

DOI: 10.24850/j-tyca-2024-05-02

Artículos

## **Análisis de frecuencias de crecientes trivariado ( $Q$ , $V$ , $D$ ) a través de funciones Cópula**

### **Trivariate flood frequency analysis ( $Q$ , $V$ , $D$ ) through Copula Functions**

Daniel Francisco Campos-Aranda<sup>1</sup>, ORCID: <https://orcid.org/0000-0002-9876-3967>

<sup>1</sup>Profesor jubilado de la Universidad Autónoma de San Luis Potosí, San Luis Potosí, México, [campos\\_aranda@hotmail.com](mailto:campos_aranda@hotmail.com)

Autor para correspondencia: Daniel Francisco Campos-Aranda, [campos\\_aranda@hotmail.com](mailto:campos_aranda@hotmail.com)

#### **Resumen**

El análisis de frecuencias de crecientes *trivariado*, del gasto máximo ( $Q$ ), el volumen escurrido ( $V$ ) y la duración total ( $D$ ) permite estimar con mayor exactitud el hidrograma de la creciente de diseño. Para procesar registros anuales conjuntos de  $Q$  y  $V$  disponibles se propuso estimar  $D$  como la duración del hidrograma Gamma hasta el 0.1 % del  $Q$ . Después, a cada registro de  $Q$ ,  $V$  y  $D$  se le busca su distribución de probabilidades idónea para obtener las funciones marginales. En seguida, se adopta la

*función Cópula (FC)* que mejor representa a las variables conjuntas  $Q$ - $V$ ,  $Q$ - $D$  y  $V$ - $D$ . Para estas búsquedas y las trivariadas subsecuentes, se trabajó con las *FC* de Clayton, Frank, Gumbel-Hougaard y Joe. En ambos casos, la selección de la mejor *FC* se basa en los errores de ajuste entre las probabilidades empíricas y teóricas. A las ternas de datos  $Q$ ,  $V$  y  $D$  se les buscó las *FC* de mejor ajuste simétricas y asimétricas de las cuatro familias citadas. A continuación se calculan los periodos de retorno conjuntos de tipo OR, AND y de Kendall. Estos últimos permiten la estimación de los *eventos de diseño* de  $Q$ ,  $V$  y  $D$ . Se describe el análisis de frecuencias trivariado para las 55 crecientes anuales de la estación hidrométrica La Cuña de la Región Hidrológica No. 12-3 (Río Santiago), México. Por último, se formulan las conclusiones, que destacan la sencillez de los análisis de frecuencias trivariados cuando se realizan con *FC*.

**Palabras clave:** funciones Cópula, cociente tau de Kendall, dependencia observada, funciones Cópula multivariadas simétricas, funciones Cópula trivariadas asimétricas, periodos de retorno conjuntos, eventos de diseño.

## Abstract

The *trivariate* flood frequencies analysis, of the maximum flow ( $Q$ ), the runoff volume ( $V$ ) and the total duration ( $D$ ), allows to estimate with greater accuracy the Hydrograph of the Design Flood. To process available joint annual records of  $Q$  and  $V$ , it was proposed to estimate  $D$  as the duration of the Gamma hydrograph up to 0.1 % of  $Q$ . Then, each record of  $Q$ ,  $V$  and  $D$  is searched for its ideal probability distribution to obtain the marginal functions. Next, the *Copula Function (CF)* that best represents the joint variables  $Q$ - $V$ ,  $Q$ - $D$  and  $V$ - $D$  is adopted. For these searches and the subsequent trivariates, we worked with the *CFs* of Clayton, Frank,

Gumbel-Hougaard and Joe. In both cases, the selection of the best  $CF$  is based on the fit errors between the empirical and theoretical probabilities. For the triads of data  $Q$ ,  $V$ , and  $D$ , the symmetrical and asymmetrical best-fit  $CF$ s of the four families mentioned were sought. Next, the joint return periods of type OR, AND and Kendall are calculated. The latter allow the estimation of the *design events* of  $Q$ ,  $V$  and  $D$ . The trivariate frequency analysis is described for the 55 annual floods of the La Cuña hydrometric station of the Hydrological Region No. 12-3 (Santiago River), Mexico. Finally, the conclusions are formulated, which highlight the simplicity of the trivariate frequency analysis, when performed with  $CF$ .

**Keywords:** Copula functions, Kendall's tau ratio, observed dependence, symmetric multivariate Copula Functions, asymmetric trivariate Copula Functions, joint return periods, design events.

Recibido: 19/09/2022

Aceptado: 04/04/2023

Publicado Online: 17/04/2023

## Introducción

### Generalidades

Los primeros análisis de frecuencias de crecientes *multivariados* datan de finales del siglo pasado. Tales estudios procesaron tres variables: gasto máximo o pico ( $Q$ ), volumen escurrido ( $V$ ) y duración total ( $D$ ), pero aceptaron varias consideraciones restrictivas. Las más comunes en tales

estudios *trivariados* fueron las siguientes: (1) considerar que las variables  $Q$ ,  $V$  y  $D$  son independientes; (2) aceptar que las distribuciones marginales son iguales, y (3) adoptar que tales distribuciones son normales, o bien, aplicar una transformación hacia la normalidad (Goel, Seth, & Chandra, 1998; Zhang & Singh, 2007).

Yue, Ouarda, Bobée, Legendre y Bruneau (1999) aplican una distribución bivariada de tipo extremo, la Gumbel, para procesar dos parejas de variables  $Q$ - $V$  y  $V$ - $D$ . Este estudio considera la correlación entre las variables, pero sus marginales son del mismo tipo.

En realidad, las variables  $Q$ ,  $V$  y  $D$  de las crecientes anuales están correlacionadas y no proceden de distribuciones Normales (Zhang & Singh, 2007). Por otra parte, el *hidrograma de diseño* que permiten formar debe tener una magnitud o probabilidad de excedencia que garantice una seguridad al embalse que se diseña o revisa acorde con los estándares establecidos (Gräler *et al.*, 2013; Xu, Yin, Guo, Liu, & Hong, 2016).

Para superar las desventajas citadas, inherentes a las *funciones de distribuciones probabilidad* (FDP), que se aplican en los análisis de frecuencias de crecientes multivariados, se ha recurrido en años recientes a la *funciones Cópula* (FC), modelos matemáticos que se basan en la correlación que exhiben las variables  $Q$ ,  $V$  y  $D$ , y que permiten construir la FDP multivariada exclusivamente a partir de las funciones marginales previamente adoptadas (Salvadori & De Michele, 2004; Salvadori & De Michele, 2007; Favre, El Adlouni, Perreault, Thiémonge, & Bobée, 2004; Meylan, Favre, & Musy, 2012; Genest & Chebana, 2017; Campos-Aranda, 2024).

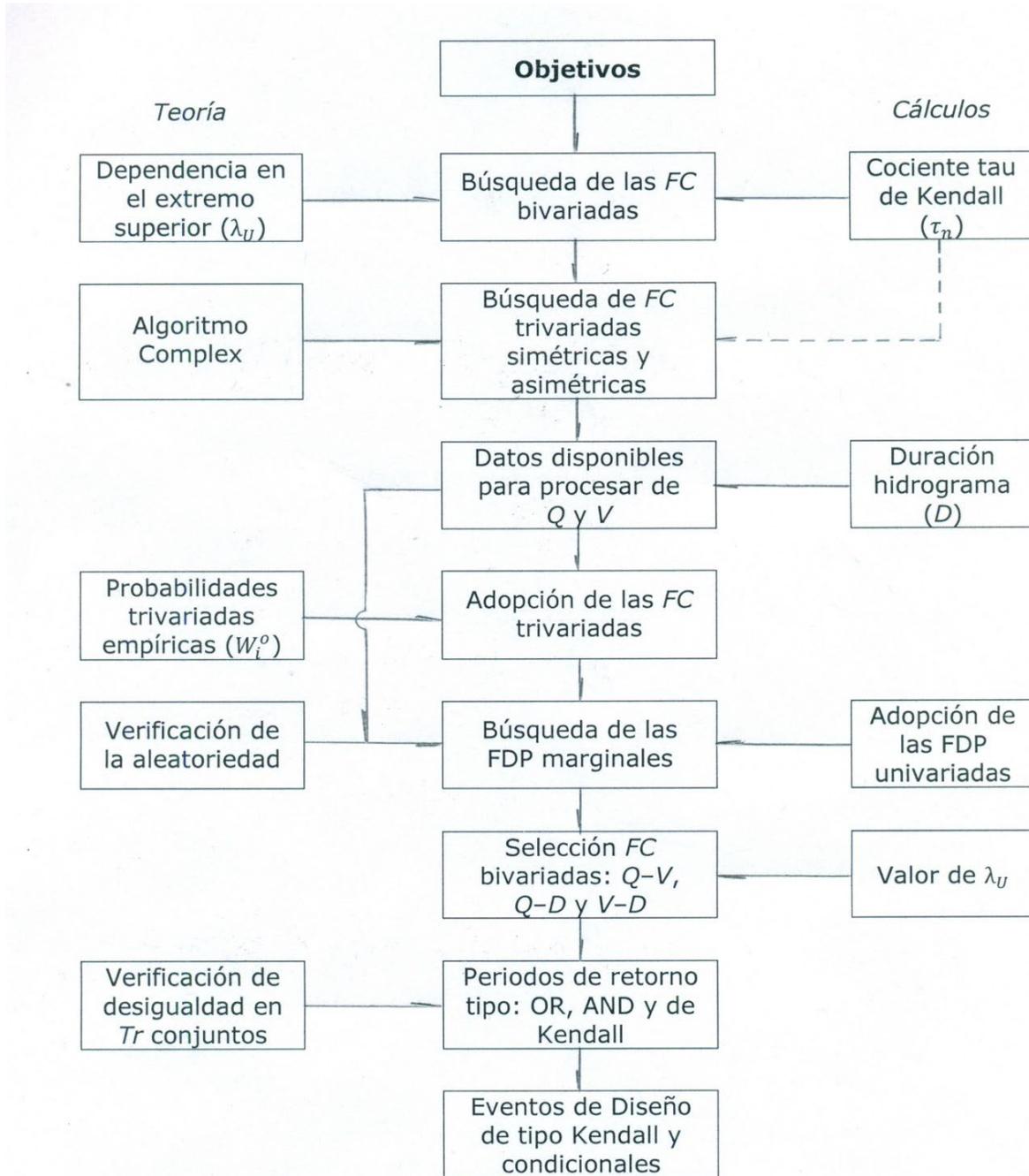
Respecto a los análisis de frecuencias *trivariados*, sus inicios se encuentran en Grimaldi y Serinaldi (2006a), quienes emplearon siete *FC* multivariadas simétricas para obtener un hietograma de diseño. Zhang y Singh (2007) procesan datos de  $Q$ ,  $V$  y  $D$  anuales con base en la *FC* de Gumbel-Hougaard; sus resultados se contrastan contra los de la distribución normal trivariada. Grimaldi y Serinaldi (2006b) describen y aplican las *FC* multivariadas asimétricas, las cuales se exponen posteriormente. Ma, Song, Ren, Jiang y Song (2013) procesan con el enfoque trivariado las sequías anuales, usando *FC* Gaussiana y de la  $t$  de Student. Por último, Zhang y Singh (2019) abordan de manera exhaustiva los análisis de frecuencias de crecientes trivariados mediante *FC*.

## Objetivos

Para este estudio se formularon los siguientes siete objetivos: (1) al disponer de datos conjuntos de  $Q$  y  $V$ , se introdujo como tercera variable aleatoria de las crecientes su duración total ( $D$ ); (2) se trabajó con cuatro familias de *FC* de Arquímedes, para los análisis bivariados  $Q$ - $V$ ,  $Q$ - $D$  y  $V$ - $D$ : Clayton, Frank, Gumbel-Hougaard y Joe; (3) las familias citadas de *FC* se aplicaron en los análisis trivariados, con sus versiones multivariadas denominadas simétricas; (4) las familias de *FC* citadas, de tipo anidado o asimétricas con dos parámetros de asociación, también se emplearon en los análisis trivariados; (5) se estimaron periodos de retorno trivariados de tipo OR, AND y de Kendall; (6) se estiman eventos de diseño, a partir del periodo de retorno conjunto de Kendall, y (7) la teoría expuesta se aplicó al registro conjunto de  $Q$  y  $V$  de las 55 crecientes anuales

registradas en la estación hidrométrica La Cuña de la Región Hidrológica No. 12-3 (Río Santiago), México.

En la Figura 1 se esquematiza la secuencia de temas teóricos y sus cálculos complementarios, con la idea de formular un diagrama de flujo de todo el estudio.



**Figura 1.** Diagrama de flujo de conceptos teóricos y sus cálculos complementarios realizados en el estudio.

## Tópicos de las funciones Cópula

### Ventajas y definición

La ventaja fundamental de las *funciones Cópula (FC)* consiste en permitir formar y expresar la *distribución conjunta* de variables aleatorias que están correlacionadas, como una función de sus distribuciones marginales, previamente adoptadas. Entonces, una *FC* enlaza o relaciona las distribuciones marginales univariadas para formar una *distribución multivariada*. Otra ventaja básica de las *FC* al formar distribuciones multivariadas es el hecho de que ellas separan el efecto de la dependencia entre las variables aleatorias de los efectos de las distribuciones marginales en la modelación conjunta.

Por lo anterior, la construcción de la distribución multivariada se reduce al estudio de la relación que guardan las variables correlacionadas si las distribuciones marginales univariadas son conocidas. El uso de las *FC* ofrece entera libertad para adoptar o seleccionar las distribuciones marginales univariadas que mejor representen a los datos (Shiau, Wang, & Tsai, 2006; Salvadori, De Michele, Kottegodda, & Rosso, 2007; Meylan *et al.*, 2012; Zhang & Singh, 2019).

Como en este estudio aplicarán las *FC* en el *análisis de frecuencias trivariado* de las crecientes anuales, la definición siguiente se refiere a tres variables aleatorias  $X$ ,  $Y$  y  $Z$  correlacionadas, cuya función de distribución de probabilidades acumuladas conjuntas es  $F_{XYZ}(x, y, z)$  con distribuciones marginales univariadas  $F_X(x)$ ,  $F_Y(y)$  y  $F_Z(z)$ ; entonces, la *FC* existe, es  $C[\cdot]$  y es tal que:

$$F_{XYZ}(x, y, z) = C[F_X(x), F_Y(y), F_Z(z)] \quad (1)$$

La ecuación anterior define el concepto básico para el desarrollo de las *FC* y se conoce el teorema de Sklar expuesto en 1959 (Nelsen, 2006; Shiau *et al.*, 2006; Meylan *et al.*, 2012; Zhang & Singh, 2019; Chowdhary & Singh, 2019).

## Familias de Cópulas

Las *funciones Cópula (FC)* que han sido desarrolladas se han clasificado en cuatro clases: de Arquímedes, de valores extremos, elípticas y misceláneas. También se clasifican en Cópulas de un parámetro o de varios, dependiendo de la amplitud con la cual la estructura de la dependencia entre las variables *Q*, *V* y *D* es definida (Meylan *et al.*, 2012; Chowdhary & Singh, 2019). Salvadori *et al.* (2007) exponen un resumen amplio y útil de *FC* que se han aplicado en el campo de la hidrología.

## Cópulas de Arquímedes bivariadas

Las Cópulas de Arquímedes han tenido aplicación amplia debido a su construcción simple, un único parámetro, rango amplio y aceptación de ambos tipos de dependencia (positiva y negativa). Designando a  $F_X(x) = u$ ,  $F_Y(y) = v$  y  $\theta$  el parámetro que mide la dependencia o asociación entre  $u$  y  $v$ , se tienen las siguientes cuatro familias de Cópulas de Arquímedes, las cuales aceptan dependencia positiva y/o negativa (Shiau *et al.*, 2006;

Genest & Favre, 2007; Salvadori *et al.*, 2007; Zhang & Singh, 2019; Chen & Guo, 2019; Chowdhary & Singh, 2019).

**1. Clayton.** Esta función Cópula es llamada de Cook-Johnson por Zhang y Singh (2006). Su ecuación y espacio de variación de  $\theta$  son:

$$C(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \quad [-1, \infty) \setminus \{0\} \quad (2)$$

Para la dependencia positiva  $\theta > 0$  y para la negativa  $-1 \leq \theta < 0$ , con  $\theta = 0$  para la independencia entre  $u$  y  $v$ . La relación de  $\theta$  con el cociente tau de Kendall es la siguiente:

$$\tau_n = \frac{\theta}{\theta+2} \quad (3)$$

**2. Frank.** Su ecuación y espacio de variación de  $\theta$  son:

$$C(u, v) = -\frac{1}{\theta} \ln \left[ 1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right] \quad (-\infty, \infty) \setminus \{0\} \quad (4)$$

Para la dependencia negativa  $0 \leq \theta < 1$  y para la positiva  $\theta > 1$ , con  $\theta = 1$  para la independencia entre  $u$  y  $v$ . La relación de  $\theta$  con el  $\tau_n$  es la siguiente:

$$\tau_n = 1 + \frac{4}{\theta} [D_1(\theta) - 1] \quad (5)$$

siendo  $D_1(\theta)$  la función Debye de orden 1, cuya expresión es:

$$D_1(\theta) = \frac{1}{\theta} \int_0^\theta \frac{s}{e^s - 1} ds \quad (6)$$

en la cual  $s$  es la variable aleatoria unitaria  $0 < s \leq 1$ . La ecuación anterior se estimó con integración numérica, con base en la Ecuación (13), ratificando sus resultados con los valores tabulados por Stegun (1972).

**3. Gumbel-Hougaard**, la cual acepta sólo dependencia positiva. Su ecuación y espacio de variación de  $\theta$  son:

$$C(u, v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right\} [1, \infty) \quad (7)$$

Con  $\theta = 1$  se tiene independencia entre  $u$  y  $v$ . La relación de  $\theta$  con el cociente tau de Kendall es la siguiente:

$$\tau_n = \frac{\theta - 1}{\theta} \quad (8)$$

**4. Cópula de Joe**. La ecuación básica de esta familia de  $FC$ , que sólo acepta dependencia positiva, es la siguiente (Joe, 1993; Salvadori *et al.*, 2007; Chowdhary & Singh, 2019):

$$C(u, v) = 1 - [(1 - u)^\theta + (1 - v)^\theta - (1 - u)^\theta(1 - v)^\theta]^{1/\theta} \quad (9)$$

cuyo parámetro de dependencia es  $\theta \geq 1$ , con  $\theta = 1$ , para el caso de independencia entre  $u$  y  $v$ . Como no tiene una expresión que relacione su parámetro  $\theta$  con el cociente tau de Kendall, éste se estima en función de su ecuación del generador  $\varphi(s)$  de la FC y de su derivada  $\varphi'(s)$ , cuya expresión es (Salvadori *et al.*, 2007; Chowdhary & Singh, 2019):

$$\tau_n = 1 + 4 \int_0^1 \frac{\varphi(s)}{\varphi'(s)} ds \quad (10)$$

en la cual  $s$  es la variable aleatoria unitaria  $0 < s \leq 1$ . Las ecuaciones de su generador y de la derivada de éste son (Salvadori *et al.*, 2007; Zhang & Singh, 2019):

$$\varphi(s) = -\ln[1 - (1 - s)^\theta] \quad (11)$$

$$\varphi'(s) = \frac{\theta(1-s)^{\theta-1}}{(1-s)^\theta - 1} \quad (12)$$

Teniendo como dato el valor de cociente tau de Kendall, la Ecuación (10) se integró numéricamente con base en la Ecuación (13) para obtener por tanteos el valor de  $\theta$  que la cumple.

## Integración numérica

Para cuantificar las integrales, por ejemplo, las ecuaciones (6) y (10), se realizó una integración numérica con base en el método de la cuadratura Gauss-Legendre, cuya ecuación operativa es (Nieves & Domínguez, 1998; Campos-Aranda, 2003):

$$\int_a^b f(x) dx \cong \frac{b-a}{2} \sum_{i=1}^{np} w_i \cdot f \left[ \frac{(b-a)h_i + b + a}{2} \right] \quad (13)$$

en la cual  $w_i$  son los coeficientes del método cuyas abscisas son  $h_i$  y  $np$  el número de parejas en donde se evalúa la función  $f(x)$ , con el argumento indicado en  $f[\cdot]$ . En Davis y Polonsky (1972) se obtuvieron las 12 parejas utilizadas de  $w_i$  y  $h_i$  con 15 dígitos, que son aceptables en el lenguaje *Basic* como variables de doble precisión.

## Indicador numérico de asociación

### Concordancia

Como la *FC* caracteriza la *dependencia* entre las variables aleatorias  $u$  y  $v$ , es necesario el estudio de las medidas de asociación para disponer de un método que permita estimar su parámetro  $\theta$ . En términos generales, una variable aleatoria es *concordante* con otra cuando sus grandes valores están asociados con los grandes valores de la otra y los valores pequeños de una con los valores reducidos de la otra (Salvadori *et al.*, 2007; Chowdhary & Singh, 2019).

Unas variables con correlación lineal directa serán concordantes, pues al aumentar una la otra también lo hace. Variables con correlación lineal inversa serán *discordantes*, ya que a grandes valores de una le corresponderán pequeños valores de la otra y viceversa. Lo anterior implica que son *concordantes* ( $c$ ) las parejas  $(x_1 - x_2)(y_1 - y_2) > 0$  y *discordantes* ( $d$ ) cuando  $(x_1 - x_2)(y_1 - y_2) < 0$  (Salvadori *et al.*, 2007; Chowdhary & Singh, 2019).

### Cociente tau de Kendall

Es un indicador numérico no paramétrico que mide la probabilidad de tener parejas concordantes, por lo cual es el *cociente* de  $c - d$  entre  $c + d$ . Su expresión para estimarlo con datos bivariados es (Zhang & Singh, 2006; Zhang & Singh, 2019):

$$\tau_n = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{signo}[(x_i - x_j)(y_i - y_j)] \quad (14)$$

en la ecuación anterior,  $n$  es el número de observaciones y el  $\text{signo}[\cdot]$  es +1 si tales parejas son concordantes y -1 si son discordantes.

Genest y Favre (2007) exponen una prueba para el cociente tau, que acepta la hipótesis nula  $H_0$  de  $X$  y  $Y$  *independientes*, y entonces su estadístico tiene distribución aproximadamente normal con media 0 y varianza  $2(2n + 5) / [9n(n - 1)]$ . Por lo tanto,  $H_0$  será rechazada con un nivel de confianza  $\alpha = 5 \%$  si:

$$\sqrt{\frac{9n(n-1)}{2(2n+5)}} |\tau_n| > Z_{\alpha/2} = 1.96 \quad (15)$$

## Dependencia en los extremos de las *FC* bivariadas

### Generalidades

El criterio más importante que se aplica para seleccionar una *FC* bivariada es el basado en la magnitud de la dependencia en la cola superior de la distribución conjunta, lo cual tiene impacto en la veracidad de las predicciones extremas. La dependencia en la cola superior derecha ( $\lambda_U$ ) es la probabilidad condicional de que  $Y$  sea mayor que un cierto percentil ( $s$ ) de  $F_Y(y)$ , dado que  $X$  es mayor que tal percentil en  $F_X(x)$ , conforme  $s$  se aproxima a la unidad. La dependencia en la cola inferior izquierda ( $\lambda_L$ ) compara que  $Y$  sea menor que  $X$ , cuando  $s$  se aproxima a cero (Chowdhary & Singh, 2019).

En relación con las *FC* bivariadas expuestas, la de Frank tiene dependencias insignificantes en sus zonas extremas: por ello,  $\lambda_L = 0$  y  $\lambda_U = 0$ . La de Clayton tiene dependencia significativa en su cola inferior igual a  $\lambda_L = 2^{-1/\theta}$ . Por el contrario, las Cópulas de Gumbel-Hougaard y de Joe tienen dependencia significativa en la cola superior igual a:

$$\lambda_U = 2 - 2^{1/\theta} \quad (16)$$

Dupuis (2007) probó seis familias de Cópulas y encontró que su capacidad para estimar eventos extremos varía de mala a buena, con el

siguiente orden: Clayton, Frank, Normal, *t*-Student, Gumbel-Hougaard y de Clayton Asociada (*Survival Clayton*). A conclusiones similares llegan Poulin, Huard, Favre y Pugin (2007), al comparar las mismas seis familias de Cópulas y la denominada A12, que tiene dependencia significativa en su cola derecha.

### Estimación de la dependencia observada

Para abordar la estimación de la dependencia en la cola superior ( $\lambda_U$ ) que muestran los datos disponibles, se debe definir primero la llamada *Cópula Empírica*. Como la *FC* que caracteriza la dependencia entre las variables aleatorias  $X$  y  $Y$ , entonces el par de rangos  $R_i$  y  $S_i$  procedentes de tales variables son el estadístico que retiene la mayor cantidad de información, y su escalamiento con el factor  $1/(n + 1)$  genera una serie de puntos en el cuadrado unitario  $[0, 1]^2$ , formando el dominio de la Cópula empírica (Chowdhary & Singh, 2019), definida así:

$$C_n(u, v) = \frac{1}{n} \sum_{i=1}^n 1\left(\frac{R_i}{n+1} \leq u, \frac{S_i}{n+1} \leq v\right) \quad (17)$$

En la ecuación anterior,  $1(\cdot)$  indica una función de las variables aleatorias  $U$  y  $V$ , que son una transformación en aumento continuo de  $X$  y  $Y$ , en relación con las integrales de probabilidad empírica  $F_n(X)$  y  $F_n(Y)$ , cuyas ecuaciones son:

$$U_i = \frac{\text{Rango}(X_i)}{n+1} = F_n(X_i) \quad V_i = \frac{\text{Rango}(Y_i)}{n+1} = F_n(Y_i) \quad (18)$$

Poulin *et al.* (2007) utilizan el estimador propuesto por Frahm, Junker y Schmidt (2005), el cual está basado en una muestra aleatoria obtenida de la Cópula empírica. Su expresión es:

$$\lambda_U^{CFG} = 2 - 2 \exp \left\{ \frac{1}{n} \sum_{i=1}^n \ln \left[ \sqrt{\ln \frac{1}{U_i} \cdot \ln \frac{1}{V_i}} / \ln \left( \frac{1}{\max(U_i, V_i)^2} \right) \right] \right\} \quad (19)$$

Este estimador acepta que la *FC* se puede aproximar por una de las clases de valores extremos y tiene la ventaja de no requerir un valor umbral para su estimación, como los cuatro expuestos por AghaKouchak, Sellars y Sorooshian (2013).

## Cóputas de Arquímedes trivariadas

### Cóputas de Arquímedes simétricas

Chen y Guo (2019) indican para variables aleatorias multivariadas mayores de dos ( $d \geq 3$ ) y correlacionadas, que la familia de Cópulas de Arquímedes se dividen en simétricas y asimétricas. Las primeras son fáciles de construir y tienen un único parámetro de asociación ( $\theta$ ), lo cual obliga a que todos los pares de variables muestren la misma estructura y grado de dependencia (Zhang & Singh, 2019).

Las Cópulas de Arquímedes *multivariadas simétricas* ( $d \geq 3$ ) más comunes son cuatro, para las cuales se indica su rango de  $\theta$  y sus funciones generadoras  $\varphi(s)$ , y su primera y segunda derivadas  $\varphi'(s)$ ,

$\varphi''(s)$ , donde  $s$  es una variable aleatoria en el intervalo de 0 a 1 (Grimaldi & Serinaldi, 2006a; Xu *et al.*, 2016; Chen & Guo, 2019; Zhang & Singh, 2019).

**1. Cópula multivariada de Clayton**, el rango de  $\theta$  es  $(0, +\infty)$  y el límite de  $\theta = 0$  corresponde a la condición de independencia en  $u_k$ :

$$C(u_1, u_2, \dots, u_d) = \left( \sum_{k=1}^d u_k^{-\theta} - d + 1 \right)^{-1/\theta} \quad (20)$$

$$\varphi(s) = \frac{1}{\theta} (s^{-\theta} - 1) \quad (21)$$

$$\varphi'(s) = -s^{-1-\theta} \quad (22)$$

$$\varphi''(s) = (\theta + 1)/s^{\theta+2} \quad (23)$$

**2. Cópula multivariada de Frank**, el rango de  $\theta$  es  $(0, +\infty)$  y el valor de  $\theta = 1$  corresponde a la condición de independencia en  $u_k$ :

$$C(u_1, u_2, \dots, u_d) = -\frac{1}{\theta} \ln \left[ 1 + \frac{\prod_{k=1}^d (e^{-\theta u_k} - 1)}{(e^{-\theta} - 1)^{d-1}} \right] \quad (24)$$

$$\varphi(s) = -\ln \left( \frac{e^{-\theta s} - 1}{e^{-\theta} - 1} \right) \quad (25)$$

$$\varphi'(s) = \frac{\theta}{1-e^{\theta s}} \quad (26)$$

$$\varphi''(s) = \frac{\theta^2}{e^{\theta s} - 2 + e^{-\theta s}} \quad (27)$$

**3. Cópula multivariada de Gumbel-Hougaard.** El rango de  $\theta$  es  $(1, +\infty)$  y el límite de  $\theta = 1$  corresponde a la condición de independencia en  $u_k$ :

$$C(u_1, u_2, \dots, u_d) = \exp\left\{-\left[\sum_{k=1}^d (-\ln u_k)^\theta\right]^{1/\theta}\right\} \quad (28)$$

$$\varphi(s) = [-\ln(s)]^\theta \quad (29)$$

$$\varphi'(s) = -\frac{\theta}{s} [-\ln(s)]^{\theta-1} \quad (30)$$

$$\varphi''(s) = \frac{\theta}{s^2} \{(\theta - 1)[- \ln(s)]^{\theta-2} + [- \ln(s)]^{\theta-1}\} \quad (31)$$

**4. Cópula multivariada de Joe.** El rango de  $\theta$  es  $[1, +\infty)$  y el límite de  $\theta = 1$  corresponde a la condición de independencia en  $u_k$ :

$$C(u_1, u_2, \dots, u_d) = 1 - \left[1 - \prod_{k=1}^d (1 - (1 - u_k)^\theta)\right]^{1/\theta} \quad (32)$$

Las ecuaciones de su función generadora y su primera derivada ya fueron expuestas como ecuaciones (11) y (12). La ecuación de la segunda derivada de  $\varphi(s)$  es la siguiente:

$$\varphi''(s) = \frac{\theta(\theta-1)(1-s)^{\theta-2} + \theta(1-s)^{2\theta-2}}{[(1-s)^{\theta}-1]^2} \quad (33)$$

La aplicación de las *FC* de Arquímedes trivariadas simétricas se justifica con fines de comparación, pues las asimétricas deberán conducir a mejores ajustes.

## Cóputas de Arquímedes asimétricas

Para modelar diferentes estructuras de dependencia en variables aleatorias multivariadas, Chen y Guo (2019) recurren al enfoque de Grimaldi y Serinaldi (2006b) de aplicar Cóputas de Arquímedes *anidadas* (*nested*, en inglés). Con tal enfoque, las Cóputas de Arquímedes trivariadas *asimétricas* de dos parámetros de asociación ( $\theta_1$  y  $\theta_2$ ) más comunes tienen la fórmula general:  $C(u, v, w) = C_{\theta_1}(w, C_{\theta_2}(u, v))$  y son las cuatro siguientes (Joe, 1993; Grimaldi & Serinaldi, 2006b; Xu *et al.*, 2016; Zhang & Singh, 2019; Chen & Guo, 2019):

**1. Cóputa asimétrica de Clayton trivariada**, con  $\theta_2 \geq \theta_1 \in [0, \infty)$  y  $\tau_{12}, \tau_{13}, \tau_{23} \in [0, 1]$  para tres variables aleatorias con dependencia positiva:

$$C(u, v, w) = \left[ (u^{-\theta_2} + v^{-\theta_2} - 1)^{-\theta_1/\theta_2} + w^{-\theta_1} - 1 \right]^{-1/\theta_1} \quad (34)$$

**2. Cópula asimétrica de Frank trivariada,** con rango de parámetros de asociación idénticos a los anteriores:

$$C(u, v, w) = -\frac{1}{\theta_1} \ln \left\{ 1 - F_1^{-1} \left( 1 - \left[ 1 - F_2^{-1} (1 - e^{-\theta_2 u}) (1 - e^{-\theta_2 v}) \right]^{\theta_1/\theta_2} (1 - e^{-\theta_1 w}) \right) \right\} \quad (35)$$

siendo  $F_1 = 1 - e^{-\theta_1}$  y  $F_2 = 1 - e^{-\theta_2}$ .

**3. Cópula asimétrica de Gumbel-Hougaard trivariada,** con  $\theta_2 \geq \theta_1 \in [1, \infty)$  y  $\tau_{12}, \tau_{13}, \tau_{23} \in [0, 1]$  para tres variables aleatorias con dependencia positiva:

$$C(u, v, w) = \exp \left\{ - \left[ \left( (-\ln u)^{\theta_2} + (-\ln v)^{\theta_2} \right)^{\theta_1/\theta_2} + (-\ln w)^{\theta_1} \right]^{1/\theta_1} \right\} \quad (36)$$

**4. Cópula asimétrica de Joe trivariada,** con rango de parámetros de asociación idénticos a los anteriores:

$$C(u, v, w) = 1 - \left\{ \left[ (1-u)^{\theta_2} (1 - (1-v)^{\theta_2}) + (1-v)^{\theta_2} \right]^{\theta_1/\theta_2} (1 - (1-w)^{\theta_1}) + (1-w)^{\theta_1} \right\}^{1/\theta_1} \quad (37)$$

## Probabilidades empíricas trivariadas

Las probabilidades de no excedencia empíricas univariadas y trivariadas se estimaron con base en la fórmula de Gringorten, que ha sido sugerida por diversos autores en los análisis de frecuencias bivariados, y por Zhang y Singh (2007) para los trivariados. Su expresión es:

$$p = \frac{i-0.44}{n+0.12} \quad (38)$$

siendo  $i$  el número del dato cuando se ordenan de menor a mayor, y  $n$  su número total o de años de los registros de gasto pico, volumen y duración total anuales.

Para el caso de las probabilidades trivariadas se trabajó en el espacio tridimensional, con el gasto y el volumen en el plano  $x, y$ , y las duraciones en el eje perpendicular ( $z$ ). El proceso numérico comienza guardando los registros históricos de gasto máximo ( $Q$ ), volumen ( $V$ ) y duración ( $D$ ) anuales en archivos  $Qh, Vh$  y  $Dh$ ; además, se ordenaron en forma progresiva de magnitud en archivos  $Qo, Vo$  y  $Do$ . En seguida se procesa cada dato anual para comparar el valor histórico contra el ordenado y se contabilizan las veces que el segundo resultó menor o igual, y se designa  $NQ, NV$  y  $ND$ . Lo anterior equivale a cambiar el dato original, de cada terna de valores anuales históricos, por su número de orden o *rango*.

Después se compara cada terna histórica de rangos contra todas las demás y se contabilizan las veces en que los tres rangos (condición AND) resultan menores; tal cantidad se denomina  $NQVD$ , es decir, el número

de ocurrencias de combinaciones de  $q$ ,  $v$  y  $d$  menores en el espacio tridimensional. Por último, se aplica la fórmula de posición gráfica de Gringorten. Para el caso trivariado es la siguiente:

$$F_e(x, y, z) = P(Q \leq q, V \leq v, D \leq d) = \frac{NQVD_i^{-0.44}}{n+0.12} \quad (39)$$

## Selección de la función Cópula

Un enfoque simple de selección de la función Cópula se basa en los estadísticos del error de ajuste, al comparar las probabilidades empíricas observadas ( $W_i^o$ ) con las teóricas calculadas ( $W_i^c$ ) con la función Cópula que se prueba. Los indicadores aplicados son el error medio estándar (*EME*), el error medio absoluto (*EMA*) y el error absoluto máximo (*EAM*); sus expresiones son (Chowdhary & Singh, 2019; Chen & Guo, 2019):

$$EME = \sqrt{\frac{1}{n} \sum_{i=1}^n (W_i^o - W_i^c)^2} \quad (40)$$

$$EMA = \frac{1}{n} \sum_{i=1}^n |W_i^o - W_i^c| \quad (41)$$

$$EAM = \max_{i=1:n} |W_i^o - W_i^c| \quad (42)$$

## Estimación del parámetro de dependencia $\theta$

El método más simple para estimar el parámetro  $\theta$  de las *FC* trivariadas simétricas (ecuaciones (20), (24), (28) y (32)) es por tanteos, buscando que los estadísticos del error de ajuste (ecuaciones (40) a (42)) sean mínimos.

## Estimación de los parámetros de dependencia $\theta_1$ y $\theta_2$

La búsqueda del valor mínimo de la Ecuación (40) o error medio estándar, para el ajuste de las *FC* trivariadas asimétricas de Clayton, Frank, Gumbel–Hougaard y de Joe, definidas por las ecuaciones (34) a (37), se realizó con base en el algoritmo Complex de múltiples variables restringidas o acotadas, para encontrar los valores óptimos de  $\theta_1$  y  $\theta_2$ , cumpliendo con la condición  $\theta_2 > \theta_1$ .

El *algoritmo Complex* es una técnica de exploración local (Box, 1965), que se guía exclusivamente por lo que encuentra a su paso; sus antecedentes, una descripción breve de su proceso operativo y su código OPTIM en lenguaje *Basic*, se pueden consultar en Campos-Aranda (2003). En Bunday (1985) se tiene otra descripción y código de este método de búsqueda.

Las designaciones principales en el código OPTIM son NX y NY, que definen el número de variables de decisión y dependientes, función de las primeras; para el caso analizado 2 ( $\theta_1, \theta_2$ ) y 1 ( $\theta_2 > \theta_1$ ).

Una ventaja importante del código OPTIM radica en permitir un fácil acceso de los límites (L = lower, U = upper), nombres y valores iniciales

de las variables, en la subrutina citada, por medio de las designaciones siguientes:  $XL(I)$ ,  $XU(I)$ ,  $XN\$(I)$ ,  $X(I)$ ,  $YL(J)$ ,  $YU(J)$ ,  $YN\$(J)$  y  $Y(J)$ . Para el caso estudiado,  $I$  varía de 1 a 2 y  $J = 1$ . En todas las variables de decisión se usó como límite inferior 0.001 y como límite superior 5 y 10, para  $\theta_1$  y  $\theta_2$ , en la  $FC$  de Clayton y de Gumbel-Hougaard; y 10 y 20 en las  $FC$  de Frank y Joe. La única variable dependiente se definió por el cociente de  $\theta_2$  entre  $\theta_1$ , con límite inferior de 1 y superior de 5; valor que fue adoptado de manera arbitraria.

La función objetivo se denomina FO en el código OPTIM y se define al final del programa, corresponde lógicamente a la Ecuación (40), con nombre  $FO\$="EME"$ , de error medio estándar. Para los criterios de convergencia de las desviaciones absoluta y relativa de la FO se utilizaron los valores siguientes: 0.0002 y 0.00001.

## Ratificación de la función Cópula seleccionada

Esta es la etapa más importante del proceso de aplicación práctica de las  $FC$ , pues en ella se verifica que tal modelo reproduzca fielmente las probabilidades conjuntas observadas (Ecuación (39)). Yue (2000) indica la forma simple y práctica de representar las probabilidades conjuntas empíricas y teóricas, la cual consiste en llevar al eje de las abscisas la primera y al eje de las ordenadas la segunda; lógicamente, en tal gráfica, cada pareja de datos define un punto que coincide o se aleja de la recta a 45°.

Yue y Rasmussen (2002) aplican el test de Kolmogorov-Smirnov con un nivel de significancia ( $\alpha$ ) del 5 %, para aceptar o rechazar la *diferencia máxima absoluta (dma)* entre las probabilidades conjuntas

empíricas y teóricas. Para evaluar la estadística ( $D_n$ ) del test se utilizó la ecuación que exponen Meylan *et al.* (2012) para  $\alpha = 5\%$  y que es:

$$D_n = \frac{1.358}{\sqrt{n}} \quad (43)$$

$n$  es el número de datos. Si la  $d_{ma}$  es menor que  $D_n$ , se ratifica la  $FC$  adoptada.

## Periodos de retorno trivariados

### Tipos OR y AND

El primer *periodo de retorno trivariado* del evento ( $X, Y, Z$ ) se define bajo la condición OR, lo cual indica que los límites  $x, y$  o  $z$ , o los tres, *pueden* ser excedidos, y entonces la ecuación clásica del periodo de retorno o inverso de la probabilidad de excedencia será (Zhang & Singh, 2019):

$$T_{XYZ} = \frac{1}{P(X>x \text{ or } Y>y \text{ or } Z>z)} = \frac{1}{1-F_{XYZ}(x,y,z)} = \frac{1}{1-C[F_X(x),F_Y(y),F_Z(z)]} \quad (44)$$

en la cual  $C[F_X(x), F_Y(y), F_Z(z)] = C(u, v, w)$  es la  $FC$  seleccionada o probada.

El segundo *periodo de retorno trivariado* del evento ( $X, Y, Z$ ) está asociado con el caso en que los tres límites *son* excedidos ( $X > x, Y > y,$

$Z > z$ ), o condición AND, su ecuación es la (45) siguiente (Zhang & Singh, 2019):

$$T'_{XYZ} = \frac{1}{P(X>x \text{ and } Y>y \text{ and } Z>z)} = \frac{1}{F'_{XYZ}(x,y,z)} = \frac{1}{1-u-v-w+C(u,v)+C(u,w)+C(v,w)-C(u,v,w)} \quad (45)$$

## Secundario o de tipo Kendall

Salvadori y De Michele (2004) introducen con detalle el concepto del *periodo de retorno bivariado secundario* ( $\zeta$ ), designado así para enfatizar que el periodo de retorno conjunto  $T_{XY}$  es el primario, de donde procede, al utilizar las *isolíneas* que define la *FC* aplicada, cuya expresión es:

$$L_s = [(u, v) \in \mathbb{I}^2: C(u, v) = s] \quad (46)$$

siendo  $s$  la variable aleatoria unitaria  $0 < s \leq 1$  y  $C$  la *FC* probada. Después se define una región  $B_C(s)$  en el espacio unitario ( $\mathbb{I}^2$ ) sobre la *isolínea*, debajo de ella y a la izquierda, que será:

$$B_C(s) = \{(u, v) \in \mathbb{I}^2: C(u, v) \leq s\} \quad (47)$$

En las *FC* de la clase de Arquímedes, la distribución univariada de Kendall, designada  $K_C(s)$ , proporciona una medida de los eventos dentro de la  $B_C(s)$ ; su ecuación es (Salvadori & De Michele, 2004; Salvadori & De Michele, 2007; Salvadori *et al.*, 2007; Gräler *et al.*, 2013):

$$K_C(s) = s - \frac{\varphi(s)}{\varphi'(s)} \quad (48)$$

en la cual  $\varphi(s)$  es el generador de la FC y  $\varphi'(s)$  su derivada. Por último, el periodo de retorno secundario ( $\zeta$ ) de los eventos fuera de  $B_C(s)$  es:

$$\zeta = \frac{1}{1-K_C(s)} \quad (49)$$

cuyo denominador es la probabilidad de excedencia (*survival function*), la cual corresponde a eventos probablemente destructivos o peligrosos.

La distribución paramétrica de Kendall (Ecuación (49)), para las Cópulas de Arquímedes *trivariadas* simétricas es la siguiente (Barbe, Genest, Ghoudi, & Rémillard, 1996; Grimaldi & Serinaldi, 2006a; Zhang & Singh, 2019):

$$K_C(s) = P[C(u, v, w) \leq s] = s - \frac{\varphi(s)}{\varphi'(s)} - \frac{\varphi^2(s) \cdot \varphi''(s)}{2[\varphi'(s)]^3} \quad (50)$$

Sustituyendo las ecuaciones (29) a (31) en la (50), se obtiene una expresión para la distribución de Kendall de la FC de Gumbel-Hougaard trivariada simétrica, que será utilizada posteriormente.

Gräler *et al.* (2013) extienden la desigualdad  $T_{XYZ} \leq T'_{XYZ}$ , la cual indica que el periodo de retorno tipo OR es menor que el tipo AND e

indican que  $T_{KEN}$  está intermedio entre los dos citados.  $T_{KEN}$  se obtiene con la Ecuación (49). Entonces se tiene:

$$T_{XYZ} \leq T_{KEN} \leq T'_{XYZ} \quad (51)$$

Salvadori, De Michele y Durante (2011) destacan que la estimación de los periodos de retorno y sus eventos de diseño en los análisis de frecuencias *multivariados* son un problema difícil. Para resolverlo establecen un marco teórico basado en las *FC* y en la distribución de Kendall, el cual aplican mediante simulación numérica.

## De tipo condicional

Respecto a los *periodos de retorno trivariados condicionales*, Zhang y Singh (2007), en su apéndice, exponen un resumen de los que consideran de aplicación práctica. Zhang y Singh (2019) presentan una relación exhaustiva de los diferentes tipos; todos ellos estudiados por Serinaldi (2015). Ma *et al.* (2013) proponen la probabilidad condicional siguiente:

$$P(Q \leq q, V \leq v | D \geq d) = \frac{C(u,v) - C(u,v,w)}{1-w} \quad (52)$$

Grimaldi y Serinaldi (2006b) contrastan resultados de la Ecuación (52) con las Cópulas de Frank simétricas y asimétricas para los valores máximos del volumen y la duración ( $V, D$ ) dado que han ocurrido magnitudes del gasto máximo ( $Q$ ) con diversos periodos de retorno

univariados. La *FC* de Frank asimétrica proporciona un mejor ajuste, y valores más grandes de *V* y *D*.

## Datos por procesar

### Registro conjunto de *Q* y *V*

Este registro se integró con los 55 datos de gasto máximo (*Q*) y volumen escurrido (*V*) durante las crecientes anuales registradas en la estación hidrométrica La Cuña, del río Verde de la Región Hidrológica No. 12-3 (Río Santiago), México; con un área de cuenca de 19 097 km<sup>2</sup>. Tal registro fue expuesto por Gómez, Aparicio y Patiño (2010), y por Campos-Aranda (2024), y se reproduce en la Tabla 1.

**Tabla 1.** Registros conjuntos de gasto pico (*Q*) y volumen (*V*) de las crecientes anuales registradas en la estación hidrométrica La Cuña, México.

Año	<i>Q</i> (m <sup>3</sup> /s)	<i>V</i> (Mm <sup>3</sup> )
1947	784.0	146.80
1948	736.8	155.12
1949	510.0	111.40
1950	461.0	94.06
1951	411.0	111.55
1952	326.0	70.82
1953	349.8	144.75

Año	Q (m <sup>3</sup> /s)	V (Mm <sup>3</sup> )
1954	130.4	23.22
1955	690.0	203.31
1956	266.0	106.76
1957	199.0	45.92
1958	690.0	188.71
1959	340.6	47.91
1960	249.6	91.58
1961	350.0	130.68
1962	317.0	51.27
1963	732.6	127.90
1964	265.1	82.75
1965	743.6	295.34
1966	463.9	202.90
1967	1474.9	598.38
1968	323.0	118.25
1969	160.4	32.22
1970	763.8	187.75
1971	578.0	166.61
1972	191.8	26.39
1973	2440.0	920.30
1974	238.4	66.66
1975	622.1	249.07
1976	1374.0	527.96
1977	439.7	111.77

<b>Año</b>	<b>Q (m<sup>3</sup>/s)</b>	<b>V (Mm<sup>3</sup>)</b>
1978	280.2	66.23
1979	267.2	45.80
1980	287.3	99.60
1981	280.7	28.70
1982	156.5	35.37
1986	698.2	193.51
1987	184.7	55.39
1988	595.2	242.21
1989	110.2	42.49
1990	523.9	248.07
1991	1 636.3	443.30
1992	1 168.0	172.49
1993	295.0	96.50
1994	212.8	53.55
1995	367.4	114.61
1996	144.6	57.43
1997	78.4	16.55
1998	261.9	66.17
1999	196.3	41.15
2000	46.8	18.62
2001	313.8	75.78
2002	319.6	153.79
2003	621.1	326.28
2004	824.5	384.45

Año	Q (m <sup>3</sup> /s)	V (Mm <sup>3</sup> )
media	499.87	154.84
<i>MED</i>	340.60	111.40
<i>D.E.</i>	432.23	162.58
<i>Cs</i>	2.403	2.699
<i>Ck</i>	10.364	12.105

### Estimación de la duración total (*D*)

Para la estimación o asignación de la tercera variable aleatoria, relativa a la duración total (*D*) de cada creciente anual en horas, se recurrió a la representación de los hidrogramas con la distribución de probabilidades Gamma de dos parámetros de ajuste, que fue propuesta por Ponce (1989), y aplicada y descrita por Aldama-Rodríguez y Ramírez-Orozco (1998).

Tal método comienza por estimar un tiempo al pico ( $t_p$ ) anual con la expresión siguiente, la cual procede de los hidrogramas unitarios adimensionales definidos por Snider (1972), también citada por Aldama, Ramírez, Aparicio, Mejía-Zermeño y Ortega-Gil (2006):

$$t_p = ct \frac{3V}{4Q} \quad (53)$$

Cuando el volumen (*V*) de la creciente anual se expresa en millones de m<sup>3</sup> (Mm<sup>3</sup>) y su gasto pico (*Q*) en m<sup>3</sup>/s, el coeficiente de transformación

( $ct$ ) de la ecuación anterior es 208.3333 con el tiempo al pico ( $t_p$ ) en horas.

Ponce (1989) expone la ecuación analítica de los hidrogramas Gamma, función de los tiempos al pico ( $t_p$ ), y al centroide de tal hidrograma ( $t_g$ ):

$$Q = Q_p \left( \frac{t}{t_p} \right)^m \exp \left[ \frac{t_p - t}{t_g - t_p} \right] \text{ para } t > 0 \quad (54)$$

Donde  $Q$  es el gasto;  $Q_p$ , el gasto pico;  $t_p$ , el tiempo al pico;  $t_g$ , el tiempo al centroide del hidrograma ( $t_g > t_p$ ), y el exponente  $m = t_p / (t_g - t_p)$ . Los gastos tienen unidades de  $\text{m}^3/\text{s}$ , los tiempos están en segundos y los volúmenes en  $\text{m}^3$ .

Integrando la ecuación anterior, se obtiene una expresión para el volumen ( $V$ ), la cual se puede resolver por tanteos para estimar el  $t_g$ , ya que el volumen es dato (Aldama-Rodríguez & Ramírez-Orozco, 1998):

$$V = \int_0^{\infty} Q(t) \cdot dt = Q_p \cdot \exp(m) \left( \frac{1}{m} \right)^m (t_g - t_p) \cdot \Gamma(1 + m) \quad (55)$$

Para la estimación de la función Gamma  $\Gamma(\omega)$  se utilizó la fórmula de Stirling (Davis, 1972):

$$\Gamma(\omega) \cong e^{-\omega} \cdot \omega^{\omega - \frac{1}{2}} \cdot (2\pi)^{1/2} \cdot F1 \quad (56)$$

Siendo:

$$F1 = \left( 1 + \frac{1}{12 \cdot \omega} + \frac{1}{288 \cdot \omega^2} - \frac{139}{51840 \cdot \omega^3} - \frac{571}{2488320 \cdot \omega^4} + \dots \right) \quad (57)$$

Teniendo como datos  $Q$  y  $V$  de la Tabla 1, se estima el tiempo al pico ( $t_p$ ) con la Ecuación (53) y después por tanteos el  $t_g$  con la Ecuación (55). Por último, se realizan las estimaciones del tiempo final ( $t = t_f$ ), también por tanteos, hasta el 0.1% del gasto pico ( $Q_p=Q$ ) con la Ecuación (54). En la Tabla 2 se exponen los resultados obtenidos para los datos de la Tabla 1.

**Tabla 2.** Tiempos al pico ( $t_p$ ), al centroide ( $t_g$ ) y al final ( $t_f$ ) en los hidrogramas Gamma que representan a las crecientes anuales de la estación hidrométrica La Cuña, México.

Año	$t_p$ horas	$t_g$ horas	$t_f = D$ horas	$Q_f$ m <sup>3</sup> /s
1947	39.009	49.56	169.0	0.79
1948	43.861	55.73	190.2	0.74
1949	45.507	57.82	197.4	0.51
1950	42.507	54.01	184.5	0.46
1951	56.544	71.84	245.0	0.41
1952	45.258	57.50	196.0	0.33
1953	86.210	109.53	374.0	0.35
1954	37.097	47.13	161.3	0.13
1955	61.386	77.99	266.2	0.69
1956	83.615	106.23	362.5	0.27
1957	48.074	61.08	207.7	0.20
1958	56.978	72.39	247.1	0.69
1959	29.305	37.23	127.2	0.34
1960	76.439	97.12	331.8	0.25

<b>Año</b>	<b><math>t_p</math> horas</b>	<b><math>t_g</math> horas</b>	<b><math>t_f = D</math> horas</b>	<b><math>Q_f</math> m<sup>3</sup>/s</b>
1961	77.786	98.83	337.6	0.35
1962	33.695	42.81	146.2	0.32
1963	36.372	46.21	157.8	0.73
1964	65.030	83.03	286.4	0.26
1965	82.745	105.13	359.0	0.74
1966	91.121	115.77	395.6	0.46
1967	84.523	107.39	368.1	1.40
1968	76.271	96.90	331.3	0.32
1969	41.849	53.17	181.9	0.16
1970	51.211	65.06	222.1	0.76
1971	60.053	76.30	260.4	0.58
1972	28.665	36.42	124.6	0.19
1973	78.578	99.83	341.1	2.40
1974	58.253	74.01	252.8	0.24
1975	83.410	105.97	361.9	0.62
1976	80.052	101.71	347.2	1.37
1977	52.957	67.28	229.7	0.44
1978	49.243	62.56	213.7	0.28
1979	35.710	45.37	154.9	0.27
1980	72.224	91.76	313.3	0.29
1981	21.301	27.06	92.4	0.28
1982	47.085	59.82	204.2	0.16
1986	57.741	73.36	250.4	0.70
1987	62.477	79.38	272.0	0.18

Año	$t_p$ horas	$t_g$ horas	$t_f = D$ horas	$Q_f$ m <sup>3</sup> /s
1988	84.779	107.71	368.0	0.59
1989	80.327	102.06	281.1	1.10
1990	98.647	125.33	428.2	0.52
1991	56.441	71.71	245.2	1.60
1992	30.767	39.09	133.4	1.17
1993	68.150	86.58	296.2	0.29
1994	52.426	66.61	227.9	0.21
1995	64.989	82.57	281.9	0.37
1996	82.743	105.12	360.6	0.14
1997	43.979	55.88	189.5	0.08
1998	52.636	66.87	228.6	0.26
1999	43.673	55.49	190.2	0.19
2000	82.888	105.30	360.4	0.05
2001	50.324	63.94	218.7	0.31
2002	100.249	127.37	435.1	0.32
2003	109.443	139.05	474.8	0.62
2004	97.142	123.42	421.5	0.82
MAX	109.443	139.05	474.8	2.40

## Verificación de aleatoriedad de los registros

El Test de Wald-Wolfowitz es una prueba no paramétrica que ha sido utilizada por Bobée y Ashkar (1991), Rao y Hamed (2000), y Meylan *et al.* (2012), para verificar *independencia* y *estacionariedad* en registros de gastos máximos anuales ( $X_i$ ). Por lo anterior, se propuso aplicar esta

prueba a los registros conjuntos de  $Q$ ,  $V$  y  $D$  que deben ser muestras *aleatorias*.

## Errores de ajuste

El primer criterio aplicado para la selección de la mejor FDP a unos datos o serie disponible fueron los llamados *errores de ajuste* (Kite, 1977; Willmott & Matsuura, 2005; Chai & Draxler, 2014). Este criterio y el descrito para evitar probabilidades negativas permitirán adoptar una FDP adecuada entre los diversos modelos probados.

Cambiando en las ecuaciones (40) y (41) las probabilidades observadas por los datos ordenados de la serie analizada ( $x_i$ ,  $y_i$  o  $z_i$ ) y las probabilidades calculadas por los valores estimados con la solución inversa de la FDP que se prueba o contrasta, se obtienen el error estándar de ajuste (*EEA*) y el error absoluto medio (*EAM*). Los valores que se estiman ( $\hat{x}_i$ ,  $\hat{y}_i$  y  $\hat{z}_i$ ) se buscan para la misma probabilidad de no excedencia, asignada a los datos con la fórmula empírica de Gringorten (Ecuación (38)).

Zhang y Singh (2007) también aplican otros dos estadísticos de calidad del ajuste: el primero, cuando las FDP se ajustan mediante el método de máxima verosimilitud emplean el criterio de información de Akaike (AIC, por sus siglas en inglés); el segundo fue el test de Kolmogorov-Smirnov, expuesto en la Ecuación (43).

## Resultados y su discusión

### Distribuciones marginales adoptadas

#### Verificación de la aleatoriedad de los registros

Los registros conjuntos de gasto máximo y volumen anual de la Tabla 1 y del tiempo final ( $t_f = D$ , en horas) de las columnas 4 y 9 de la Tabla 2 resultaron series *aleatorias*, cuyo estadístico ( $U$ ) del Test de Wald-Wolfowitz resultó de 0.284, 0.213 y 1.139, respectivamente.

#### FDP de los gastos máximos y volúmenes anuales

Campos-Aranda (2024) procesó los registros de gasto máximo y volumen de la Tabla 1, con las tres FDP más convenientes de acuerdo con el diagrama de cocientes de momentos L de Hosking y Wallis (1997), y las distribuciones de Kappa y Wakeby. Todas las FDP aplicadas se ajustaron con el método de momentos L, excepto la Log-Pearson tipo III (LP3), que se ajustó por momentos en los dominios logarítmico y real.

Campos-Aranda (2024) encontró para los gastos máximos anuales que las dos mejores distribuciones, la Log-Normal y la General de Valores Extremos (GVE), reportan los errores de ajuste más bajos y similares; pero sus predicciones son elevadas en los cuatro periodos de retorno extremos ( $Tr > 500$  años). Por otra parte, los errores de ajuste de la distribución Kappa son casi idénticos a los de la Log-Normal, pero con predicciones intermedias, y por ello fue la adoptada.

Los parámetros de ubicación ( $u_1$ ), escala ( $a_1$ ) y forma ( $k_1$  y  $h_1$ ) de la distribución Kappa seleccionada para los gastos máximos son 258.7462, 228.381,  $-0.2685394$  y  $0.2888472$ , cuya expresión es:

$$F(x) = \left\{ 1 - h_1 \left[ 1 - \frac{k_1(x-u_1)}{a_1} \right]^{1/k_1} \right\}^{1/h_1} \quad (58)$$

Para los volúmenes anuales de la Tabla 1, Campos-Aranda (2024) encontró que la mejor distribución, la LP3, conduce a los errores de ajuste más bajos; pero sus predicciones son bastante menores que las de la GVE. En cambio, la distribución Kappa conduce a errores de ajuste bajos y sus predicciones son intermedias, por ello fue adoptada.

Los parámetros de ubicación ( $u_2$ ), escala ( $a_2$ ) y forma ( $k_2$  y  $h_2$ ) de la distribución Kappa adoptada para los volúmenes anuales son 60.39858, 78.31082,  $-0.3155518$  y  $0.4287021$ , cuya expresión es:

$$F(y) = \left\{ 1 - h_2 \left[ 1 - \frac{k_2(y-u_2)}{a_2} \right]^{1/k_2} \right\}^{1/h_2} \quad (59)$$

## FDP de las duraciones totales

Aplicando el criterio de la distancia absoluta ponderada, expuesto por Campos-Aranda (2024), al registro de duraciones totales ( $D$ ) de las crecientes anuales de la Tabla 2, se obtuvo que las tres mejores FDP fueron la Pareto Generalizada (PAG), la LP3 y la GVE. Además, en la Tabla

3 se muestran los resultados de los modelos Log-Normal (LN3), Kappa y Wakeby.

**Tabla 3.** Errores de ajuste y predicciones ( $m^3/s$ ) de las tres mejores FDP y tres de aplicación generalizada en el registro de duraciones de las crecientes anuales ( $D$ ) de la estación hidrométrica La Cuña, México.

FDP	EEA ( $m^3/s$ )	EAM ( $m^3/s$ )	Periodos de retorno, en años					
			50	100	500	1 000	5 000	10 000
PAG	10.7	8.7	441	448	456	457	459	459
LP3	13.1	9.4	483	517	587	614	671	694
GVE	12.4	9.8	477	507	563	584	622	636
LN3	12.9	10.2	479	513	587	616	682	709
Kappa	9.4	7.8	453	466	482	486	491	492
Wakeby	10.6	8.3	448	458	468	470	473	473

Las FDP de la Tabla 3 definen para una probabilidad de no excedencia ( $p$ ) del 1 % los valores siguientes: 123.1, 93.2, 76.4, 74.4, 110.5 y 126.9 horas, con lo cual el valor más bajo del registro de duraciones totales  $D = 92.4$  horas de la Tabla 2 conduciría a un valor negativo de la probabilidad, excepto con los modelos GVE y LN3. Se adopta la distribución GVE por tener errores de ajuste menores que la función LN3.

Los parámetros de ubicación ( $u_3$ ), escala ( $a_3$ ) y forma ( $k_3$ ) de la distribución GVE adoptada son 227.791, 86.95609, 0.1676378, cuya ecuación es:

$$F(z) = \exp\left\{-\left[1 - k_3 \frac{(z-u_3)}{a_3}\right]^{1/k_3}\right\} \quad (60)$$

El valor positivo del  $k_3$  indica una distribución Weibull con límite superior. En la Tabla 4 se exponen las predicciones obtenidas con las FDP adoptadas para los registros de  $Q$ ,  $V$  y  $D$  de la Tabla 1 y Tabla 2.

**Tabla 4.** Predicciones obtenidas con las FDP adoptadas en los registros de  $Q$ ,  $V$  y  $D$  de las crecientes anuales registradas en la estación hidrométrica La Cuña, México.

Registro	FDP	Periodos de retorno bajos, en años					Periodos de retorno altos, en años				
		2	5	10	25	50	100	500	1 000	5 000	10 000
Gasto pico (m <sup>3</sup> /s)	Kappa	372	692	971	1 419	1 835	2 335	3 920	4 844	7 784	9 489
Volumen (Mm <sup>3</sup> )	Kappa	104	217	321	495	663	873	1 576	2 007	3 459	4 350
Duración (horas)	GVE	259	343	391	443	477	507	563	584	622	636

### Estimación de las probabilidades empíricas

Estas probabilidades se estiman con base en la fórmula de Gringorten (Ecuación (38)) aplicada en el espacio tridimensional, según proceso descrito. Los resultados para la Ecuación (39) se exponen en la Tabla 5.

**Tabla 5.** Duración ( $D$ ) estimada con el hidrograma tipo Gamma para las crecientes anuales de la estación hidrométrica La Cuña, México, su número de combinaciones  $Q$ ,  $V$  y  $D$  menores en el espacio tridimensional y la probabilidad empírica observada (Ecuación (39)).

No.	$D$ (horas)	$NQVD_i (W_i^o)$
1	169.0	8(0.1372)
2	190.2	12(0.2097)
3	197.4	12(0.2097)
4	184.5	8(0.1372)
5	245.0	17(0.3004)
6	196.0	9(0.1553)
7	374.0	28(0.5000)
8	161.3	1(0.0102)
9	266.2	24(0.4274)
10	362.5	17(0.3004)
11	207.7	7(0.1190)
12	247.1	21(0.3730)
13	127.2	3(0.0464)
14	331.8	12(0.2097)
15	337.6	25(0.4456)
16	146.2	3(0.0464)
17	157.8	6(0.1009)
18	286.4	13(0.2279)
19	359.0	37(0.6633)
20	395.6	34(0.6089)
21	368.1	47(0.8447)
22	331.3	21(0.3730)

No.	D (horas)	NQVD <sub>i</sub> (W <sub>i</sub> <sup>o</sup> )
23	181.9	2(0.0283)
24	222.1	19(0.3367)
25	260.4	22(0.3911)
26	124.6	1(0.0102)
27	341.1	41(0.7358)
28	252.8	9(0.1553)
29	361.9	34(0.6089)
30	347.2	40(0.7177)
31	229.7	17(0.3004)
32	213.7	9(0.1553)
33	154.9	2(0.0283)
34	313.3	17(0.3004)
35	92.4	1(0.0102)
36	204.2	3(0.0464)
37	250.4	22(0.3911)
38	272.0	5(0.0827)
39	368.0	35(0.6270)
40	281.1	2(0.0283)
41	428.2	36(0.6451)
42	245.2	26(0.4637)
43	133.4	4(0.0646)
44	296.2	17(0.3004)
45	227.9	8(0.1372)
46	281.9	20(0.3549)
47	360.6	5(0.0827)
48	189.5	1(0.0102)

No.	$D$ (horas)	$NQVD_i (W_i^o)$
49	228.6	9(0.1553)
50	190.2	5(0.0827)
51	360.4	1(0.0102)
52	218.7	11(0.1916)
53	435.1	25(0.4456)
54	474.8	40(0.7177)
55	421.5	47(0.8447)
$\bar{X}$	265.6	-
$MED$	250.4	-
$D.E.$	91.6	-
$Cs$	0.272	-
$Ck$	2.359	-

## Selección y ratificación de las $FC$ bivariadas

Para la aplicación de la Ecuación (45) del periodo de retorno trivariado de tipo AND se requieren las  $FC$  bivariadas:  $C(u,v)$ ,  $C(u,w)$  y  $C(v,w)$ . En la parte superior de la Tabla 6 se citan las medidas de asociación y concordancia de las tres variables bivariadas, y posteriormente se indican los estadísticos de ajuste de las  $FC$  de Clayton, Frank, Gumbel-Hougaard y Joe, definidas en las ecuaciones (2) a (12).

**Tabla 6.** Errores de ajuste de las cuatro *FC* aplicadas a las tres variables *bivariadas* indicadas de las crecientes anuales de la estación hidrométrica La Cuña, México.

Medidas de asociación y errores de ajuste	Variables bivariadas		
	Q-V	Q-D	V-D
Coeficiente de correlación ( $r_{xy}$ )	0.9302	0.1840	0.4648
Cociente tau de Kendall ( $\tau_n$ )	0.7199	0.1367	0.4168
Coeficiente rho de Spearman ( $\rho_n$ )	0.9088	0.2116	0.5742
Dependencia observada ( $\lambda_U^{CFG}$ )	0.7819	0.1835	0.4530
Cópula de Clayton			
Parámetro asociación( $\theta$ )	5.1394	0.3167	1.4296
Error medio estándar ( <i>EME</i> )	0.0245	0.0276	0.0273
Error absoluto medio ( <i>EAM</i> )	0.0190	0.0213	0.0215
Error máximo absoluto ( <i>EMA</i> )	0.0714	0.0821	0.0627
Dependencia superior ( $\lambda_U$ )	0.0000	0.0000	0.0000
Cópula de Frank			
Parámetro asociación( $\theta$ )	12.3850	1.2490	4.3970
Error medio estándar ( <i>EME</i> )	0.0227	0.0272	0.0258
Error absoluto medio ( <i>EAM</i> )	0.0169	0.0210	0.0203
Error máximo absoluto ( <i>EMA</i> )	0.0582	0.0772	0.0593
Dependencia superior ( $\lambda_U$ )	0.0000	0.0000	0.0000
Cópula de Gumbel-Hougaard			
Parámetro asociación( $\theta$ )	3.5697	1.1583	1.7148
Error medio estándar ( <i>EME</i> )	0.0255	0.0289	0.0297
Error absoluto medio ( <i>EAM</i> )	0.0192	0.0224	0.0223
Error máximo absoluto ( <i>EMA</i> )	0.0676	0.0741	0.0792
Dependencia superior ( $\lambda_U$ )	0.7857	0.1808	0.5019
Cópula de Joe			
Parámetro asociación( $\theta$ )	5.8866	1.2786	2.3104
Error medio estándar ( <i>EME</i> )	0.0289	0.0302	0.0340
Error absoluto medio ( <i>EAM</i> )	0.0225	0.0233	0.0264
Error máximo absoluto ( <i>EMA</i> )	0.0788	0.0733	0.0920
Dependencia superior ( $\lambda_U$ )	0.8750	0.2804	0.6501

En la Tabla 6 se observa que la *FC* de Frank conduce al mejor ajuste a los datos bivariados y la *FC* de Joe al peor ajuste. Al tomar en cuenta la dependencia observada ( $\lambda_{ij}^{CFG}$ ) de las variables bivariadas, no existe duda para seleccionar la *FC* de Gumbel-Hougaard, que prácticamente reproduce dos de ellas y queda ligeramente excedida en las variables *V-D*.

## Selección y ratificación de la *FC* trivariada

### Funciones Cópula simétricas

Los datos *conjuntos* anuales de gasto máximo (*Q*), volumen escurrido (*V*) y duración total (*D*) del hidrograma de cada creciente, tomados de la Tabla 1 y Tabla 5, se ajustaron con las *FC* trivariadas ( $d = 3$ ) de Clayton, Frank, Gumbel-Hougaard y de Joe, definidas por las ecuaciones (20), (24), (28) y (32). Tal ajuste se realizó por tanteos del valor del su parámetro de asociación ( $\theta$ ), buscando los menores errores de ajuste según las ecuaciones (40) a (42). Los resultados obtenidos se muestran en la Tabla 7.

**Tabla 7.** Indicadores estadísticos del ajuste de las funciones Cópula trivariadas *simétricas* indicadas en las crecientes anuales de la estación hidrométrica La Cuña, México.

Cópula	$\theta$	EME	EAM	No.DP	No.DN	MDP	MDN
Clayton	1.959	0.0346	0.0275	18	37	0.1029	-0.0631
Frank	5.775	0.0300	0.0222	29	26	0.0832	-0.0659
G-H	2.100	0.0351	0.0246	31	24	0.0830	-0.0917
Joe	3.250	0.0419	0.0296	33	22	0.0692	-0.1196

DP, DN = diferencias positivas y negativas.

MDP, MDN = máxima diferencia positiva y negativa.

Por otra parte, la Ecuación (43) define a  $D_n = 0.1831$  y como la diferencia máxima absoluta de la *FC* de Frank simétrica en la Tabla 7 es 0.0832, el test de Kolmogorov-Smirnov ratifica su adopción. El coeficiente de correlación ( $r_{xy}$ ) entre las probabilidades empíricas (Tabla 5) y las teóricas, estimadas con la *FC* de Frank simétrica, resultó de 0.9927; por lo tanto, se tiene un muy buen ajuste.

## Funciones Cópulas asimétricas

La aplicación de las *FC* asimétricas trivariadas, con dos parámetros de asociación ( $\theta_1, \theta_2$ ), a los datos de las Tabla 1 y Tabla 5, se realizó con base en el algoritmo Complex de múltiples variables acotadas. Los valores iniciales en la *FC* de Clayton fueron 1.5 y 2.0; en la de Frank 3 y 7; en la de Gumbel-Hougaard 1.5 y 3.5, y en la de Joe 1.5 y 4.5. Los valores

óptimos encontrados de  $\theta_1$  y  $\theta_2$  y sus indicadores de ajuste se han concentrado en la Tabla 8.

**Tabla 8.** Indicadores estadísticos del ajuste de las funciones Cópula trivariadas *asimétricas* indicadas en las crecientes anuales de la estación hidrométrica La Cuña, México.

Cópula	No. EFO	$\theta_1$	$\theta_2$	EME	EAM	No. DP	No. DN	MDP	MDN
Clayton	37	0.9698	4.8490	0.0281	0.0208	24	31	0.0781	-0.0585
Frank	80	3.2736	11.1166	0.0255	0.0194	32	23	0.0657	-0.0567
G-H	39	1.3805	6.9013	0.0304	0.0226	32	23	0.0582	-0.0799
Joe	33	1.7483	8.7409	0.0351	0.0260	35	20	0.0635	-0.0953

EFO = evaluaciones de la función objetivo.

$\theta_1, \theta_2$  = parámetros de asociación de FC asimétrica.

Nuevamente, la FC de Frank define el mejor ajuste. Como ya se indicó, la Ecuación (43) define a  $D_n = 0.1831$ , y como la diferencia máxima absoluta de la FC de Frank asimétrica en la Tabla 8 es 0.0657, el test de Kolmogorov-Smirnov ratifica su adopción. El coeficiente de correlación ( $r_{xy}$ ) entre las probabilidades empíricas (Tabla 5) y las teóricas, estimadas con la FC de Frank asimétrica, resultó de 0.9949; por lo tanto, se tiene un ajuste excelente.

## Adopción de una *FC* trivariada

El resultado de la Tabla 6, de ser la *FC* de Gumbel-Hougaard la adoptada, debido a la reproducción que hace del valor de  $\lambda_{ij}^{CFG}$  para las tres bivariables *Q-V*, *Q-D* y *V-D*, orienta su selección para el caso trivariado.

Tal selección no se considera inapropiada, pues según se observa en la Tabla 7 y Tabla 8, tal *FC* de Gumbel-Hougaard muestra ajustes bastante similares a los de la *FC* de Frank simétrica y asimétrica. Lo anterior se verificó con base en los coeficientes de correlación entre las probabilidades empíricas (Tabla 5) y las teóricas trivariadas de la *FC* de Gumbel-Hougaard simétrica y asimétrica, cuyos valores fueron 0.9916 y 0.9939.

En la Tabla 9 y Tabla 10 se muestra una parte de las probabilidades de no excedencia trivariadas, empíricas observadas ( $w_i^o$ ), tomadas de la Tabla 5, y las teóricas calculadas ( $w_i^e$ ) con la *FC* de Gumbel-Hougaard simétrica y asimétrica. También se indican sombreadas las diferencias máximas positiva y negativa.

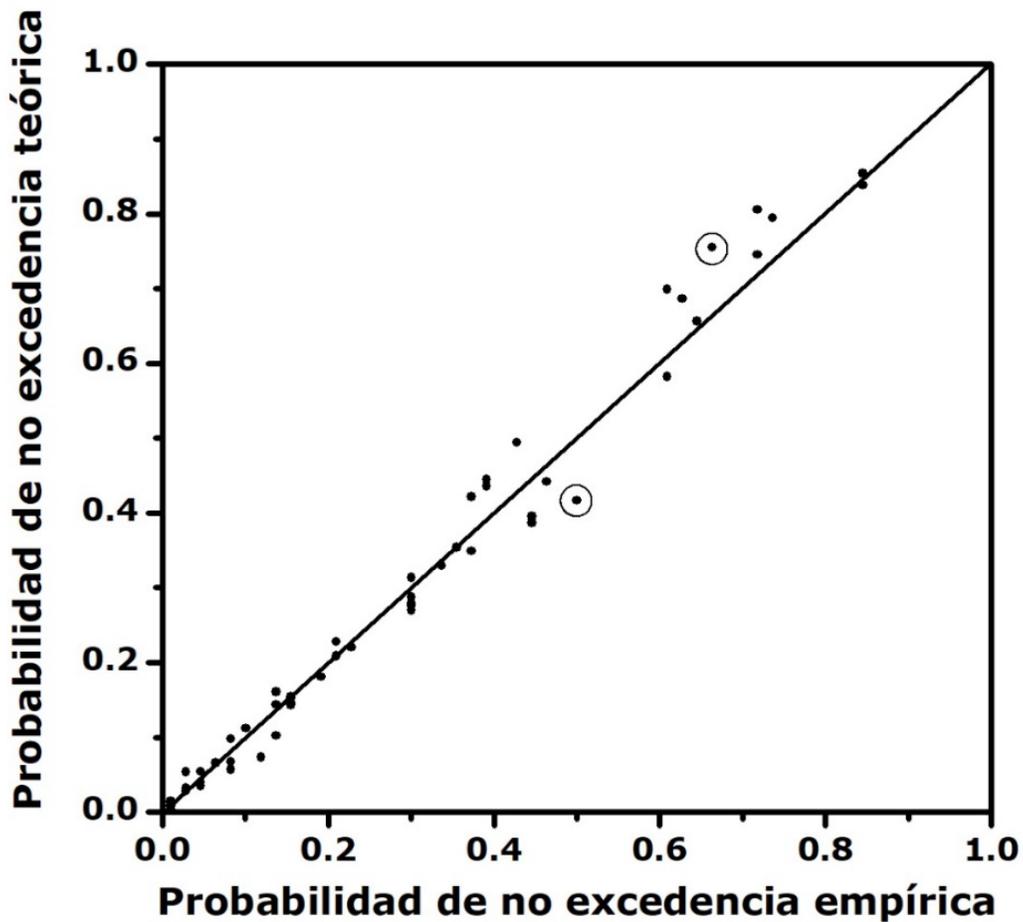
**Tabla 9.** Una parte de las probabilidades de no excedencia trivariadas y sus diferencias, calculadas con la *FC* de Gumbel-Hougaard *simétrica*, para las crecientes de anuales de la estación hidrométrica La Cuña, México.

No.	$w_i^o$	$w_i^c$	Diferencias
1	0.1372	0.1435	-0.0063
7	0.5000	0.4170	0.0830
10	0.3004	0.2701	0.0303
15	0.4456	0.3955	0.0501
19	0.6633	0.7550	-0.0917
25	0.3911	0.4446	-0.0535
30	0.7177	0.8057	-0.0880
35	0.0102	0.0087	0.0015
40	0.0283	0.0281	0.0002
45	0.1372	0.1021	0.0351
50	0.0827	0.0570	0.0257
55	0.8447	0.8380	0.0067

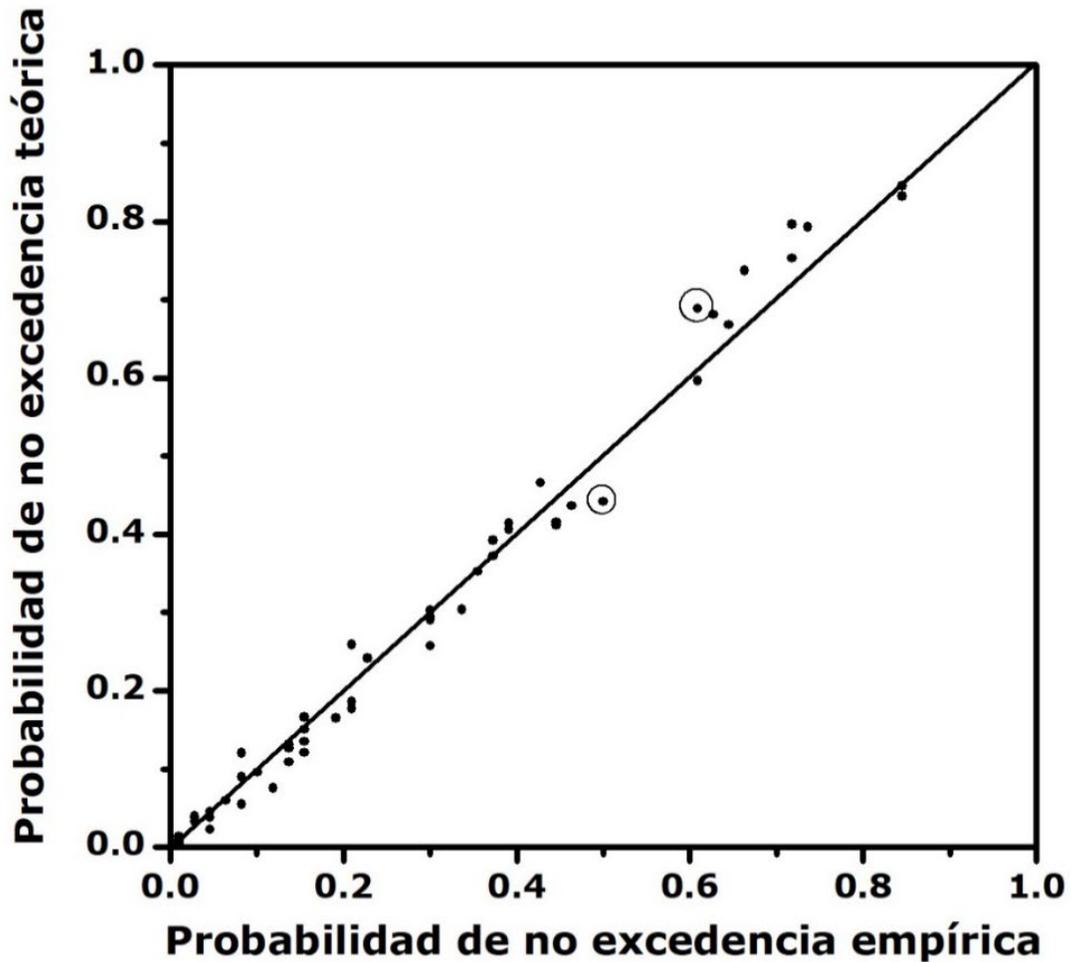
**Tabla 10.** Una parte de las probabilidades de no excedencia trivariadas y sus diferencias, calculadas con la *FC* de Gumbel-Hougaard *asimétrica*, para las crecientes de anuales de la estación hidrométrica La Cuña, México.

No.	$w_i^o$	$w_i^c$	Diferencias
1	0.1372	0.1264	0.0108
7	0.5000	0.4418	0.0582
10	0.3004	0.3018	-0.0014
15	0.4456	0.4154	0.0302
20	0.6089	0.5961	0.0128
25	0.3911	0.4142	-0.0231
29	0.6089	0.6888	-0.0799
35	0.0102	0.0044	0.0058
40	0.0283	0.0372	-0.0089
45	0.1372	0.1089	0.0283
50	0.0827	0.0549	0.0278
55	0.8447	0.8323	0.0124

Como ya se indicó, la Ecuación (43) define a  $D_n = 0.1831$ , y como las diferencias máximas absolutas de la Tabla 9 y Tabla 10 son 0.0917 y 0.0799, respectivamente, el test de Kolmogorov-Smirnov ratifica a las *FC* de Gumbel-Hougaard simétrica y asimétrica adoptadas. El contraste gráfico entre ambas probabilidades para ratificar su adopción se expone en la Figura 2 y Figura 3 para los datos completos de la Tabla 9 y Tabla 10.



**Figura 2.** Contraste gráfico de probabilidades empíricas y teóricas estimadas con la *FC* de Gumbel-Hougaard trivariada *simétrica* para las crecientes anuales registradas en la estación La Cuña, México.



**Figura 3.** Contraste gráfico de probabilidades empíricas y teóricas estimadas con la *FC* de Gumbel-Hougaard trivariada *asimétrica* para las crecientes anuales registradas en la estación La Cuña, México.

## Estimación de los periodos de retorno trivariados

Con base en las ternas de predicciones mostradas en la Tabla 4, se aplicaron las ecuaciones (44) y (45), para estimar los periodos de retorno conjuntos de tipo OR y AND, mostrados en la Tabla 11. En tales expresiones se aplican las *FC* de Gumbel-Hougaard bivariadas, definidas en la Tabla 6; la trivariada simétrica, cuyo parámetro de asociación ( $\theta$ ) está indicado en la Tabla 7; para *FC* trivariada asimétrica sus parámetros  $\theta_1$  y  $\theta_2$  se citaron en la Tabla 8. Para la estimación del periodo de retorno secundario ( $T_{KEN}$ ) se aplicó la Ecuación (50) para la *FC* Gumbel-Hougaard.

**Tabla 11.** Periodos de retorno conjuntos de tipo OR, AND y secundario estimados con las *FC* trivariadas de Gumbel-Hougaard simétrica y asimétrica para las crecientes anuales de la estación hidrométrica La Cuña, México.

<i>Tr</i> (años)	Tipo de <i>Tr</i> con la <i>FC</i> simétrica			<i>FC</i> asimétrica	
	OR	Secundario	AND	OR	AND
2	1.5	3.8	3.9	1.4	3.7
5	3.2	11.3	17.1	3.1	15.0
10	6.2	23.9	44.8	6.0	37.2
25	15.0	61.5	137.0	14.6	107.2
50	29.9	124.1	294.5	29.0	226.9
100	59.7	249.4	613.7	58.0	469.3
500	295.1	1 251.8	3 216.5	285.5	2 358.0
1 000	596.1	2 504.1	6 335.8	579.2	4 836.3
5 000	2 957.4	12 529.7	32 202.0	2 865.9	23 899.2
10 000	5 938.8	25 040.6	64 035.2	5 771.3	48 771.0

La similitud entre los periodos de retorno *trivariados* o *conjuntos* de tipo OR de la Tabla 11 es notable, resultando ligeramente inferiores los estimados con la *FC* asimétrica; por lo tanto, los más exactos. Para los de tipo AND, se tienen mayores diferencias y, lógicamente, los más confiables son los estimados con la *FC* asimétrica.

La semejanza que mostraron los periodos de retorno trivariados de tipo OR permite considerar que la *FC* de Gumbel-Hougaard trivariada *simétrica* ofrece una descripción aceptable y confiable de los datos conjuntos trivariados, y por lo tanto el periodo de retorno *secundario* se puede utilizar para obtener los eventos de diseño conjuntos, cuyas variables  $Q$ ,  $V$  y  $D$  se obtendrán con las funciones o distribuciones marginales.

## Estimación de los eventos de diseño

Suponiendo que en las cercanías de aguas abajo de la estación hidrométrica La Cuña y sobre el Río Verde se van a construir diques para protección de planicies de inundación con fines agrícolas e industriales y un puente para cruzarlo, entonces se requiere estimar *eventos de diseño* con periodos de retorno conjuntos o trivariados de 50, 100, 500 y 1 000 años. Además, se realizará una revisión de la seguridad hidrológica del embalse en proyecto, en el sitio de la estación de aforos, con periodos de retorno conjuntos de 5 000 y 10 000 años. Por lo anterior, es necesario estimar ternas de  $Q$ ,  $V$  y  $D$  con los seis periodos de retorno conjuntos  $T_{KEN}$  citados.

Con base en la Ecuación (50) de la distribución de Kendall trivariada, establecida para la *FC* de Gumbel-Hougaard, se buscó por tanteos el periodo de retorno univariado ( $Tr$ ) y su respectiva probabilidad de no excedencia ( $s$ ), que definen un periodo de retorno *secundario* igual al conjunto o trivariado de diseño. Encontrado tal valor de la variable unitaria ( $s$ ), se obtienen con las soluciones inversas de las distribuciones marginales (ecuaciones (58) a (60)), las variables  $Q$ ,  $V$  y  $D$  respectivas. En la Tabla 12 se exponen los resultados.

**Tabla 12.** Eventos de diseño obtenidos para el periodo de retorno secundario indicado en las crecientes anuales de la estación hidrométrica La Cuña, México.

<b><math>Tr</math> de diseño secundario</b>	<b><math>Tr</math> univariado</b>		<b>Eventos de diseño</b>		
	<b>(años)</b>	<b>Prob. <math>s</math></b>	<b><math>Q</math> (<math>m^3/s</math>)</b>	<b><math>V</math> (<math>Mm^3</math>)</b>	<b><math>D</math> (horas)</b>
50	20	0.950	1 300	448	431
100	40	0.9750	1 693	605	466
500	200	0.9950	2 935	1 132	533
1 000	400	0.9975	3 657	1 456	556
5 000	1 996	0.9994989	5 952	2 542	601
10 000	3 989	0.9997493	7 288	3 208	617

## Estimación de periodos de retorno condicionales

De acuerdo con Grimaldi y Serinaldi (2006b), y Ma *et al.* (2013), una *probabilidad condicional (PC)* interesante está definida por la Ecuación (52) y para el caso estudiado su planteamiento será el siguiente:

$$P(V \leq v, D \leq d | Q \geq q) = PC_{1,2} = \frac{C(v,w) - C(u,v,w)}{1-u} \quad (61)$$

Esta expresión permite encontrar por tanteos los valores del volumen ( $V$ ) y la duración ( $D$ ), dado que ha ocurrido un gasto máximo ( $Q$ ) de un determinado periodo de retorno, el cual debe ser igualado por el *periodo de retorno condicional* siguiente, función de la  $PC_{1,2}$ :

$$Tr_{1,2} = 1/(1 - PC_{1,2}) \quad (62)$$

En la Ecuación (61),  $u$ ,  $v$  y  $w$  se calculan para los valores de  $Q$ ,  $V$  y  $D$  con las FDP marginales, según las ecuaciones (58) a (60). En la Ecuación (61),  $C(v,w)$  es la *FC* de Gumbel-Hougaard bivariada, definida en la Tabla 6 con  $\theta = 1.7148$  y  $C(u,v,w)$ ; son las *FC* de Gumbel-Hougaard trivariadas simétrica ( $\theta = 2.1$ ) y asimétrica ( $\theta_1 = 1.3805$  y  $\theta_2 = 6.9013$ ), las que definen las  $PC_{1,2}$ . Los resultados para los cuatro periodos de retorno indicados se tienen en la Tabla 13.

**Tabla 13.** Valores de parejas de  $V$  y  $D$  condicionados a la ocurrencia de  $Q$ , cuyo periodo de retorno se indica para las crecientes anuales de la estación La Cuña, México.

Variable aleatoria	Periodos de retorno del $Q$ en años.			
	10	25	50	100
$Q$ ( $m^3/s$ ) asignado	971	1 419	1 835	2 335
$V$ ( $Mm^3$ ) buscado	1 500	2 500	3 500	4 500
$D_1$ (horas) buscada	507	600	651	654
$D_2$ (horas) buscada	494	594	649	651

En la Tabla 13, los valores altos de la duración ( $D_1$ ) se estimaron con la *FC* de Gumbel-Hougaard trivariada simétrica y los bajos ( $D_2$ ) con la *FC* asimétrica.

## Conclusiones

Los análisis de frecuencias de crecientes *trivariados* de las variables gasto máximo ( $Q$ ), volumen escurrido ( $V$ ) y duración total ( $D$ ) permitirán una estimación más exacta del hidrograma de la creciente de diseño, asociada con un periodo de retorno conjunto.

En este estudio, para procesar los registros anuales conjuntos de  $Q$  y  $V$  disponibles, se asignó como tercera variable de las crecientes a  $D$ , estimada con el hidrograma de tipo Gamma, para un gasto final del 0.1 % del  $Q$ .

El uso de las funciones Cópula (*FC*) en los análisis de frecuencias multivariados permite construir la distribución multivariada a partir de las

funciones marginales. Por lo anterior, las distribuciones de probabilidad idóneas de  $Q$ ,  $V$  y  $D$  se definen con la máxima precisión posible, y pueden ser diferentes y de cualquier tipo.

La estimación del periodo de retorno trivariado de tipo AND requiere de las distribuciones bivariadas en el caso estudiado de las variables conjuntas  $Q$ - $V$ ,  $Q$ - $D$  y  $V$ - $D$ . Por lo anterior, primero se buscan  $FC$  que reproduzcan la dependencia observada ( $\lambda_U^{CFG}$ ) y muestren un buen ajuste con las variables conjuntas citadas.

En la aplicación numérica descrita, con 55 crecientes anuales registradas en la estación de aforos La Cuña, de la Región Hidrológica No. 12-3 (río Santiago), México, se emplearon las siguientes cuatro familias de  $FC$ : Clayton, Frank, Gumbel-Hougaard y Joe.

Para las ternas de datos anuales de  $Q$ ,  $V$  y  $D$ , se aplicaron  $FC$  de Arquímedes multivariadas simétricas con un parámetro de asociación ( $\theta$ ) y trivariadas asimétricas con dos parámetros de asociación ( $\theta_1$ ,  $\theta_2$ ) de las familias citadas. Finalmente, se estimaron periodos de retorno conjuntos de tipo OR, AND y Kendall. Estos últimos permiten obtener los eventos de diseño de  $Q$ ,  $V$  y  $D$  mostrados en la Tabla 12.

Los análisis de frecuencias de crecientes trivariados descritos tienen gran sencillez y no presentan complicaciones de cómputo cuando se realizan con base en las  $FC$ .

## Referencias

- 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.
- Aldama-Rodríguez, A. A., & Ramírez-Orozco, A. I. (1998). Parametrización de hidrogramas mediante interpolantes hermitianos. *Ingeniería Hidráulica en México*, 13(3), 19-28.
- Barbe, P., Genest, C., Ghoudi, K., & Rémillard, B. (1996). On Kendall's Process. *Journal of Multivariate Analysis*, 58(2), 197-229.
- 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. (2024). Análisis de frecuencias bivariado de Crecientes Anuales mediante enfoque práctico de las funciones Cópula. *Tecnología y ciencias del agua*, 15(2), 1-56. DOI: 10.24850/j-tyca-15-02-01
- 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.
- 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
- Chen, L., & Guo, S. (2019). Chapter 2. Copula theory, and Chapter 3. Copula-based Flood Frequency Analysis. *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. (1972). Chapter 6. Gamma Function and related functions. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (pp. 253-296). New York, USA: iDover Publications.

- Davis, P. J., & Polonsky, I. (1972). Chapter 25. Numerical interpolation, differentiation and integration. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (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)
- Favre, A. C., El Adlouni, S., Perreault, L., Thiémonge, N., & Bobée, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, 40(1), 1-12. DOI: 10.1029/2003WR002456
- 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* (pp. 30.1-30.10), 2<sup>nd</sup> ed. 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. DOI: 10.1061/(ASCE)0733-9429(1998)124:2(146)

- Gómez, J. F., Aparicio, M., & Patiño, C. (2010). Capítulo 6. Análisis de frecuencias bivariado para la estimación de avenidas de diseño. En: *Manual de análisis de frecuencias en hidrología* (pp. 106-127). Jiutepec, México: Instituto Mexicano de Tecnología del Agua.
- 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., & 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, UK: Cambridge University Press.
- Joe, H. (1993). Parametric families of multivariate distributions with given margins. *Journal of Multivariate Analysis*, 46(2), 262-282.
- Kite, G. W. (1977). Chapter 12. Comparison of frequency distributions. *Frequency and risk analyses in hydrology* (pp. 156-168). Fort Collins, USA: Water Resources Publications.

- Ma, M., Song, S., Ren, L., Jiang, S., & Song, J. (2013). Multivariate drought characteristics using trivariate Gaussian and Student t copulas. *Hydrological Processes*, 27(8), 1175-1190. DOI: 10.1002/hyp.8432
- 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 2. Definitions and Basic Properties. In: *An introduction to Copulas* (pp. 7-49), 2<sup>nd</sup> ed. 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 (pp. 416-434). En: *Métodos numéricos aplicados a la Ingeniería*. México, DF, México: Compañía Editorial Continental.
- Ponce, V. M. (1989). Section 2.4. Runoff. In: *Engineering hydrology. principles and practices* (pp. 62-84). Englewood Cliffs, USA: Prentice Hall.
- 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. In: *Flood frequency analysis* (pp. 12-21). Boca Raton, USA: CRC Press.

- Salvadori, G., De Michele, C., Kottegoda, N. T., & Rosso, R. (2007). Chapter 3. Bivariate analysis via Copulas and Appendix C. Families of Copulas. In: *Extremes in Nature. An approach using Copulas* (pp. 131-175, 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
- 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. (2004). Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resources Research*, 40(W12511), 1-17. DOI: 10.1029/2004WR003133
- Serinaldi, F. (2015). Dismissing return periods! *Stochastic Environmental Research and Risk Assessment*, 29(4), 1179-1189. DOI: 10.1007/s00477-014-0916-1
- Shiau, J. T., Wang, H. Y., & Tsai, C. T. (2006). Bivariate frequency analysis of floods using copulas. *Journal of the American Water Resources Association*, 42(6), 1549-1564. DOI: 10.1111/j.1752-1688-2006.tb06020.x
- Snider, D. (1972). Chapter 16. Hydrographs. In: *National Engineering Handbook, Section 4: Hydrology* (pp. 16.1-16.26). Washington, DC, USA: Soil Conservation Service, U. S. Department of Agriculture.

- Stegun, I. A. (1972). Chapter 27. Miscellaneous functions. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (pp. 997-1010). New York, USA: Dover Publications.
- 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.
- 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. (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. In: *Copulas and their applications in water resources engineering* (pp. 62-122, 123-171). Cambridge, United Kingdom. Cambridge University Press.
- 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)