con el enfoque *trivariado*, procesando un registro base de gastos *QX* y otros dos auxiliares, *QY* y

QZ, que están correlacionados con el primero y tienen igual amplitud. Se describe con detalle cómo se verifica que los gastos *QX*, *QY* y *QZ* sean *simultáneos*, es decir, que pertenezcan al mismo evento analizado. La FDP *conjunta* trivariada de gastos se obtuvo mediante la *función Cópula* de Gumbel-Hougaard, que mostró excelente ajuste y reprodujo la dependencia observada en los gastos. La aplicación numérica expuesta procesó 43 gastos anuales y se realizó en las estaciones hidrométricas Tempoal como base, y El Cardón y Terrerillos como auxiliares del sistema del río Tempoal de la Región Hidrológica No. 26 (Pánuco), México. Para la búsqueda de las FDP *marginales* idóneas se utilizó el diagrama de cocientes de momentos L, y además se aplicaron para contraste de predicciones las FDP Kappa y Wakeby. Por último, se formulan las conclusiones, las cuales destacan la importancia del enfoque trivariado, basado en la dependencia regional, para validar el comportamiento en magnitudes de las CD estimadas con el enfoque *univariado*.

Palabras clave: funciones Cópula (*FC*), *FC* de Frank y Gumbel-Hougaard, *FC* multivariadas simétricas, *FC* trivariadas asimétricas, cociente tau de Kendall, dependencia en el extremo superior y observada, periodo de retorno secundario, eventos de diseño.

Received: 15/02/2023

Accepted: 16/09/2023

Published Online: 11/10/2023

Introduction

Generalities

The Mexican Republic is located under the influence zone of hurricanes or cyclones which are originated in the Caribbean Sea and the Pacific Ocean, generating local convective storms and extensive orographic storms. In addition, it is also affected by cold fronts, which are meteorological phenomena with ample spatial reach. These atmospheric events cause *Floods* or *Maximum Avenues* that inundate various regions of the country; resulting in loss of human lives and enormous economic and environmental damage (Aldama, Ramírez, Aparicio, Mejía-Zermeño, & Ortega-Gil, 2006).

The basic hydrological study for estimating floods is called *Frequency Analysis* (FA, by its acronym in Spanish), which defines the *Design Floods* (CD, by its acronym in Spanish), consisting on the maximum river flows associated with low probabilities of being exceeded. The CDs allow sizing for hydrological safety reasons, various protective hydraulic works such as retaining walls and dams, bridges, rectifications and channeling of rivers and urban drainage.

For the estimation of the CD to be reliable, the maximum annual flow record processed must be random, the Probability Distribution Function (FDP, by its acronym in Spanish) used to obtain the desired *predictions* must be ideal and the method used for its adjustment efficient. Furthermore, the selection of results must be objective.

The Afs initiated in the middle of the last century and were initially *univariate* in nature, generally processing the maximum annual flow.

Towards the end of that century, *bivariate* FA began being used, with two approaches. The first one used other variables of the annual floods, such as their runoff volume and total duration (Goel, Seth, & Chandra, 1998; Yue, Ouarda, Bobée, Legendre, & Bruneau, 1999). The second one used auxiliary records, with regional proximity that showed dependence or correlation. These AF were taken to the *trivariate* level (Escalante-Sandoval & Raynal-Villaseñor, 1994).

The first bivariate FAs were based on a joint PDF, which had equal marginal distributions (Normal or Gumbel) and a common recording period in its variables. The first trivariate AF also applied a joint PDF with equal marginals, but processed their variables with different recording periods. The fit of such joint distribution was carried out by maximum likelihood, with a complex algebraic process that was solved with numerical optimization (Escalante-Sandoval & Raynal-Villaseñor, 2008).

In this study, the FA is addressed under the *trivariate* approach, with a base record (*QX*) and two auxiliaries (*QY* and *QZ*), which show dependence, which means they are correlated and have the same number of years of record. The PDF that represents the triplet of records is constructed based on their previously adopted ideal univariate distributions, by means of a *Copula Function* (*FC*, by its acronym in Spanish). The predictions estimated with the adopted *FC* are contrasted against those obtained with a univariate PDF, fitted to the complete *QX* record.

Objectives

For this study, the following five *objectives* were formulated: (1) to conduct trivariate analyses, working with two *FC* families of the Archimedean class: Frank and Gumbel-Hougaard; (2) to apply the aforementioned *FC* families with their multivariate versions called symmetric; (3) to use the selected *FC* families, nested or asymmetric type with two association parameters; (4) to estimate the design events from the joint Kendall return period and (5) to apply the exposed theory in the Tempoal river system, of Hydrological Region No. 26 (Panuco), Mexico. Tempoal was processed as a base station and El Cardón, Los Hules and Terrerillos as auxiliaries, with a joint record of 43 annual floods.

Study organization

Due to the extensive and diverse theoretical concepts and calculations involved, it is convenient to describe their organization for a better understanding of the study or research. For this reason, there are three chapters described: (1) Copula Functions and Processed Data; (2) the Results and their discussion and (3) the Conclusions.

The first chapter presents a summary of the theoretical aspects applied in the study, hence the beginning with a section that cites the advantages and operational aspects of the Copula Functions (*FC*). Then, the bivariate *FC* that will be used are presented, where the fitting uses Kendall's tau quotient and its selection is based on the right tail probabilistic dependence. This first theoretical part concludes with the description of the trivariate *FC*s that will be applied.

Afterwards, three concepts of frequency analysis are addressed, in which their application allows the fit and selection of *FC*s, these are: (a) the estimation of univariate and trivariate empirical probabilities; (b) the search for the optimal marginal FDPs, based on the fitting errors and (c) concepts and equations of the trivariate OR type, AND and secondary or Kendall return periods. Finally, this chapter exposes the flows of the annual maximum floods that will be processed and describes how the simultaneity of such regional events was verified.

Regarding the Results chapter and its discussion, Figure 1 outlines the sequence of theoretical topics and their calculations with the idea of formulating an explanatory flow diagram.

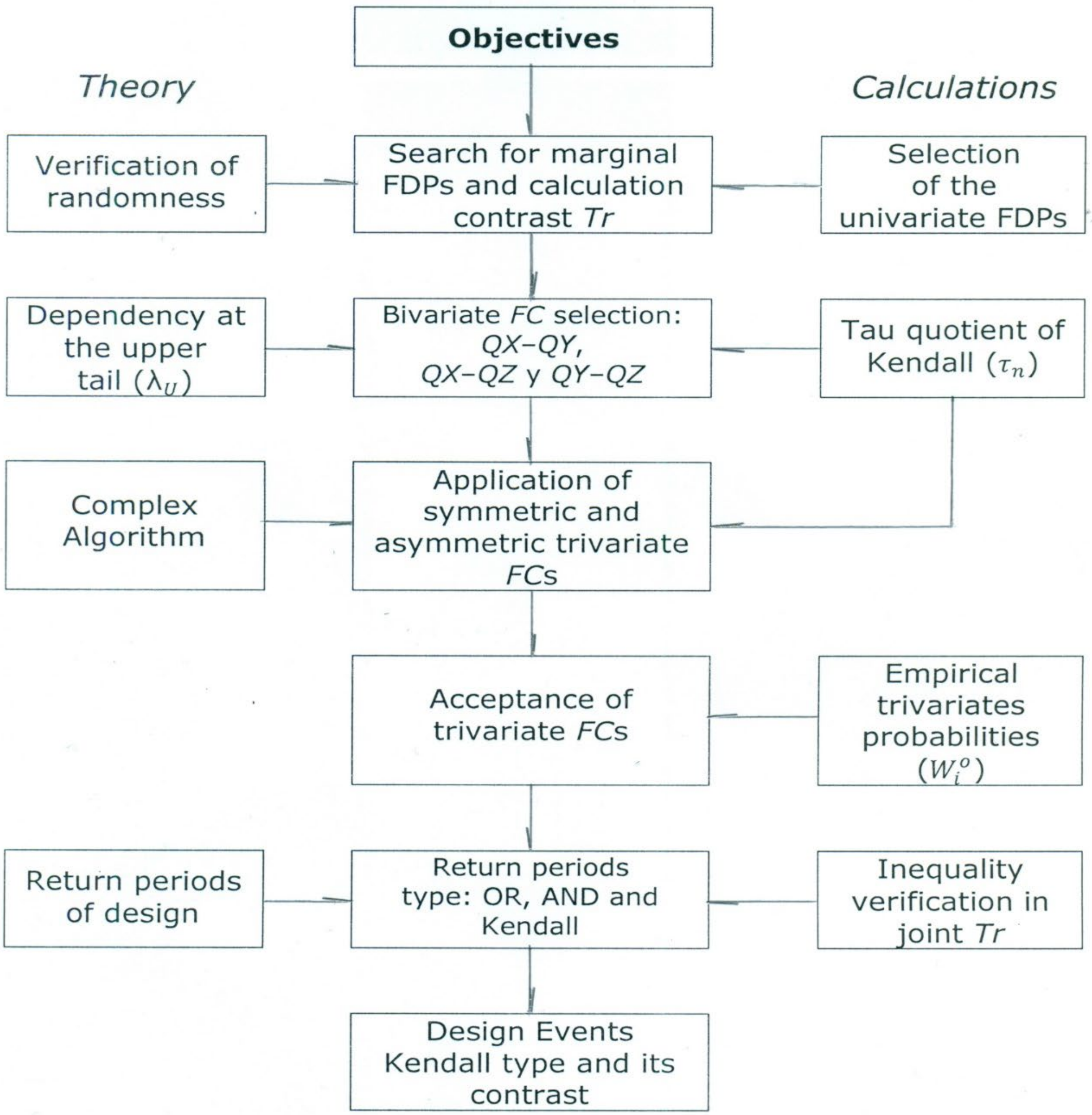

Figure 1. Flow diagram of theoretical concepts and calculations, carried out in the Results and its discussion chapter.

Tecnología y Ciencias del Agua
OPEN ACCESS
Check for updates

Copula functions and processed data

Advantages

The basic advantage of *Copula Functions* (*FC*) is that it allows the formation and expression of the *joint* or *multivariate distribution* of random variables that are correlated, as a function of their marginal distributions, previously adopted. Therefore, a *FC* links or relates the univariate marginal distributions to form the *multivariate distribution*.

The application of *FC* offers complete freedom to adopt or select the univariate marginal distributions that best represent the data (Salvadori, De Michele, Kottegoda, & Rosso, 2007; Meylan, Favre, & Musy, 2012; Genest & Chebana, 2017; Zhang & Singh, 2019).

Another advantage of *FC*s when forming multivariate distributions is that they separate the effect of dependence or correlation between random variables from the effects of marginal distributions in joint modeling.

Families of Copulas

The Copula functions (*FC*) that have been developed are divided into four classes: Archimedean, extreme value, elliptical and miscellaneous. They are also classified into *FC* of one or several parameters, depending on the extent to which the structure of the dependence between the correlated random variables is defined (Meylan *et al.*, 2012; Genest & Chebana, 2017; Chowdhary & Singh, 2019). Salvadori *et al*. (2007) present a broad and useful summary of *FC*, which have been applied in the field of hydrology.

Bivariate Archimedean Copulas

Archimedean Copulas have had wide application due to their simple construction, a single parameter, wide range and acceptance of both types of dependence (positive and negative). Designating $F_X(x) = u$, $F_Y(y) = v$ the marginal FDPs and θ the parameter that measures the dependence or association between u and v, the following two families of Archimedean Copulas are exposed (Genest & Favre, 2007; Salvadori *et al.*, 2007; Zhang & Singh, 2019; Chen & Guo, 2019; Chowdhary & Singh, 2019).

The first *FC* was selected to serve as a contrast of a good fit to the data (Chowdhary & Singh, 2019), but lacking the ability to reproduce the dependence observed in the right tail of such data. With such an approach, the Clayton, Planckett or Raftery *FC*s could have been used, which are easier to fit. The second selected *FC* posesses such capability.

(1) Frank's FC family

Its equation and variation space of θ are:

$$C(u,v) = \frac{-1}{\theta}\ln\left[1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right] \quad (-\infty,\infty) \setminus \{0\} \tag{1}$$

For the negative dependence $0 \leq \theta < 1$ and for the positive $\theta > 1$, with $\theta = 1$ for the independence between *u* and *v*. The relationship of θ with Kendall's tau quotient (τ_n) is the following:

$$\tau_n = 1 + \frac{4}{\theta}[\mathrm{D}_1(\theta) - 1] \qquad (2)$$

where $\mathrm{D}_1(\theta)$ is the Debye function of order 1, expressed as:

$$D_1(\theta) = \frac{1}{\theta}\int_0^{\theta} \frac{s}{e^s - 1} ds \qquad (3)$$

The previous equation was evaluated with numerical integration, ratifying its results against the values tabulated by Stegun (1972). The Gauss-Legendre quadrature method was applied, whose operating equation is (Nieves & Domínguez, 1998; Campos-Aranda, 2003):

$$\int_a^b f(x)\, dx \cong \frac{b-a}{2} \sum_{i=1}^{np} w_i \cdot f\left[\frac{(b-a)h_i + b + a}{2}\right] \qquad (4)$$

in which,

w_i = coefficients of the method

h_i = abscissa

np = number of pairs in which the function $f(x)$ is evaluated, with the argument indicated in $f\,[\cdot]$

In Davis and Polonsky (1972), the 12 used pairs of w_i and h_i with 15 digits, which are acceptable in the *Basic* language, were obtained as double precision variables.

(2) The Gumbel-Hougaard *FC*

Also belongs to the family of extreme values and only accepts positive dependence. Its equation and variation space of θ are:

$$C(u,v) = \exp\left\{-\left[(-lnu)^{\theta} + (-lnv)^{\theta}\right]^{1/\theta}\right\} \quad (5)$$

With θ = 1 there is independence between *u* and *v*. The relationship of θ with Kendall's tau quotient is as follows:

$$\tau_n = \frac{\theta - 1}{\theta} \quad (6)$$

Numerical association indicator

Concordance

As the *FC* characterizes the *dependence* between the random variables *u* and *v*, it is necessary to study the association measures to have a method that allows estimating its parameter θ. In general terms, a random variable is *concordant* with another, when its large values are associated with the large magnitudes of the other and the small values of one with the reduced values of the other (Salvadori *et al.*, 2007; Chowdhary & Singh, 2019).

Some variables with a direct linear correlation will be concordant, since as one increases, the other also increases. Variables with inverse linear correlation will be *discordant*, since large values of one will correspond to small values of the other. The above implies that the pairs $(x_i - x_j)(y_i - y_j) > 0$ are *concordant* (c) and *discordant* (d) when $(x_i - x_j)(y_i - y_j) < 0$ (Salvadori *et al.*, 2007; Chowdhary & Singh, 2019).

Kendall tau quotient

This is a non-parametric numerical indicator that measures the probability of having concordant couples; the expression to estimate it with bivariate data is (Zhang & Singh, 2006; Zhang & Singh, 2019):

$$\tau_n = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{signo}\left[(x_i - x_j)(y_i - y_j)\right] \quad (7)$$

In the above equation:

n = number of observations

sign[·] = +1 if such pairs are concordant and −1 if they are discordant

Genest and Favre (2007) present a test for the tau quotient, which accepts the null hypothesis H_0 of independent X and Y and then its statistics have an approximately Normal distribution with zero mean and variance $2(2n + 5) / [9n(n - 1)]$. Therefore, H_0 will be rejected with a confidence level α = 5 % if:

$$\sqrt{\frac{9n(n-1)}{2(2n+5)}}|\tau_n| > Z_{\alpha/2} = 1.96 \quad (8)$$

Extreme dependence on the bivariate *FCs*

Generalities

The most important criterion applied to select a bivariate *FC* is the one based on the magnitude of the upper tail dependence in the joint distribution, which has an impact on the veracity of the extreme predictions. The upper right tail dependence (λ_U) is the conditional probability that *Y* is greater than a certain percentile (*s*) of $F_Y(y)$, given that *X* is greater than such percentile in $F_X(x)$, as *s* approaches unity. The lower left tail dependence (λ_L) compares *Y* to being less than *X*, as *s* approaches zero (Chowdhary & Singh, 2019).

In relation to the bivariate *FC* exposed, Frank's has insignificant extreme zones dependencies: therefore, $\lambda_L = 0$ and $\lambda_U = 0$. On the other hand, the Gumbel-Hougaard Copula has a significant dependence on the upper tail, equal to:

$$\lambda_U = 2 - 2^{1/\theta} \quad (9)$$

Dupuis (2007) tested six bivariate *FC* families and found that their ability to estimate extreme events ranges from poor to good, with the following order: Clayton, Frank, Normal, *t*-Student, Gumbel-Hougaard, and Clayton Associated (*Survival Clayton*). Poulin, Huard, Favre and Pugin (2007) reaches similar conclusions, when comparing the same six families of Copulas and the one called A12 (Nelsen, 2006), which has significant dependence on its right tail.

Estimation of observed dependence

To address the estimation of the upper tail dependence (λ_U) shown by the available data, the *Empirical Copula* must be defined first. Since the *FC* characterizes the dependence between the random variables *X* and *Y*; then the pair of ranges R_i and S_i coming from such variables are the statistic that retains the greatest amount of information and its scaling with the factor 1/(*n*+1) generates a series of points in the unit square $[0,1]^2$, forming the domain of Empirical Copula (Chowdhary & Singh, 2019), defined as follows:

$$C_n(u,v) = \frac{1}{n}\sum_{i=1}^{n} 1\left(\frac{R_i}{n+1} \leq u, \frac{S_i}{n+1} \leq v\right) \qquad (10)$$

In the above equation, 1(·) indicates a function of the random variables *U* and *V*, which are a continuously increasing transformation of *X* and *Y*, relative to the empirical probability integrals $F_n(X)$ and $F_n(Y)$, with the following equations:

$$U_i = \frac{\text{Range}(X_i)}{n+1} = F_n(X_i) \quad V_i = \frac{\text{Range}(Y_i)}{n+1} = F_n(Y_i) \qquad (11)$$

Poulin *et al.* (2007), and Requena, Mediero and Garrote (2013) use the estimator proposed by Frahm, Junker and Schmidt (2005), which is based on a random sample obtained from the Empirical Copula, its designation comes from its authors Capéraà, Fougères and Genest (1997), and is expressed as:

$$\lambda_U^{CFG} = 2 - 2\exp\left\{\frac{1}{n}\sum_{i=1}^{n} \ln\left[\sqrt{\ln\frac{1}{U_i}\cdot\ln\frac{1}{V_i}}/\ln\left(\frac{1}{\max(U_i,V_i)^2}\right)\right]\right\} \quad (12)$$

This estimator accepts that the *FC* can be approximated by one of the class of extreme values and has the advantage of not requiring a threshold value for its estimation, as is the case of the four estimators presented by AghaKouchak, Sellars and Sorooshian (2013).

Trivariate Archimedean copulas

Symmetrical Archimedean copulas

Chen and Guo (2019) indicate that for multivariate random variables, greater than two ($d \geq 3$) and correlated, the family of Archimedean Copulas are divided into *symmetric* and *asymmetric*. The former are easy to construct and have a single association parameter (θ), which requires that all pairs of variables show the same structure and degree of dependence (Zhang & Singh, 2019).

For the two families of Archimedean Copulas exposed and their *symmetric multivariate* versions ($d \geq 3$), their range of θ and the generating functions $\varphi(s)$ and their first and second derivatives $\varphi'(s)$, $\varphi''(s)$

are indicated, where s is the random variable in the interval from 0 to 1. (Grimaldi & Serinaldi, 2006a; Xu, Yin, Guo, Liu, & Hong, 2016; Chen & Guo, 2019; Zhang & Singh, 2019).

(3) Frank's multivariate Copula family

The range of θ is (0, +∞) and the value of θ =1 indicates the independence condition in u_k:

$$C(u_1, u_2, \cdots, u_d) = \frac{-1}{\theta}\ln\left[1 + \frac{\prod_{k=1}^{d}(e^{-\theta u_k}-1)}{(e^{-\theta}-1)^{d-1}}\right] \tag{13}$$

$$\varphi(s) = -\ln\left(\frac{e^{-\theta s}-1}{e^{-\theta}-1}\right) \tag{14}$$

$$\varphi'(s) = \frac{\theta}{1-e^{\theta s}} \tag{15}$$

$$\varphi''(s) = \frac{\theta^2}{e^{\theta s}-2+e^{-\theta s}} \tag{16}$$

(4) Family of multivariate Gumbel-Hougaard Copula

The range of θ is (1, +∞) and the limit of θ = 1 corresponds to the independence condition in u_k:

$$C(u_1, u_2, \cdots, u_d) = \exp\left\{-\left[\sum_{k=1}^{d}(-\ln u_k)^{\theta}\right]^{1/\theta}\right\} \tag{17}$$

$$\varphi(s) = [-ln(s)]^{\theta} \quad (18)$$

$$\varphi'(s) = \frac{-\theta}{s}[-ln(s)]^{\theta-1} \quad (19)$$

$$\varphi''(s) = \frac{\theta}{s^2}\{(\theta-1)[-ln(s)]^{\theta-2} + [-ln(s)]^{\theta-1}\} \quad (20)$$

Asymmetric Archimedean copulas

To model different dependence structures, in multivariate random variables, Chen and Guo (2019) resort to the approach of Grimaldi and Serinaldi (2006b), of applying *nested* Archimedean Copulas. With such an approach, the most common asymmetric trivariate Archimedean Copulas with two association parameters (θ_1 and θ_2) have the general formula $(u,v,w) = C_{\theta_1}\left(w, C_{\theta_2}(u,v)\right)$ and the following two are presented (Grimaldi & Serinaldi, 2006b; Xu *et al.*, 2016; Zhang & Singh, 2019; Chen & Guo, 2019):

(5) Trivariate Frank Asymmetric Copula Family

With $\theta_2 \geq \theta_1 \in [0,\infty$ and $\tau_{12}, \tau_{13}, \tau_{23} \in [0,1]$ for three random variables with positive dependence:

$$C(u,v,w) = \frac{-1}{\theta_1}\ln\left\{1 - \mathrm{F}_1^{-1}\left(1 - \left[1 - \mathrm{F}_2^{-1}\left(1-e^{-\theta_2 u}\right)\left(1-e^{-\theta_2 v}\right)\right]^{\theta_1/\theta_2}\right)\left(1-e^{-\theta_1 w}\right)\right\} \quad (21)$$

being:

$F_1 = 1 - e^{-\theta_1}$

$F_2 = 1 - e^{-\theta_2}$

(6) Trivariate Gumbel-Hougaard Asymmetric Copula Family

With $\theta_2 \geq \theta_1 \in [0,\infty$ y $\tau_{12}, \tau_{13}, \tau_{23} \in [0,1]$ for three random variables with positive dependence:

$$C(u,v,w) = \exp\left\{-\left[\left((-lnu)^{\theta_2} + (-lnv)^{\theta_2}\right)^{\theta_1/\theta_2} + (-lnw)^{\theta_1}\right]^{1/\theta_1}\right\} \quad (22)$$

Empirical probability estimation

The univariate and trivariate *empirical* non-exceedance probabilities were estimated based on the Gringorten formula (Equation (23)), which has been suggested by several authors for bivariate frequency analyzes, and by Zhang and Singh (2007) for trivariate ones. Such equation is:

$$F(x) = \frac{i-0.44}{n+0.12} \quad (23)$$

where *i* is the number of data sorted from lowest to highest and *n* is the total number or number of years of maximum annual flow records. In bivariate analyzes it is possible to work graphically in the two-dimensional plane, as Campos-Aranda (2023) has explained.

In the case of trivariate probabilities, we worked in three-dimensional space, with the maximum annual expenses *QX* and *QY* in the *x*, *y* plane, and *QZ* in the perpendicular axis (*z*). The numerical process begins by saving the historical records of annual maximum flow (*QX*, *QY* and *QZ*) in files *QXh*, *QYh* and *QZh*; In addition, they were sorted in progressive order of magnitude in files *QXo*, *QYo* and *QZo*. Each annual data is then processed to compare the historical value against the ordered one and the times that the second was less than or equal to the order are counted and designated *NQX*, *NQY* and *NQZ*. The above is equivalent to changing the original data, in each list of historical annual values, for its order or *range* number.

Then each historical set of three ranges is compared against all the others and the times in which the three ranges are smaller are counted (AND condition) and such quantity is called *NQXYZ*; that is, the number of occurrences of combinations of minor *qx*, *qy*, and *qz* in three-dimensional space. Finally, Gringorten's graphic position formula is applied; for the trivariate case it is as follows:

$$F_e(x,y,z) = P(QX \leq qx, QY \leq qy, QZ \leq qz) = \frac{NQXYZ_i - 0.44}{n + 0.12} \quad (24)$$

Copula Function Selection

A simple approach to selecting the Copula Function is based on the fit error statistics, by comparing the observed empirical probabilities (W_i^o)

with the calculated theoretical ones (W_i^c) with the Copula Function being tested. The indicators applied are the standard mean error (*EME*, by its acronym in Spanish), the mean absolute error (*EMA*) and the maximum absolute error (*EAM*); Their expressions are (Chowdhary & Singh, 2019; Chen & Guo, 2019):

$$EME = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(W_i^o - W_i^c\right)^2} \quad (25)$$

$$EMA = \frac{1}{n}\sum_{i=1}^{n}|(W_i^o - W_i^c)| \quad (26)$$

$$EAM = \max_{i=1:n}|(W_i^o - W_i^c)| \quad (27)$$

Search for optimal marginal distributions

The search for the ideal *marginal* distributions, took into account the statistical characteristics of the hydrological data to be processed in Table 1. The above, through the L quotients of asymmetry (t_3) and kurtosis (t_4), which allow defining in the L Quotient Diagram of Hosking and Wallis (1997), the three best distributions, due to their closer proximity to the *five curves* shown in such graph.

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

Table 1. Maximum annual flows (m^3/s) available and estimated at the indicated hydrometric stations of the Tempoal River basin in the common period from 1960 to 2002.

No.	Year Month	Tempoal	El Cardon	Los Hules	Terrerillos	*NXYZ*
1	1960 NOV	1277.0	1080.0	320.8	247.1	2
2	1961 JUN	852.9	303.5	434.5	525.0	6
3	1962 JUN	739.2	246.7	457.5	529.3	5
4	1963 JUL	1800.0	481.0	947.4	895.9	18
5	1964 DIC	748.0	122.6	258.0	397.1	1
6	1965 AGO	792.7	202.0	283.9	659.4	5
7	1966 JUN	1778.0	287.0	742.2	1121.7	16
8	1967 SEP	2245.0	854.2	1009.4	1153.0	25
9	1968 SEP	1145.0	476.0	1096.0	611.2	11
10	1969 SEP	1948.0	555.8	825.0	2224.2	28
11	1970 SEP	1418.0	339.9	800.0	1049.3	16
12	1971 OCT	1630.0	720.4	1064.0	1488.5	22
13	1972 JUL	989.0	185.8	450.0	529.0	6
14	1973 JUN	1668.0	387.0	749.0	1740.0	20
15	1974 SEP	4950.0	1198.3	1950.0	3187.8	37
16	1975 SEP	4040.0	1204.2	2470.0	2085.0	33
17	1976 OCT	1275.0	185.0	472.0	792.3	6
18	1977 OCT	514.0	179.1	559.0	162.9	1
19	1978 SEP	3725.0	1390.0	2874.0	2152.3	34
20	1979 SEP	1655.9	667.0	1082.0	514.2	9
21	1980 SEP	1162.0	357.0	583.2	994.1	14

No.	Year Month	Tempoal	El Cardon	Los Hules	Terrerillos	*NXYZ*
22	1981 AGO	2020.0	733.9	1650.3	(1151.3)	24
23	1982 SEP	539.6	133.1	268.8	491.4	2
24	1983 JUL	868.0	269.8	544.0	743.5	8
25	1984 SEP	4030.0	572.0	2834.9	2981.0	30
26	1985 JUL	1882.0	457.0	938.4	1487.7	22
27	1986 JUN	476.0	130.0	308.0	434.0	1
28	1987 JUL	1765.0	346.8	1440.0	2635.0	19
29	1988 SEP	3265.0	356.0	4350.0	3710.0	21
30	1989 SEP	649.0	306.0	644.0	2100.0	5
31	1990 AGO	1611.0	141.8	(3463.7)	204.5	1
32	1991 OCT	3532.0	1248.0	(1072.2)	2860.0	36
33	1992 OCT	2291.0	790.0	762.8	1607.5	29
34	1993 SEP	6120.0	865.5	1684.1	3422.5	37
35	1994 SEP	1133.0	412.0	723.8	1237.9	15
36	1995 AGO	741.9	381.6	440.9	474.0	3
37	1996 AGO	683.0	218.0	804.0	507.6	4
38	1997 OCT	905.0	85.7	428.4	362.5	1
39	1998 SEP	1266.9	(271.9)	204.3	994.4	13
40	1999 OCT	2693.7	602.9	630.9	3328.3	30
41	2000 JUN	641.2	(185.5)	84.9	753.4	4
42	2001 SEP	1847.9	498.3	278.5	1512.2	24
43	2002 SEP	926.4	134.0	496.7	822.2	5
Estadístico *U*		0.686	0.460	0.538	0.993	–

These distributions have three fit parameters: Generalized Logistic (LOG), General Extreme Values (GVE), Log-Normal (LGN), Pearson type III (PE3) and Generalized Pareto (PAG). The PE3 curve allows testing the Log-Pearson type III (LP3) distribution. This process has been described by Campos-Aranda (2023).

In addition, the Kappa and Wakeby distributions were applied, which have shown great versatility and universality to represent extreme hydrological data series, since they have four and five adjustment parameters (Hosking, 1994; Hosking & Wallis, 1997; Kjeldsen, Ahn, & Prosdocimi, 2017).

The LP3 distribution was the only one that was applied with the method of moments in the logarithmic (WRC, 1977) and real (Bobeé, 1975) domains, selecting the one with the best fit. The remaining seven were fitted to the maximum flow records of the annual floods, through the method of L moments, according to procedures set forth by Hosking and Wallis (1997), and Stedinger (2017).

Fit errors

The first criterion applied to select the best PDF for the available data or series were the so-called *fit errors* (Kite, 1977; Willmott & Matsuura, 2005; Chai & Draxler, 2014). This criterion will be applied after selecting the three best FDPs in the L-ratio diagram, according to their minimum absolute distance and having applied the Kappa and Wakeby distributions.

By changing in equations (25) and (26), the probabilities observed by the ordered data of the analyzed series (x_i) and the probabilities calculated by the values estimated with the FDP that is tested or contrasted, the standard error of fit (*EEA*) is obtained and the mean absolute error (*EAM*). The estimated ($\hat{x}_i$) values are sought for the same probability of non-exceedance, assigned to the data by Gringorten's empirical formula (Equation (23)).

Estimation of the dependence parameter θ

The simplest method to estimate the parameter θ of the bivariate *FC* (Equation (1)) is by trial and error, equating Kendall's tau ratio and in the symmetric trivariate *FC* (equations (13) and (17)), seeking that the fit error statistics (equations (25) to (26)) are minimal.

Estimation of the dependence parameters θ_1 and θ_2

The search for the minimum value of Equation (25) or mean standard error, for the fit of the asymmetric trivariate *FC*s defined by equations (21) and (22), was carried out based on the Complex algorithm of multiple restricted or bounded variables, to find the optimal values of θ_1 and θ_2, satisfying the condition $\theta_2 > \theta_1$.

The *Complex algorithm* is a local exploration technique (Box, 1965), which is guided exclusively by what it finds in its path; Its background, a brief description of its operating process and its OPTIM code in *Basic* language can be consulted in Campos-Aranda (2003). In Bunday (1985) there is another description and code of this search method.

The main designations in the OPTIM code are NX and NY, which define the number of decision and dependent variables, depending on the former; for the case analyzed, two (θ_1, θ_2) and one (θ_2>θ_1). An important advantage of the OPTIM code lies in allowing easy access to the limits (L = lower, U = upper), names and initial values of the variables, in the aforementioned subroutine, through the following designations: XL(I), XU(I), XN$(I), X(I), YL(J), YU(J), YN$(J) and Y(J). For the case studied, I varies from 1 to 2 and J = 1.

In all decision variables, 0.001 was used as the lower limit and 10 and 20 as the upper limit in Frank's *FC* for θ_1 and θ_2, and 5 and 15 in the Gumbel-Hougaard *FC*. The only dependent variable was defined by the ratio of θ_2 to θ_1, with a lower limit of one and an upper limit of 5; value that was adopted arbitrarily.

The objective function is called FO in the OPTIM code and is defined at the end of the program, it logically corresponds to Equation (25), with the name FO$=*EME*, standard mean error. For the convergence criteria of the absolute and relative deviations of the FO, the following values were used: 0.0002 and 0.00001.

Ratification of the selected Copula Function

This is the most important stage of the process of the practical application of *FC*, since it verifies that this model faithfully reproduces the observed trivariate joint probabilities (Equation (24)). Yue (2000) indicates a simple and practical way to represent empirical and theoretical joint probabilities. This consists of taking the first to the abscissa axis and the second to the ordinate axis; Logically, in such a graph, each pair of data defines a point that coincides with or departs from the line at 45º.

Yue and Rasmussen (2002) apply the Kolmogorov-Smirnov test with a significance level (α) of 5 %, to accept or reject the *maximum absolute difference* (*dma*) between the empirical and theoretical joint probabilities. To evaluate the statistic (D_n) of the test, the expression presented by Meylan *et al*. (2012) was used for α = 5 %, this is:

$$D_n = \frac{1.358}{\sqrt{n}} \quad (28)$$

n is the number of data. If the *dma* is less than D_n, the adopted *FC* is ratified.

Trivariate return periods

OR and AND types

The first trivariate return period of the event (*X*,*Y*,*Z*) is defined under the OR condition, which indicates that the limits *x*, *y* or *z*, or all three can be exceeded and then, the classical return period equation or inverse of the probability of exceedance will be (Genest & Chebana, 2017; Zhang & Singh, 2019):

$$T_{XYZ} = \frac{1}{P(X>x \vee Y>y \vee Z>z)} = \frac{1}{1-F_{XYZ}(x,y,z)} = \frac{1}{1-C[F_X(x),F_Y(y),F_Z(z)]} \quad (29)$$

in which, $C[F_X(x), F_Y(y), F_Z(z)] = C(u,v,w)$ is the selected or proven *FC*.

The second trivariate return period of the event (X, Y, Z) is associated with the case in which the three limits are exceeded ($X > x$, $Y > y$, $Z > z$) or AND condition, its equation is the following (Zhang & Singh, 2019):

$$T'_{XYZ} = \frac{1}{P(X>x \wedge Y>y \wedge Z>z)} = \frac{1}{F'_{XYZ}(x,y,z)} = \frac{1}{1-u-v-w+C(u,v)+C(u,w)+C(v,w)-C(u,v,w)} \quad (30)$$

For the application of Equation (30) above, it is observed that the three bivariate *FC* and the trivariate *FC* are required.

Secondary or Kendall type

Salvadori and De Michele (2004) introduce in detail the concept of the *Secondary bivariate Return Period* (ζ), designated as such to emphasize that the joint return period T_{XY} is the primary one, from which it comes by using the *isolines* defined by the applied *FC*, which is expressed as:

$$L_s = [(u,v) \in I^2 : C(u,v) = s] \quad (31)$$

Where:

s = unitary random variable $0 < s \leq 1$

C = tested *FC*

Then, a region $B_C(s)$ is defined in the unit space (I^2) above the isoline, below it and to the left, which will be:

$$B_C(s) = \{(u,v) \in I^2 : C(u,v) \leq s\} \tag{32}$$

In *FC*s of the Archimedean class, the univariate Kendall distribution, designated $K_C(s)$, provides a measure of the events within the $B_C(s)$; Its equation is (Salvadori & De Michele, 2004; Salvadori & De Michele, 2007; Salvadori *et al.*, 2007; Gräler *et al.*, 2013):

$$K_C(s) = s - \frac{\varphi(s)}{\varphi\prime(s)} \tag{33}$$

in which, $\varphi(s)$ is the generator of the *FC* and $\varphi'(s)$ its derivative. Finally, the secondary return period (ζ) of events outside $B_C(s)$ is:

$$\zeta = \frac{1}{1-K_C(s)} \tag{34}$$

where the denominator is the probability of exceedance (*survival function*), which corresponds to probably destructive or dangerous events.

The parametric Kendall distribution (Equation (33)), for the symmetric trivariate Archimedean Copulas is the following (Barbe, Genest, Ghoudi, & Rémillard, 1996; Grimaldi & Serinaldi, 2006a; Zhang & Singh, 2019):

$$K_C(s) = P[C(u,v,w) \leq s] = s - \frac{\varphi(s)}{\varphi'(s)} - \frac{\varphi^2(s) \cdot \varphi''(s)}{2[\varphi\prime(s)]^3} \tag{35}$$

Substituting equations (18) to (20) into Equation (35), we obtain an expression for the Kendall distribution of the symmetric trivariate Gumbel-Hougaard *FC*, which will be used later.

Gräler *et al*. (2013) extend the inequality $T_{XYZ} \leq T'_{XYZ}$, which indicates that the OR type return period is always less than the AND type and indicate that T_{KEN} is intermediate between the two mentioned. T_{KEN} is obtained with Equation (34). Then we have:

$$T_{XYZ} \leq T_{KEN} \leq T'_{XYZ} \tag{36}$$

Salvadori, De Michele and Durante (2011) highlight that the estimation of return periods and their design events in *multivariate* frequency analyzes is a difficult problem. To solve it, they establish a theoretical framework based on the *FC* and the Kendall distribution, which they apply through numerical simulation.

Annual Flood Records to be processed

The Tempoal River is one of the important tributaries of the Moctezuma River, which together with the Tampaón River form the Panuco River, of Hydrological Region No. 26 of Mexico. The Tempoal River has five hydrometric stations: El Cardon, Los Hules, Terrerillos, Platon Sanchez and Tempoal, whose basin areas are: 609, 1269, 1493, 4700 and 5275 km^2. Figure 2 shows the location and morphology of the Tempoal River, taken from Campos-Aranda (2015).

Tecnología y Ciencias del Agua
OPEN ACCESS
Check for updates

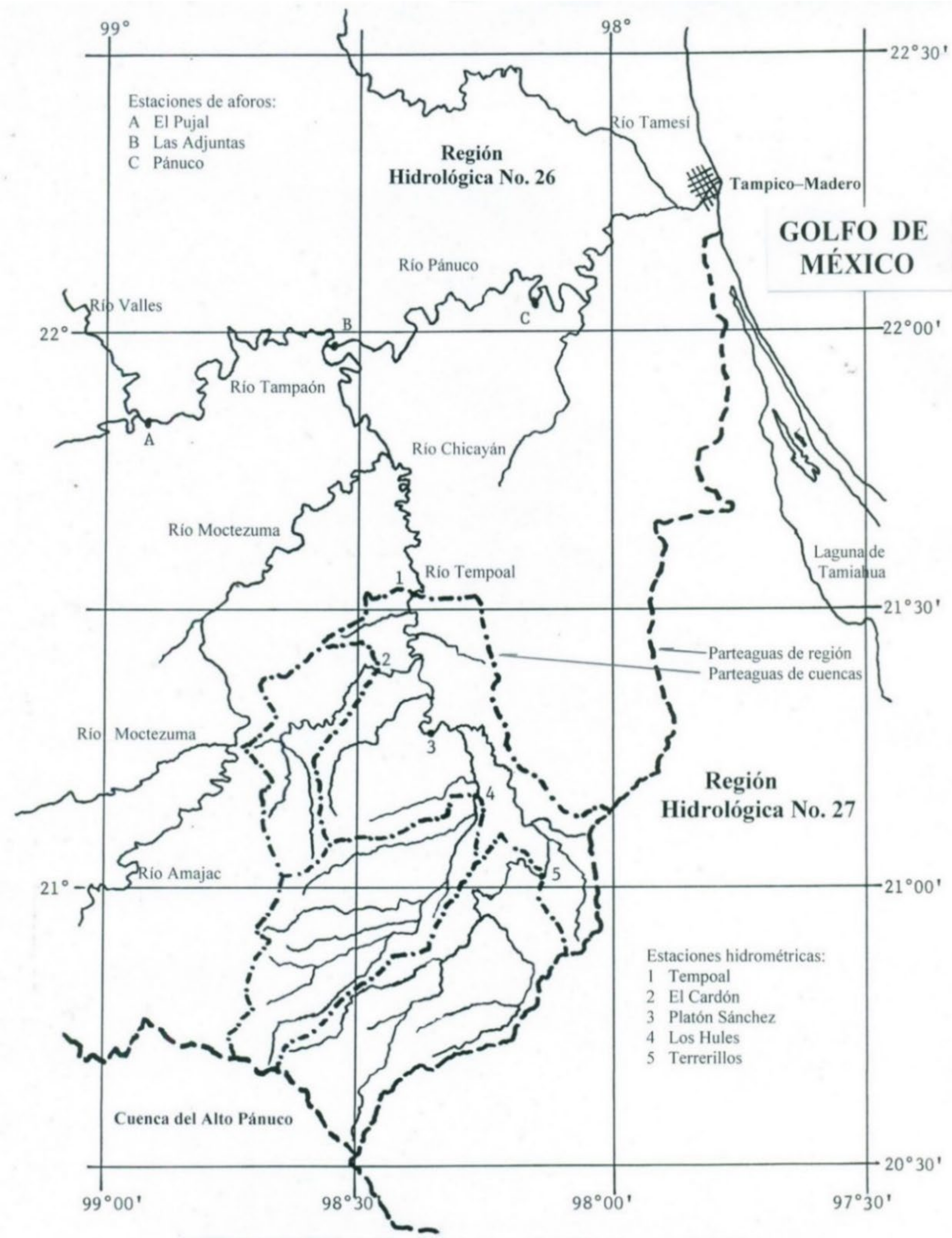

Figure 2. Geographic location and morphology scheme of the Tempoal River, Hydrological Region No. 26 (Panuco), Mexico.

Campos-Aranda (2015) presented the available records of maximum annual flow in m^3/s, in the period from 1960 to 2002 in the five hydrometric stations of the Tempoal River, from the BANDAS system (IMTA, 2003), from the *Monthly Hydrometric Data* file, with nine years per page. Such records are incomplete, with two missing years in 1998 and 2000 in El Cardon, two more in Los Hules in 1990 and 1991, one in Terrerillos in 1981 and 18 missing years in the Platon Sanchez gauging station in the period from 1960 to 1977.

Based on the algorithm of Beale and Little (1975), the missing data and the period without registration of the Platon Sanchez station were estimated simultaneously. The annual flows estimated by Campos-Aranda (2015) are shown in parentheses. In the final column of Table 1 there is the number of occurrences of Equation (24), for the Tempoal-El Cardon-Terrerillos triple, which is defined later.

Column 3 of Table 1 shows the available data and their months of occurrence are in column 2, for the annual floods (*QX*) of the Tempoal station, which is the base. On the other hand, columns 4, 5 and 6 show the annual floods in the auxiliary stations *simultaneous per month*, as explained in the following section.

Simultaneity of events to be analyzed

At the beginning of this century, when frequency analyses of trivariate increases began, for example, those of Zhang and Singh (2007), the processed variables, maximum flow (*Q*), runoff volume (*V*) and total duration (*D*), were derived from the annual flood's hydrograph. The above implies, as already indicated, that the maximum annual flows of the auxiliary stations must correspond to the same event registered in the base station, that is, a *simultaneity of events* must be met.

For the Tempoal base station, the number of months in column 2 of Table 1 establishes a perfectly defined *wet season* from June to October, since only two floods occur outside of it in November 1960 and December 1964. month with the most occurrences is September with 18 events; However, in the remaining months between 5 and 7 events occur. The existence of a clearly defined five-month wet season makes it possible to verify the *simultaneity of floods* per month.

In Table 1, the floods that are not the annual maximums are indicated in shaded form; Therefore, they establish a lack of monthly *simultaneity* with those of the Tempoal station. There are 13 at the El Cardón station, 6 in Los Hules and 10 in Terrerillos.

Regarding the maximum extreme values (*outliers*) of the joint record (1960-2002) in Tempoal, four values greater than 4000 m^3/s are observed, in the years 1974, 1975, 1984 and 1993.

Results and their discussion

Verification of randomness

Based on the Wald-Wolfowitz Test (Bobée & Ashkar, 1991; Rao & Hamed, 2000; Meylan *et al.*, 2012), the independence and stationarity of the four maximum flow records in Table 1 were tested. *U* statistics of such a test are found in the last row of Table 1; since U < 1.96, it can be inferred that the records are random.

Marginal distribution at the Tempoal station

Table 2 shows the results of the fit of the three ideal FDPs according to Weighted Absolute Distance in the L Quotient Diagram (Hosking & Wallis, 1997) and the Kappa and Wakeby models. With respect to the predictions, an excellent similarity is observed, except for the Log-Normal distribution that leads to high values.

Table 2. Fitting errors and predictions (m^3/s) of the three suitable FDPs and two widely used FDPs in the record of maximum annual flows (1960-2002) from the *Tempoal* hydrometric station, Mexico.

FDP	*EEA*	*EAM*	Return periods in years					
			25	50	100	500	1 000	5 000
PAG	152.0	108.1	4 720	5 694	6 695	9 103	10 183	12 792
PE3	151.9	108.1	4 704	5 670	6 656	9 023	10 073	12 581
LN3	183.0	121.3	4 689	5 830	7 097	10 588	12 354	17 148
Kappa	162.7	112.0	4 720	5 766	6 880	9 755	11 130	14 680
Wakeby	158.9	114.6	4 724	5 727	6 765	9 325	10 493	13 373
PAG mod	153.5	111.1	4 652	5 512	6 356	8 256	9 049	10 836

The best distribution, the Generalized Pareto (PAG), establishes low fit errors as seen in Table 2. However, for a probability of non-exceedance

(p) of 1 %, it defines a value of 496.8 m^3/s, which, the lowest value in the record, which is 476.0 m^3/s, generates a negative probability.

To correct the above, the modified version of the L-moment method was applied, presented by Rao and Hamed (2000), and Campos-Aranda (2014), which leads to a value of 458.0 m^3/s for a p = 1 %, the predictions shown in the final row of Table 2 and location (u_1), scale (a_1) and shape (k_1) parameters as follows: 444.2926, 1364.267, 0.026739, its equation is:

$$F(x) = 1 - \left[1 - \frac{k_1(x-u_1)}{a_1}\right]^{1/k_1} \quad (37)$$

Marginal distribution at El Cardon station

Similarly, Table 3 shows the results of the fit in the record of annual floods at the El Cardón station. Again, the best FDP, the PAG defines a value of 93.9 m^3/s for a p = 1 %, which is higher than the minimum recorded of 85.7 m^3/s.

Table 3. Fitting errors and predictions (m^3/s) of the three suitable FDPs and two widely used FDPs in the record of maximum annual flows (1960-2002) of the *El Cardon* hydrometric station, Mexico.

FDP	***EEA***	***EAM***	**Return periods in years**					
			25	**50**	**100**	**500**	**1 000**	**5 000**
PAG	52.3	28.0	1274	1488	1689	2111	2275	2619
PE3	60.6	31.6	1267	1506	1747	2313	2561	3145
LP3	65.7	42.5	1194	1394	1590	2024	2204	2603
Kappa	44.0	29.5	1266	1442	1594	1869	1961	2128
Wakeby	55.1	30.0	1274	1493	1701	2145	2321	2696
PAGmod	45.1	28.3	1246	1425	1585	1891	1999	2207

To correct the above, the modified version of the L-moment method was applied, presented by Rao and Hamed (2000), and Campos-Aranda (2014), which leads to a value of 79.5 m3/s for a p = 1 %, the predictions shown in the final row of Table 3 and location (u_2), scale (a_2) and shape (k_2) parameters as follows: 74.78195, 471.314, 0.168322 with the following equation:

$$F(y) = 1 - \left[1 - \frac{k_2(y-u_2)}{a_2}\right]^{1/k_2} \quad (38)$$

The Pearson type III FDP also leads to a value lower than 85.7 m^3/s for p = 1 % of 73.8 m^3/s, but its fitting errors are high as are its predictions, which is why it was not adopted.

Tecnología y Ciencias del Agua
OPEN ACCESS
Check for updates

Marginal distribution at the Terrerillos station

Finally, based on a procedure similar to that described for the two previous records, the results shown in Table 4 were obtained. It is observed that the Kappa distribution leads to the lowest fitting errors, but all its predictions are considered reduced and furthermore, for *p* = 1 % it defines a value of 212.0 m^3/s, which is higher than the lowest value in the record which was 162.9 m^3/s.

Table 4. Fitting errors and predictions (m^3/s) of the three suitable FDPs and two of general use in the record of maximum annual flows (1960-2002) of the *Terrerillos* hydrometric station, Mexico.

FDP	*EEA*	*EAM*	Return periods in years					
			25	50	100	500	1 000	5 000
PAG	174.9	105.8	3 603	4 173	4 695	5 741	6 129	6 906
LP3	197.2	143.3	3 352	3 871	4 358	5 370	5 759	6 566
PE3	211.7	132.7	3 586	4 256	4 928	6 498	7 182	8 789
Kappa	113.6	95.7	3 524	3 861	4 105	4 441	4 523	4 636
Wakeby	209.5	130.2	3 585	4 241	4 895	6 405	7 053	8 553
PAGmod	150.2	107.2	3 530	4 018	4 443	5 230	5 499	5 997

On the other hand, the PAG distribution also defines a higher value for *p* = 1 % with 180.9 m^3/s. Again, to correct the above, the modified version of the L-moment method was applied, presented by Rao and Hamed (2000), and Campos-Aranda (2014), which leads to a value of

144.1 m³/s for a p = 1 %, the predictions shown in the final row of Table 4 and the following location (u_3), scale (a_3) and shape (k_3) parameters: 129.786, 1430.494, 0.199064, with the following equation:

$$F(z) = 1 - \left[1 - \frac{k_3(z-u_3)}{a_3}\right]^{1/k_3} \quad (39)$$

Predictions for contrast at the Tempoal station

To contrast the *predictions* obtained with the *FC*, in the trivariate frequency analyzes at the Tempoal hydrometric station, with the nearby ones that showed regional dependence, first the predictions for the complete data period at the Tempoal station were estimated, which spanned from 1954 to 2006 (n = 49).

The six data included were the annual increases from 1954 to 1959, which are: 2 110.0, 6 000.0, 4 424.0, 449.0, 4 100.0 and 1 507.6 m³/s. These floods establish a new minimum value of 449.0 m³/s; but the most relevant thing is that in a period of six years, a flood similar to the maximum of the entire 43-year record of 6 120 m³/s occurs, and two more greater than 4 000 m³/s.

The above will surely give rise to larger predictions, as there are a greater number of extreme events. The weighted absolute distances define the first three as ideal FDPs, which are shown in Table 5, results.

Table 5. Fitting errors and predictions (m^3/s) of the three suitable FDPs and two widely used FDPs in the record of maximum annual flows (1954-2002) of the *Tempoal* hydrometric station, Mexico.

FDP	***EEA***	***EAM***	**Return periods in years**					
			25	**50**	**100**	**500**	**1 000**	**5 000**
PE3	253.4	148.4	5 262	6 339	7 436	10 056	11 215	13 976
PAG	246.5	145.5	5 282	6 336	7 394	9 866	10 937	13 438
LP3	261.6	195.4	4 967	5 885	6 807	8 962	9 895	12 063
Kappa	221.3	165.0	5 256	6 158	7 003	8 758	9 433	10 835
Wakeby	253.1	152.3	5 281	6 340	7 404	9 897	10 981	13 519
PAGmod	236.1	146.1	5 253	6 260	7 256	9 524	10 481	12 663

The first two ideal distributions lead to low fitting errors, mainly the mean absolute error, but define values with a probability of non-exceedance of 1 % greater than the minimum of the record (449.0 m^3/s); therefore, they are not acceptable.

The modified version of the L-moment method for fitting the PAG distribution, presented by Rao and Hamed (2000), and Campos-Aranda (2014), leads to a value of 433.0 m^3/s for p = 1 % and the predictions shown in the final row of Table 5, which will be used in the contrasts.

As anticipated, by increasing the six years of registration, the increases in the return periods from 100 to 5 000 years, in Table 5, are higher than those in Table 2, from 14.2 to 16.9 %.

Kendall tau quotients

Based on Equation (7), the Kendall tau ratios shown in Table 6 were calculated for the data from the five hydrometric stations of the Tempoal river system (Table 1), analyzed by pairs. The lowest Kendall tau ratio, between the Los Hules and Terrerillos stations, leads to a value of 3.61 in Equation (8); therefore, the stations are not independent.

Table 6. Bivariate Kendall tau ratios for the annual floods (1960-2002) of the four hydrometric stations of the Tempoal river system, Mexico.

Hydrometric stations	El Cardon	Los Hules	Terrerillos
Tempoal	0.6013	0.5127	0.5991
El Cardon	1	0.4640	0.4839
Los Hules		1	0.4219

For the Tempoal station, its best pairs, with greater correlation or *regional dependence*, are formed with El Cardon and Terrerillos; Logically, the triple to be analyzed is: Tempoal-El Cardon-Terrerillos. It is worth noting that both correlations are similar in magnitude and positive in sign. From the above, it can be intuited that the fit of the symmetrical and asymmetric trivariate *FC* will be similar.

Dependence observed by pairs

Based on Equation (22) and the data in Table 1, the values of the dependence observed in the right tail were calculated, shown in Table 7. It is observed that the highest values of λ_U^{CFG} occur between the record of floods at the Tempoal stations with Terrerillos and with El Cardon; Therefore, it resembles the behavior of Kendall's tau quotients.

Table 7. Values of the dependence in the right tail (λ_U^{CFG}) the annual floods (1960-2002) of the four hydrometric stations of the Tempoal river system, Mexico.

Hydrometric stations	El Cardon	Los Hules	Terrerillos
Tempoal	0.6451	0.5942	0.6693
El Cardón	1	0.5500	0.5186
Los Hules		1	0.5619

When searching in Table 4 presented by Campos-Aranda (2023), approximate values to those defined in Tables 6 and 7 for the Kendall tau and the right tail dependence(λ_U^{CFG}), it was found that the Gumbel-Hougaard *FC* seems to reproduce those magnitudes, therefore, this will be the *FC* to be used for modelling the pairs: Tempoal-El Cardon, Tempoal-Terrerillos and El Cardon-Terrerillos; as well as the Tempoal-El Cardon-Terrerillos triple.

Fitting of *FCs* in bivariate analyzes

As already indicated, for the application of Equation (30), regarding the trivariate return period of type AND, the bivariate *FC*s $C(u, v)$, $C(u, w)$ and $C(v, w)$ are required; which are adopted based on the following process. The last column of Table 8 shows the values of the observed dependence estimator (Equation (22)), for the indicated pairs.

Table 8. Statistical fit indicators of the bivariate Copula functions between the annual floods at the indicated hydrometric stations of the Tempoal river system, Mexico.

FC	***θ***	***EME***	***EAM***	**DP**	**DN**	**MDP**	**MDN**	**(**λ_U^{CFG}**)** λ_U
Estaciones: Tempoal-El Cardon								(0.6451)
Frank	7.964	0.0316	0.0237	17	26	0.0782	−0.0575	0.0000
G-H	2.5083	0.0306	0.0231	20	23	0.0685	−0.0569	0.6817
Estaciones: Tempoal-Terrerillos								(0.6693)
Frank	7.905	0.0372	0.0283	19	24	0.0766	−0.0891	0.0000
G-H	2.4945	0.0381	0.0306	22	21	0.0758	−0.0831	0.6797
Estaciones: El Cardon-Terrerillos								(0.5186)
Frank	5.451	0.0255	0.0194	27	16	0.0493	−0.0652	0.0000
G-H	1.9378	0.0296	0.0222	25	18	0.0641	−0.0743	0.5700

Meaning of the new acronyms:

DP, DN: Number of positive and negative differences.

MDP, MDN: Maximum positive and negative difference.

It is observed that in the three cases the Gumbel-Hougaard *FC* (G-H) provides an extreme right tail dependence (Equation (9)), slightly greater than that observed; Based on the above, it is concluded that the selection of this *FC* was correct.

Selection and ratification of the trivariate *FC*

Symmetric Copula Functions

The trivariate *FC* (d = 3) of Frank and Gumbel-Hougaard, defined by equations (13) and (17), were fitted to the *joint* annual maximum flow data of the Tempoal, El Cardon and Terrerillos stations, taken from Table 1. Such fit was carried out by trial and error of the value of its association parameter (θ), seeking the smallest fitting errors according to expressions 25 to 27. Such calculations were carried out based on a computer program in *Basic*, developed specifically. The results obtained are shown in Table 9.

Table 9. Statistical fit indicators of the indicated *symmetric* trivariate Copula functions, in the triples of annual floods from the Tempoal-El Cardón-Terrerillos stations, Mexico.

FC	***θ***	***EME***	***EAM***	**DP**	**DN**	**MDP**	**MDN**
Frank	7.995	*0.0277*	*0.0213*	20	23	0.0614	−0.0692
G-H	2.795	0.0294	0.0225	19	24	0.0492	−0.0792

Meaning of the new acronyms:

DP, DN: Number of positive and negative differences.

MDP, MDN: Maximum positive and negative difference.

On the other hand, Equation (28) defines D_n = 0.2071 and since the maximum absolute difference of the Frank and symmetric Gumbel-Hougaard *FC*s in Table 9 is 0.0792, the Kolmogorov-Smirnov test allows to adopt any of them. The correlation coefficients (r_{xy}) between the empirical probabilities (Equation (24)) and the theoretical ones, estimated with the Frank and symmetric G-H *FC*s, were 0.9953 and 0.9944; therefore, both *FC*s define good fits.

Asymmetric Copulas Functions

The application of the trivariate asymmetric *FC*, with two association parameters (θ_1, θ_2), to the data in Table 1 for the processed triple, was carried out based on the Complex algorithm of multiple bounded variables. The initial values in Frank's *FC* 2 and 8; in the Gumbel-Hougaard 1.5 and 4. The optimal values found for θ_1 and θ_2 and their fit indicators have been concentrated in Table 10.

Table 10. Statistical indicators of the fit of the *asymmetric* trivariate Copula functions in the triples of annual floods from the Tempoal-El Cardon-Terrerillos stations, Mexico.

FC	θ_1	θ_2	***EME***	***EAM***	**DP**	**DN**	**MDP**	**MDN**
Frank	7.1294	10.4163	*0.0275*	*0.0208*	19	24	0.0611	−0.0659
G-H	2.3475	3.8288	0.0290	0.0227	16	27	0.0573	−0.0773

Meaning of the new acronyms:

θ_1, θ_2: association parameters of asymmetric *FC*.

Again, Frank's *FC* defines the best fit. As already indicated, Equation (28) defines D_n = 0.2071 and since the maximum absolute difference of the Frank and asymmetric Gumbel-Hougaard *FC*s in Table 10 is 0.0773, the Kolmogorov-Smirnov test confirms the adoption of any of them. The correlation coefficients (r_{xy}) between the empirical probabilities (Equation (24)) and the theoretical ones, estimated with Frank's *FC* and asymmetric G-H, were 0.9955 and 0.9945; therefore, both *FC*s show excellent fit.

Adoption of a trivariate *FC*

The result of Table 8, if the Gumbel-Hougaard *FC* is the one adopted, due to the reproduction it makes of the value of λ_U^{CFG} for the three pairs that are established and that were analyzed, guides your selection for the trivariate case.

Such selection is not considered inappropriate, since as observed in Tables 9 and 10, such Gumbel-Hougaard *FC* shows quite similar fits to those of the symmetrical and asymmetric Frank *FC*. The above was verified based on the correlation coefficients (r_{xy}) between the empirical and theoretical trivariate probabilities of both *FC*s, which were practically the same.

Return periods of design floods

Assuming that in the vicinity of waters below the *Tempoal* hydrometric station, dikes are going to be built to protect floodplains for agricultural and industrial purposes and a bridge to cross it, then it is required to estimate design events with joint or trivariate return periods of 50, 100, 500 and 1 000 years. Therefore, it is necessary to estimate flows or *Design Floods* (*QX*) at the *Tempoal* base station with the four T_{KEN} joint return periods mentioned.

Estimation of trivariate design floods

Inequality of trivariate return periods

Once the four design joint return periods are defined, their respective non-exceedance probability (u,v,w) of 0.98, 0.99, 0.998 and 0.999 is applied in Equation (5) with the values of θ shown in Table 8. Furthermore, Equation (17) is applied with θ = 2.795 and Equation (22) with θ_1 = 2.3475 and θ_2 = 3.8288, to obtain the probabilities required by equations (29) and (30) of the OR and AND type return periods. To estimate the T_{KEN}, Equation (35) was applied.

The results in Table 11 show great similarity in the trivariate return periods of the OR and AND type, of the symmetrical and asymmetrical *FC*. Therefore, T_{KEN} can be used to obtain the annual joint or flood design events at the Tempoal base station. The results of column 3 of Table 11 allow us to verify Equation (36), of the inequality of the trivariate *Tr*.

Table 11. Joint return periods of type OR, AND and *Secondary* estimated with the symmetrical and asymmetric Gumbel-Hougaard trivariate *FC*, for the triple of annual floods of the Tempoal-El Cardón-Terrerillos stations, Mexico.

Tr **(years)**	**Type of *Tr* with the symmetrical *FC***			***Asymmetric FC***	
	OR	**Secondary**	**AND**	**OR**	**AND**
50	33.9	94.2	119.0	33.8	118.0
100	67.7	189.0	240.0	67.5	238.1
500	337.6	947.5	1208.8	336.9	1199.1
1 000	675.1	1895.7	2419.6	673.6	2399.8

Estimation of design events

Based on Equation (35) of the trivariate Kendall distribution, established for the symmetric Gumbel-Hougaard *FC*, the univariate return period (*Tr*) and its respective probability of non-exceedance (*s*) were searched by trial and error, which define a *secondary* return period equal to the design joint or trivariate. Once the unit variable value (*s*) is found, the respective *QX* variables are obtained with the inverse solution of the marginal distribution (Equation (37)). The results are presented in Table 12.

Table 12. Design events obtained with the secondary return period that equals the univariate of the *triples of joint floods* of the stations of the Tempoal river system, Mexico.

Tr **(years) univariate**	**Tempoal-El Cardon-Terrerillos**	
	Prob. (*s*) of secondary *Tr*	***QX*** **(m^3/s) value**
50	0.962075	4719
100	0.980916	5569
500	0.996164	7497
1 000	0.998081	8303

The trivariate predictions in Table 12 are lower than the univariate predictions of the Tempoal base station, shown in Table 5; as shown in Figure 3. The design values in Table 12 are lower by 24.6, 23.3, 21.3 and 20.8 %, respectively, in relation to the predictions in Table 5.

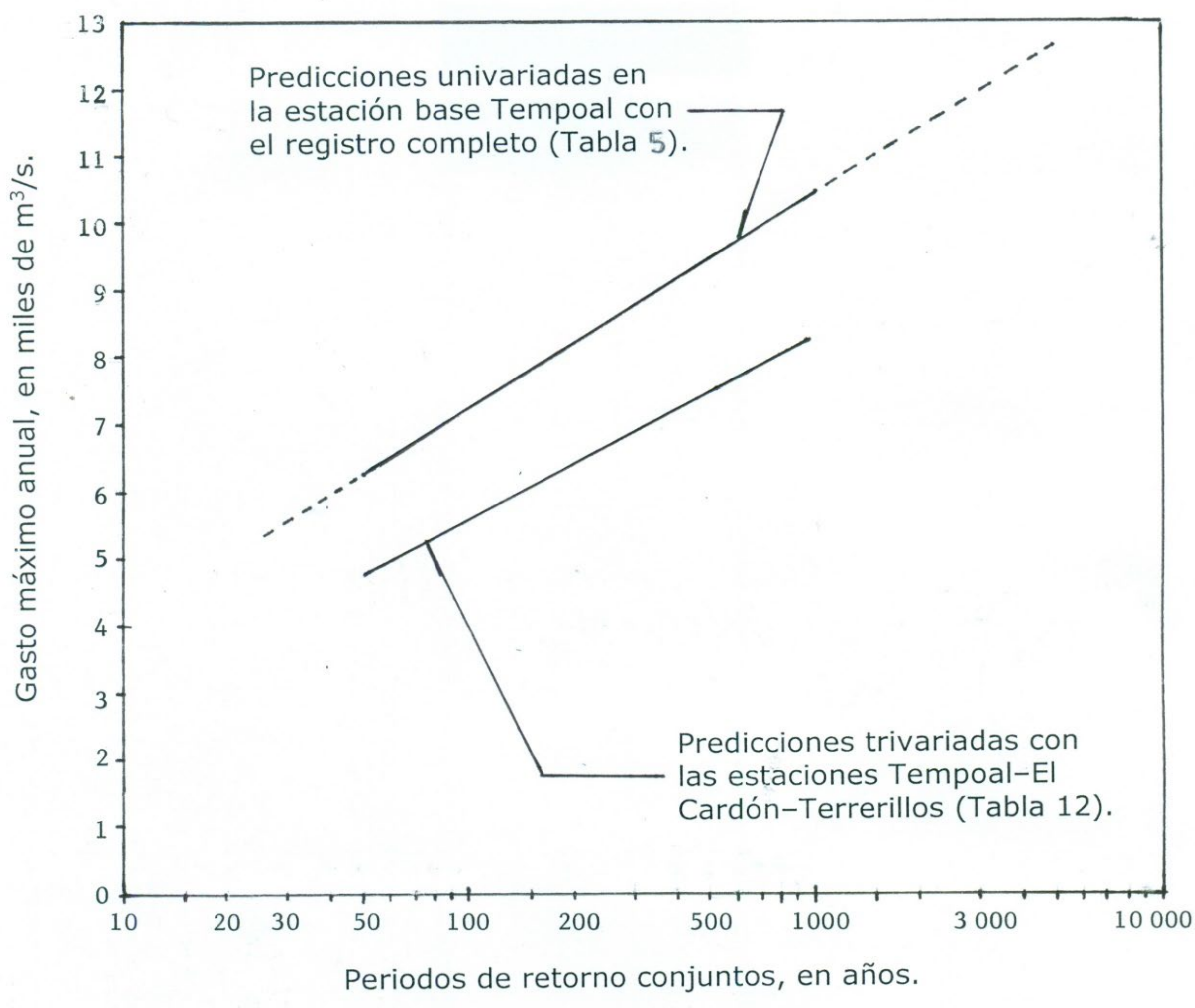

Figure 3. Graph of design predictions obtained with the univariate and trivariate approaches for the annual floods of the Tempoal base station, Mexico.

Another similar study

Campos-Aranda (2022) presents a numerical application in which the predictions obtained with the bivariate frequency analysis of floods with regional dependence were slightly higher than those estimated with the complete record of the base gauging station and therefore, they are the adopted. In these cases, the bivariate flood frequency analysis is considered a success, because it leads to more severe or critical predictions.

Future approach to *FC* application

Bivariate and trivariate flood frequency analyses at a base station or for a project of interest, with regional dependence, that is, with nearby records from auxiliary hydrometric stations that show correlation with that of the base station, should evolve to process marginal FDPs of series or *non-stationary records*. Towards such an eventuality, various approaches have already begun to be suggested, such as that of Bender, Wahl and Jensen (2014), and that of Chebana and Ouarda (2021).

Conclusions

The frequency analyzes of trivariate floods, of the maximum flow variables in the base station (*QX*) and in the auxiliary stations (*QY* and *QZ*), which exhibit a correlation or regional dependence and have the same recording amplitude, will allow a contrast of the univariate estimation of *Design Floods* of the complete record of the base station, against those obtained with the *FC* associated with a joint return period.

The use of *Copula Functions* (*FC*) in trivariate frequency analyzes allows the construction of the joint distribution based on the marginal functions. Therefore, the ideal probability distributions of *QX*, *QY* and *QZ* are defined with the maximum possible accuracy and can be different and of any type.

The estimation of the trivariate return period of the AND type requires bivariate distributions; in the case studied of the variable pairs *QX*-*QY*, *QX*-*QZ* and *QY*-*QZ*. Therefore, first we look for *FC*s that reproduce the observed dependence (λ_U^{CFG}) and show a good fit with the aforementioned joint variables.

In the numerical application described for a common period of 43 annual floods recorded in the Tempoal river system, of Hydrological Region No. 26 (Pánuco), Mexico; Tempoal was used as a base station and El Cardón and Terrerillos as auxiliaries, as they were the ones that showed the greatest correlation. Two FC families were applied: Frank and Gumbel-Hougaard.

For the annual data triples of *QX*, *QY* and *QZ*, symmetric trivariate Archimedean *FC*s were applied, with one association parameter (θ) and asymmetric trivariates, with two association parameters (θ_1, θ_2), from the aforementioned families. Finally, joint return periods of OR, AND and Kendall type were estimated. The latter allow us to obtain the *QX* design events, shown in Table 12.

This type of analysis of flood frequencies with regional dependence leads, in some cases, to predictions greater than the univariate estimates, made with the complete record of the base station. On other occasions, such as the case presented, their predictions are lower and then, they allow the trend of the return period versus design flow relationship to be verified, as shown in Figure 3.

The frequency analyzes of trivariate floods, with regional dependence, described are very simple and do not present computational complications, when carried out based on the *FC*s.

Acknowledgment

We appreciate the comments and corrections suggested by the anonymous referees C, D, E and the thematic editor, which made the work more explicit by changing the order of some themes or calculations; as well as, reinforce theoretical concepts related to Copula Functions and propose options for future research. The reference of a similar bivariate study was also included, to apply regional dependence or correlation.

References

AghaKouchak, A., Sellars, S., & Sorooshian, S. (2013). Chapter 6. Methods of tail dependence estimation. In: AghaKouchak, A., Easterling, D., Hsu, K., Schubert, S., & Sorooshian, S. (eds.). *Extremes in a changing climate* (pp. 163-179). Dordrecht, The Netherlands: Springer.

Aldama, A. A., Ramírez, A. I., Aparicio, J., Mejía-Zermeño, R., & Ortega-Gil, G. E. (2006). *Seguridad hidrológica de las presas en México.* Jiutepec, México: Instituto Mexicano de Tecnología del Agua.

Barbe, P., Genest, C., Ghoudi, K., & Rémillard, B. (1996). On Kendall's Process. *Journal of Multivariate Analysis*, 58(2), 197-229.

Beale, E. M. L. & Little, R. J. A. (1975). Missing values in multivariate analysis. *Journal of Royal Statistical Society Series B*, 37(1), 129-145. DOI: 10.1111/j.2517–6161.1975.tb01037.x

Bender, J., Wahl, T., & Jensen, J. (2014). Multivariate design in the presence of non–stationarity. *Journal of Hydrology*, 514(June), 123-130. DOI: 10.1016/j.jhydrol.2014.04.017

Bobée, B. (1975). The Log-Pearson type 3 distribution and its application to Hydrology. *Water Resources Research*, 11(5), 681-689. DOI: 10.1029/WR011i005p00681

Bobée, B., & Ashkar, F. (1991). Chapter 1. Data requirements for hydrologic frequency analysis. In: *The Gamma Family and derived distributions applied in hydrology* (pp. 1-12). Littleton, USA: Water Resources Publications.

Box, M. J. (1965). A new method of constrained optimization and a comparison with other methods. *Computer Journal*, 8(1), 42-52.

Bunday, B. D. (1985). Theme 6.2. The complex method. In: *Basic optimisation methods* (pp. 98-106). London, England: Edward Arnold publishers, Ltd.

Campos-Aranda, D. F. (2003). Capítulo 7. Integración numérica y Capítulo 9. Optimización numérica. En: *Introducción a los métodos numéricos: software en Basic y aplicaciones en hidrología superficial* (pp. 137-153, 172-211). San Luis Potosí, México: Editorial Universitaria Potosina.

Campos-Aranda, D. F. (2014). Predicción de crecientes usando la distribución Pareto Generalizada ajustada con tres métodos simples. *Tláloc*, 65, octubre-diciembre, 7-26.

Campos-Aranda, D. F. (2015). Estimación simultánea de datos hidrológicos anuales faltantes en múltiples sitios. *Ingeniería. Investigación y Tecnología*, 16(2), 295-306.

Campos-Aranda, D. F. (2022). Análisis de frecuencias de crecientes bivariado con dependencia regional y funciones Cópula. *Aqua-LAC*, 14(2), 47-61. DOI: 10.29104/phi–aqualac/2022–v14–2–11

Campos-Aranda, D. F. (2023). Selección y aplicación de funciones Cópula con dependencia en su extremo derecho al análisis de frecuencias conjunto (*Q*,*V*) de crecientes anuales. *Tecnología y ciencias del agua*, 14(5), 120-188. DOI: 10.24850/j–tyca–14–05–03

Capéraà, P., Fougéres, A. L., & Genest, C. (1997). A nonparametric estimation procedure for bivariate extreme value copulas. *Biometrika*, 84(3), 567-577. DOI: 10.1093/biomet/84.3.567

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247-1250. DOI: 10.5194/gmd–7–1247–2014

Chebana, F., & Ouarda, T. B. M. J. (2021). Multivariate non-stationary hydrological frequency analysis. *Journal of Hydrology*, 593(February), 125907. DOI: 10.1016/j.jhydrol.2020.125907

Chen, L., & Guo, S. (2019). Chapter 2. Copula theory and Chapter 3. Copula-based flood frequency analysis. In: *Copulas and its application in hydrology and water resources* (pp. 13-38, 39-71). Gateway East, Singapore: Springer.

Chowdhary, H., & Singh, V. P. (2019). Chapter 11. Multivariate frequency distributions in hydrology. In: Teegavarapu, R. S. V., Salas, J. D., & Stedinger, J. R. (eds.). *Statistical analysis of hydrologic variables* (pp. 407-489). Reston, USA: American Society of Civil Engineers.

Davis, P. J., & Polonsky, I. (1972). Chapter 25. Numerical interpolation, differentiation and integration. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (9th print.) (pp. 875-926). New York, USA: Dover Publications.

Dupuis, D. J. (2007). Using Copulas in hydrology: Benefits, cautions, and issues. *Journal of Hydrologic Engineering*, 12(4), 381-393. DOI: 10.1061/(ASCE)1084–0699(2007)12:4(381)

Escalante-Sandoval, C. A., & Raynal-Villaseñor, J. A. (1994). A trivariate extreme value distribution applied to flood frequency analysis. *Journal of Research of the National Institute of Standards and Technology*, 99(4), 369-375.

Escalante-Sandoval, C., & Raynal-Villaseñor, J. (2008). Trivariate generalized extreme value distribution in flood frequency analysis. *Hydrological Sciences Journal*, 53(3), 550-567.

Frahm, G., Junker, M., & Schmidt, R. (2005). Estimating the tail–dependence coefficient: Properties and pitfalls. *Insurance: Mathematics and Economics*, 37(1), 80-100. DOI: 10.1016/j–insmatheco.2005.05.008

Genest, C., & Favre, A. C. (2007). Everything you always wanted to know about Copula modeling but were afraid to ask. *Journal of Hydrologic Engineering*, 12(4), 347-368. DOI: 10.1061/(ASCE)1084–0699(2007)12:4(347)

Genest, C., & Chebana, F. (2017). Copula modeling in hydrologic Frequency Analysis. In: Singh, V. P. (ed.). *Handbook of Applied Hydrology* (2nd ed.) (pp. 30.1-30.10). New York, USA: McGraw-Hill Education.

Goel, N. K., Seth, S. M., & Chandra, S. (1998). Multivariate modeling of flood flows. *Journal of Hydraulic Engineering*, 124(2), 146-155.

Gräler, B., van den Berg, M. J., Vandenberghe, S., Petroselli, A., Grimaldi, S., De Baets, B., & Verhoest, N. E. C. (2013). Multivariate return periods in hydrology: A critical and practical review focusing on synthetic design hydrograph estimation. *Hydrology and Earth System Sciences*, 17(4), 1281-1296. DOI: 10.5194/hess–17–1281–2013

Grimaldi, S., & Serinaldi, F. (2006a). Design hyetograph analysis with 3–copula function. *Hydrological Sciences Journal*, 51(2), 223-238. DOI: 10.1623/hysj.51.2.223

Grimaldi, S., & Serinaldi, F. (2006b). Asymmetric copula in multivariate flood frequency analysis. *Advances in Water Resources*, 29(8), 1155–1167. DOI: 10.1016/j.advwatres.2005.09.005

Hosking, J. R. M. (1994). The four-parameter Kappa distribution. IBM *Journal of Research and Development*, 38(3), 251-258.

Hosking, J. R., & Wallis, J. R. (1997). Appendix: L-moments for some specific distributions. In: *Regional Frequency Analysis. An approach based on L–moments* (pp. 191-209). Cambridge, England: Cambridge University Press.

IMTA, Instituto Mexicano de Tecnología del Agua. (2003). *Banco Nacional de Datos de Aguas Superficiales* (*Bandas*) (8 CD). Jiutepec, México: Comisión Nacional del Agua, Secretaría de Medio Ambiente y Recursos Naturales, Instituto Mexicano de Tecnología del Agua.

Kite, G. W. (1977). Chapter 12. Comparison of frequency distributions. In: *Frequency and risk analyses in hydrology* (pp. 156-168). Fort Collins, USA: Water Resources Publications.

Kjeldsen, T. R., Ahn, H., & Prosdocimi, L. (2017). On the use de a four–parameter kappa distribution in regional frequency analysis. *Hydrological Sciences Journal*, 62(9), 1354-1363. DOI: 10.1080/02626667.2017.1335400

Meylan, P., Favre, A. C., & Musy, A. (2012). Chapter 3. Selecting and checking data series and Theme 9.2. Multivariate Frequency Analysis using Copulas. In: *Predictive hydrology. A frequency analysis approach* (pp. 29-70, 164-176). Boca Raton, USA: CRC Press.

Nelsen, R. B. (2006). Chapter 4. Archimedean Copulas. In: *An introduction to Copulas* (2nd ed.) (pp. 109-155). New York, USA: Springer Series in Statistics.

Nieves, A., & Domínguez, F. C. (1998). Secciones 6.2 y 6.3. Cuadratura de Gauss e integrales múltiples. *Métodos numéricos. Aplicados a la ingeniería* (pp. 416-434). México, DF, México: Compañía Editorial Continental.

Poulin, A., Huard, D., Favre, A. C., & Pugin, S. (2007). Importance of tail dependence in bivariate frequency analysis. *Journal of Hydrologic Engineering*, 12(4), 394-403. DOI: 10.1061/(ASCE)1084–0699(2007)12:4(394)

Rao, A. R., & Hamed, K. H. (2000). Theme 1.8. Tests on hydrologic data and Theme 8.3: The Generalized Pareto distribution. *Flood frequency analysis* (pp. 12-21, 271-290). Boca Raton, USA: CRC Press.

Requena, A. I., Mediero, L., & Garrote, L. (2013). A bivariate return period based on copulas for hydrologic dam design: Accounting for reservoir routing in risk estimation. *Hydrology and Earth System Sciences*, 17(8), 3023-3038. DOI: 10.5194/hess-17-3023-2013

Salvadori, G., & De Michele, C. (2004). Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resources Research*, 40(W12511), 1-17. DOI: 10.1029/2004WR003133

Salvadori, G., & De Michele, C. (2007). On the use of Copulas in Hydrology: Theory and Practice. *Journal of Hydrologic Engineering*, 12(4), 369-380. DOI: 10.1061/(ASCE)1084–0699(2007)12:4(369)

Salvadori, G., De Michele, C., Kottegoda, N. T., & Rosso, R. (2007). Chapter 3. Bivariate analysis via Copulas; Appendix B: Dependence, and Appendix C: Families of Copulas. *Extremes in nature. An approach using Copulas* (pp. 131-175, 219-232, 233-269). Dordrecht, The Netherlands: Springer.

Salvadori, G., De Michele, C., & Durante, F. (2011). On the return period and design in a multivariate framework. *Hydrology and Earth System Sciences,* 15(11), 3293-3305. DOI: 10.5194/hess–15–3293–2011

Stedinger, J. R. (2017). Flood frequency analysis. In: Singh, V. P. (ed.). *Handbook of applied hydrology* (2nd ed.) (pp. 76.1-76.8). New York, USA: McGraw-Hill Education.

Stegun, I. A. (1972). Chapter 27. Miscellaneous functions. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (9th print.) (pp. 997-1010). New York, USA: Dover Publications.

WRC, Water Resources Council. (1977). *Guidelines for determining flood flow frequency* (revised edition). Bulletin #17A of the Hydrology Committee. Washington, DC, USA: Water Resources Council.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79-82. DOI: 10.3354/cr030079

Xu, C., Yin, J., Guo, S., Liu, Z., & Hong, X. (2016). Deriving design flood hydrograph based on conditional distribution: A case study of Danjiangkou reservoir in Hanjiang basin. *Mathematical Problems in Engineering*, 2016(4319646), 1-16. DOI: 10.1155/2016/4319646

Yue, S., Ouarda, T. B. M. J., Bobée, B., Legendre, P., & Bruneau, P. (1999). The Gumbel mixed model for flood frequency analysis. *Journal of Hydrology*, 226(1-2), 88-100. DOI: 10.1016/S0022–1694(99)00168–7

Yue, S. (2000). Joint probability distribution of annual maximum storm peaks and amounts as represented by daily rainfalls. *Hydrological Sciences Journal*, 45(2), 315-326. DOI: 10.1080/02626660009492327

Yue, S., & Rasmussen, P. (2002). Bivariate frequency analysis: Discussion of some useful concepts in hydrological application. *Hydrological Processes*, 16(14), 2881-2898. DOI: 10.1002/hyp.1185

Zhang, L., & Singh, V. P. (2006). Bivariate flood frequency analysis using the Copula method. *Journal of Hydrologic Engineering*, 11(2), 150-164. DOI: 10.1061/(ASCE)1084–0699(2006)11:2(150)

Zhang, L., & Singh, V. P. (2007). Trivariate flood frequency analysis using the Gumbel-Hougaard Copula. *Journal of Hydrologic Engineering*, 12(4), 431-439. DOI: 10.1061/(ASCE)1084–0699(2007)12:4(431)

Zhang, L., & Singh, V. P. (2019). Chapter 3. Copulas and their properties and Chapter 4. Symmetric Archimedean Copulas. *Copulas and their applications in water resources engineering.* pp. (pp. 62-122, 123-171). Cambridge, United Kingdom: Cambridge University Press.