

DOI: 10.24850/j-tyca-13-06-11

Artículos

## Aplicación de la distribución GVE bivariada en el análisis de frecuencias conjunto de crecientes

### Application of the bivariate GEV distribution in the joint flood frequency analysis

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, México, campos\_aranda@hotmail.com

Autor para correspondencia: Daniel Francisco Campos-Aranda, campos\_aranda@hotmail.com

#### Resumen

Las crecientes que ocurren en nuestro país cada año generan daños y ponen en peligro a los embalses, cuyo dimensionamiento hidrológico está basado en el *hidrograma de la creciete de diseño*. La estimación más simple de tal hidrograma se basa en el análisis de frecuencias conjunto



del gasto pico y volumen anuales. En este estudio se ajustó la *distribución general de valores extremos bivariada (GVEb)*, al registro de 55 crecientes anuales en la estación hidrométrica La Cuña, sobre el Río Verde de la Región Hidrológica No. 12-3, México. Este proceso abarca nueve etapas: (1) selección y prueba de los registros anuales; (2) verificación de su aleatoriedad; (3) estimación de las probabilidades empíricas conjuntas; (4) ajuste de la función *GVEb* a través del método de máxima verosimilitud; (5) validación de la función *GVEb*; (6) ratificación de las marginales GVE; (7) verificación de las restricciones de probabilidad; (8) estimación de eventos de diseño univariados híbridos, y (9) estimación de eventos de diseño conjuntos y selección del subgrupo crítico. En la etapa 1 se aplica un test simple de la GVE. La etapa 2 se realiza con base en el Test de Wald-Wolfowitz. En la etapa 4 se emplea el algoritmo Complex. Las etapas 5 y 6 utilizan el Test de Kolmogorov–Smirnov. En la etapa 9 se usan las gráficas del periodo de retorno conjunto de tipo AND. Por último, se formulan las conclusiones, las cuales destacan el enfoque de maximización adoptado y las ventajas de aplicar la *GVEb*.

**Palabras clave:** tipos de crecientes de diseño, distribución *GVEb*, distribuciones condicionales, probabilidades empíricas conjuntas, algoritmo Complex, validación de la función *GVEb*, periodos de retorno univariados híbridos, periodos de retorno conjuntos.

## Abstract

The floods in our country every year cause damage and endanger the reservoirs. Therefore, its hydrological dimensioning is based on the *hydrograph of the design flood*, and its most straightforward estimation is based on the joint frequency analysis of the annual peak flow and volume. In this study, the *bivariate general extreme values distribution (GVEb)* was adjusted to the record of the 55 annual floods at the La Cuña hydrometric station on the Río Verde of Hydrological Region No. 12-3, Mexico. This study encompasses the following nine stages: (1) selection and testing of annual records; (2) verification of the randomness of the records; (3) estimation of the joint empirical probabilities; (4) adjustment of the *GVEb* function through the maximum likelihood method; (5) validation of the *GVEb* function; (6) ratification of GVE marginal functions; (7) verification of probability constraints; (8) estimation of hybrid univariate design events, and (9) estimation of joint design events and selection of the critical subgroup. In stage 1, a simple test is applied based on the shape parameter of the marginal GVE. Stage 2 is carried out based on the Wald-Wolfowitz Test. In stage 4, the Complex algorithm is used. Stages 5 and 6 use the Kolmogorov-Smirnov Test. In stage 9, the graphs of the joint return period of type AND are used. Finally, conclusions are formulated, which highlight the maximization approach adopted and the advantages of the bivariate joint frequency analysis through the *GVEb*.

**Keywords:** Types of design floods, *GVEb* distribution, conditional distributions, joint empirical probabilities, complex algorithm, *GVEb* function validation, hybrid univariate return periods, joint return periods.

Recibido: 03/02/2021

Aceptado: 16/08/2021

## Introducción

### Tipos de obras hidráulicas

En nuestro país, por su localización geográfica ocurren anualmente *crecientes* severas en diversas regiones, las cuales generan daños sociales, ambientales y materiales. Tales crecientes ponen en peligro a todas las *obras hidráulicas* de aprovechamiento y protección (Aldama, 2000). Las obras hidráulicas de aprovechamiento más importantes son los *embalses*, los cuales se clasifican desde un punto de vista hidrológico



simple en grandes y pequeños. Los grandes embalses son decenas, operan de manera multianual y tienen diversos usos; debido a su tamaño, su ruptura o colapso origina una catástrofe de enormes proporciones que no se puede permitir. En cambio, los embalses pequeños son miles y tienen operación anual y un único propósito, comúnmente el riego. Su ruptura o colapso causa gran destrucción local y debe ser evitada.

Las principales obras hidráulicas de protección son los *diques* o *muros de contención*, los cuales evitan el desbordamiento del río en sus planicies de inundación, permitiendo el uso agrícola o recreativo de tales áreas. En cuanto a los *puentes*, aunque son obras de cruce de carreteras y ferrocarriles, su dimensionamiento hidrológico correcto evita su destrucción, o los daños superficiales o laterales en el terraplén de la carretera o ferrocarril debido a su tamaño reducido.

También son obras hidráulicas de protección todas las *estructuras del drenaje urbano*, las *presas de control de crecientes* y las *presas rompempicos*. Las primeras almacenan la creciente de diseño de manera temporal y la liberan con un gasto pequeño; las segundas reducen la creciente debido a su tránsito en su vaso o almacenamiento.

## Tipos de crecientes de diseño



El proceso de estimación del impacto posible de los eventos hidrológicos extremos en una obra hidráulica, y la selección de sus dimensiones y de la política de operación para su funcionamiento correcto se conoce como *diseño hidrológico* (Ramírez-Orozco & Aldama, 2000). El concepto asociado con las obras hidráulicas para evitar daños que lleven a su ruptura o colapso se denomina *seguridad hidrológica* (Aldama, Ramírez, Aparicio, Mejía-Zermeño & Ortega-Gil, 2006). Lógicamente, existen otros conceptos asociados con la seguridad estructural y el funcionamiento hidráulico.

Para el caso de los diques o muros de contención, los puentes, las rectificaciones y canalizaciones, y algunas obras del drenaje urbano, su dimensionamiento hidrológico se basa exclusivamente en el caudal o gasto máximo asociado con una baja probabilidad de ser excedido o *creciente de diseño*.

En cambio, para todos los embalses, presas de control y rompe-picos y los estanques urbanos, su dimensionamiento hidrológico requiere el *hidrograma* de la creciente de diseño, ya que éste determina su funcionamiento y/o desempeño. El hidrograma de la creciente es la gráfica que define la evolución en el tiempo (abscisas) del gasto o caudal (ordenadas). Tiene cuatro características o variables básicas: gasto pico, volumen (área bajo el hidrograma), tiempo al gasto pico y duración total.

## Crecientes de diseño univariadas

Los métodos hidrológicos de estimación de las crecientes se dividen en dos grandes grupos: (1) *hidrometeorológicos* y (2) *probabilísticos*. Los primeros son determinísticos, se aplican por subcuencas y utilizan los datos de lluvia de las estaciones pluviográficas y pluviométricas para definir las tormentas de diseño, las cuales se transforman en gastos o caudales mediante la modelación matemática del proceso lluvia-escurrecimiento. Los hidrogramas parciales estimados se trasladan e integran el hidrograma buscado (Teegavarapu, 2012; Mujumdar & Nagesh-Kumar, 2012).

Los métodos probabilísticos son más confiables y exactos, pero requieren datos de las crecientes máximas anuales observadas en el sitio del proyecto. Cuando procesan sólo los datos de los caudales o gastos máximos anuales, esta técnica estadística se conoce como *análisis de frecuencias de crecientes* (AFC) y consiste en representar el registro disponible de gastos pico por una *función de distribución de probabilidades* (FDP) y con base en tal modelo realizar las inferencias buscadas o *predicciones* asociadas con una determinada probabilidad de excedencia (*crecientes de diseño*), cuyo recíproco es el periodo de retorno en años (Kite, 1977; Stedinger, Vogel & Foufoula-Georgiou, 1993;

Hosking & Wallis, 1997; Rao & Hamed, 2000; Khaliq, Ouarda, Ondo, Gachon & Bobée, 2006; Meylan, Favre, & Musy, 2012; Stedinger, 2017; Teegavarapu, Salas, & Stedinger, 2019).

## Crecientes de diseño bivariadas

Como ya se indicó, una creciente es un evento hidrológico extremo *multivariado*. Al respecto, se ha demostrado que los embalses no son sensibles al valor del tiempo al gasto pico ( $Qp$ ) y que la duración total está relacionada con el volumen ( $V$ ); de manera que estas dos variables ( $Qp$  y  $V$ ) que están correlacionadas son suficientes para definir, de manera aproximada, el hidrograma de la creciente de diseño (Goel, Seth & Chandra, 1998; Yue, Ouarda, Bobée, Legendre & Bruneau, 1999; Aldama, 2000; Yue & Rasmussen, 2002; Aldama *et al.*, 2006).

El análisis *bivariado* de las crecientes es el enfoque multivariado más *simple* y aun así implica diversas complicaciones matemáticas; por ejemplo: (1) se debe emplear una FDP bivariada; (2) su validación requiere la estimación de las probabilidades empíricas bivariadas; (3) ahora existen probabilidades conjuntas y condicionales; (4) se debe definir un periodo de retorno conjunto, para el cual existen *infinitas* parejas de valores de  $Qp$  y  $V$  que lo satisfacen, y (5) hay que seleccionar

entre las parejas citadas los eventos críticos o de diseño (Aldama, 2000; Ramírez-Orozco & Aldama, 2000; Escalante-Sandoval & Reyes-Chávez, 2002; Volpi & Fiori, 2012; Requena, Mediero, & Garrote, 2013).

## Objetivos

En *objetivo global* de este estudio consistió en detallar las nueve etapas que integran al análisis de frecuencias conjunto de crecientes que procesa el gasto pico y volumen anuales. Los *objetivos específicos* se pueden resumir en los seis siguientes: (1) selección y prueba de los registros de gasto máximo y volumen anuales; (2) ajuste con momentos L de la distribución GVE; (3) descripción y ajuste de la distribución general de valores extremos bivariada (*GVEb*), mediante el método de máxima verosimilitud, se incluyen la validación de la *GVEb* a los datos bivariados y la ratificación de sus marginales; (4) exposición de la teoría de los periodos de retorno univariados y conjuntos; (5) selección de los eventos críticos en la gráfica del periodo de retorno conjunto de tipo AND, y (6) la teoría descrita se aplica al registro de 55 años de la estación hidrométrica La Cuña, del Río Verde en la Región Hidrológica No. 12-3 (Río Santiago), México.

## Teoría operativa

### Distribución general de valores extremos

La distribución GVE o *general de valores extremos* incluye tres familias de modelos probabilísticos llamadas: Gumbel, Fréchet y Weibull, que son rectas, y dos tipos de curvas en el papel de probabilidad extremo Gumbel-Powell (Chow, 1964). La expresión de su FDP es (Stedinger *et al.*, 1993; Escalante-Sandoval & Raynal-Villaseñor, 1994; Hosking & Wallis, 1997; Rao & Hamed, 2000):

$$F(x) = \exp\left\{-\left[1 - k\frac{(x-u)}{\alpha}\right]_+^{1/k}\right\} \quad (1)$$

en la cual  $u$ ,  $a$  y  $k$  son los parámetros de ubicación, escala y forma, con  $-\infty < u$ ,  $k < \infty$  y  $a > 0$ . La distribución GVE está definida por el conjunto  $\{x_i: [1 - k(x_i - u) / a] > 0\}$ , lo cual se indica con el signo + afuera del paréntesis rectangular.

Coles (2001) indica que cualquier combinación de parámetros de ajuste que viola la condición anterior de *positividad* implica que al menos uno de los puntos observados ( $x_i$ ) está más allá de los puntos finales de la distribución, y entonces la función de verosimilitud es cero y la función logarítmica de verosimilitud es igual a  $-\infty$ .

Si  $k < 0$ , la GVE tiene el tipo II o Fréchet, sin límite superior ( $u + a / k < x < \infty$ ) y concavidad hacia arriba. Cuando  $k > 0$ , se define el tipo III o Weibull, con límite superior ( $\infty < x < u + a / k$ ) y concavidad hacia abajo. Finalmente, cuando  $k = 0$ , se llega, en un sentido asintótico, a la distribución tipo I o Gumbel, también llamada doble exponencial; son líneas rectas en el papel de probabilidad extremo ( $-\infty < x < \infty$ ).

Al representar con la distribución GVE los valores máximos en cada año, las *predicciones* buscadas ( $X_{Tr}$ ) para una probabilidad  $q$  de excedencia o un periodo de retorno  $Tr$  en años ( $Tr = 1/q$ ) se obtienen con base en la Ecuación (1), pues  $F(x) = 1 - q$  (Hosking & Wallis, 1997; Rao & Hamed, 2000):

$$X_{Tr} = u + \frac{\alpha}{k} \{1 - [-\ln(1 - q)]^k\} \text{ para } k \neq 0 \quad (2)$$

## Modelo probabilístico GVE bivariado

Emil Julius Gumbel propuso a inicios de la década de 1960 el llamado *modelo logístico*, el cual acepta como distribuciones *marginales* las de valores extremos de tipo Gumbel, GVE, Gumbel mixta y TCEV (Escalante-Sandoval & Raynal-Villaseñor, 1994; Ramírez-Orozco & Aldama, 2000; Escalante-Sandoval & Reyes-Chávez, 2004; Yue & Wang, 2004; Escalante-Sandoval, 2005); su expresión es:

$$F(x, y, \theta) = \exp\{-[(-\ln F_X(x))^m + (-\ln F_Y(y))^m]^{1/m}\} \quad (m \geq 1) \quad (3)$$

en la cual  $\theta$  representa a los siete parámetros de ajuste;  $m$  es el *parámetro de asociación*, que describe la dependencia entre las dos variables aleatorias  $X$ ,  $Y$ . Cuando las marginales tienen distribución Gumbel su expresión es:

$$m = \frac{1}{\sqrt{1-\rho}} \quad (0 \leq \rho \leq 1) \quad (4)$$

siendo  $\rho$  el coeficiente de correlación de Pearson, cuya ecuación es:

$$\rho = \frac{E[(X-\mu_X)(Y-\mu_Y)]}{\sigma_X \cdot \sigma_Y} \quad (5)$$

en la cual  $(\mu_X, \sigma_X)$  y  $(\mu_Y, \sigma_Y)$  son la media y la desviación estándar poblacionales de las variables aleatorias  $X$  y  $Y$ . Si las distribuciones marginales  $F_X(x)$  y  $F_Y(y)$  son funciones GVE, sus expresiones son:

$$F(x) = \exp\left\{-\left[1 - k_1 \left(\frac{x-u_1}{\alpha_1}\right)_+\right]^{1/k_1}\right\} \quad (6)$$

$$F(y) = \exp\left\{-\left[1 - k_2 \left(\frac{y-u_2}{\alpha_2}\right)_+\right]^{1/k_2}\right\} \quad (7)$$

Sustituyendo las ecuaciones (6) y (7) en la (3), se obtiene la FDP bivariada con marginales GVE (Escalante-Sandoval & Raynal-Villaseñor, 1994):

$$F(x, y, \theta) = \exp\left\{-\left[\left(1 - k_1 \left(\frac{x-u_1}{\alpha_1}\right)_+\right)^{m/k_1} + \left(1 - k_2 \left(\frac{y-u_2}{\alpha_2}\right)_+\right)^{m/k_2}\right]^{1/m}\right\} \quad (8)$$

con  $\theta = u_1, \alpha_1, k_1, u_2, \alpha_2, k_2, m$ , cuyos parámetros de ajuste se obtienen con el método estadístico de máxima verosimilitud. Las probabilidades univariadas y conjunta del modelo logístico deben cumplir con la siguiente

restricción (Escalante-Sandoval & Raynal-Villaseñor, 1994; Escalante-Sandoval, 2005):

$$F_X(x) \cdot F_Y(y) < F(x,y) < \min[F_X(x), F_Y(y)] \quad (9)$$

## Ajuste de la distribución GVE con momentos L

Los momentos L ( $\lambda_s$ ) son combinaciones lineales de los momentos de probabilidad ponderada ( $\beta_r$ ) desarrollados por Greenwood, Landwehr, Matalas y Wallis (1979), los cuales son parámetros estadísticos de los datos ordenados. En una muestra de tamaño  $n$ , con sus elementos  $x_i$  arreglados en orden ascendente ( $x_1 \leq x_2 \leq \dots \leq x_n$ ), los estimadores insesgados de  $\beta_r$  se obtienen con la expresión general siguiente:

$$\beta_r = \frac{1}{n} \sum_{j=r+1}^n \frac{(j-1) \cdot (j-2) \cdot \dots \cdot (j-r)}{(n-1) \cdot (n-2) \cdot \dots \cdot (n-r)} x_j \quad (10)$$

Los momentos L son un sistema eficiente y robusto para estimar los parámetros de ajuste de diversas FDP utilizadas en los AFC, cuyas

ecuaciones de cálculo son (Stedinger *et al.*, 1993; Hosking & Wallis, 1997; Rao & Hamed, 2000; Stedinger, 2017):

$$\lambda_1 = \beta_0 \quad (11)$$

$$\lambda_2 = 2 \cdot \beta_1 - \beta_0 \quad (12)$$

$$\lambda_3 = 6 \cdot \beta_2 - 6 \cdot \beta_1 + \beta_0 \quad (13)$$

El cociente ( $\tau$ ) de momentos L de similitud con el de asimetría es:

$$\tau_3 = \lambda_3 / \lambda_2 \quad (14)$$

Las ecuaciones que permiten estimar los tres parámetros de ajuste de la distribución GVE son:

$$k \cong 7.8590 \cdot c + 2.9554 \cdot c^2 \quad (15)$$

siendo:



$$c = \frac{2}{3+t_3} - 0.63093 \quad (16)$$

$$\alpha = \frac{\lambda_2 \cdot k}{(1-2^{-k}) \cdot \Gamma(1+k)} \quad (17)$$

$$u = \lambda_1 - \frac{\alpha}{k} [1 - \Gamma(1+k)] \quad (18)$$

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

$$\Gamma(\omega) \cong e^{-\omega} \cdot \omega^{\omega-\frac{1}{2}} \cdot (2\pi)^{1/2} \cdot F1 \quad (19)$$

siendo:

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

## Método de máxima verosimilitud

El *principio de máxima verosimilitud* se explica con diferentes niveles de detalle según el texto consultado (Kite, 1977; Metcalfe, 1997; Rao & Hamed, 2000; Coles, 2001; Kottegoda & Rosso, 2008; Meylan *et al.*, 2012). Pero en general define lo siguiente: dada una muestra aleatoria  $(X_1, X_2, \dots, X_n)$  de observaciones independientes e idénticamente distribuidas (*iid*) que siguen una FDP denominada  $F_\theta$  con parámetros de ajuste  $\theta_1, \theta_2, \dots, \theta_q$ . Entonces, por definición, la probabilidad de obtener un valor  $X_i$ , será:

$$P(x_i \leq X \leq x_i + dx_i) = f_\theta(x_i) \cdot dx_i \quad (20)$$

siendo  $f_\theta(x_i)$  la *función de densidad de probabilidad (fdp)*. Como los datos son *iid*, la probabilidad de obtener  $n$  valores  $X_i$  será la probabilidad conjunta o *función de verosimilitud*, designada  $L$  del inglés *likelihood*, y que se expresa como:

$$L(\theta) = f_\theta(x_1) \cdot f_\theta(x_2) \cdots f_\theta(x_n) = \prod_{i=1}^n f_\theta(x_i) \quad (21)$$

El método de máxima verosimilitud consiste en encontrar un vector  $\hat{\theta}$  de parámetros que hagan máxima a  $L(\theta)$  y por lo tanto a la probabilidad de obtener la muestra  $(X_1, X_2, \dots, X_n)$ . Con frecuencia resulta más conveniente tomar logaritmos y trabajar con la *función logarítmica de verosimilitud* (Coles, 2001), es decir:

$$l(\theta) = \log L(\theta) = \sum_{i=1}^n \log f_{\theta}(x_i) \quad (22)$$

lo anterior es aceptable debido a que la función logarítmica es monotónica y entonces la función  $l(\theta)$  alcanza su máximo en el mismo punto que la función  $L(\theta)$ .

## Ajuste de la GVE por máxima verosimilitud

La siguiente descripción se centra en el proceso iterativo que maximiza la función logarítmica de verosimilitud (Rao & Hamed, 2000):

$$l(\theta) = -n \cdot \ln \alpha - (1 - k) \sum_{i=1}^n v_i - \sum_{i=1}^n e^{-v_i} \quad (23)$$

siendo:

$$v_i = -\frac{1}{k} \ln \left[ 1 - k \left( \frac{x_i - u}{\alpha} \right) \right] \quad (24)$$

El proceso comienza evaluando  $P$ ,  $Q$  y  $R$  con las expresiones:

$$P = n - \sum_{i=1}^n e^{-v_i} \quad (25)$$

$$Q = \sum_{i=1}^n e^{kv_i-k} - (1-k) \sum_{i=1}^n e^{kv_i} \quad (26)$$

$$R = n - \sum_{i=1}^n v_i + \sum_{i=1}^n v_i e^{-v_i} \quad (27)$$

El cálculo de los incrementos en cada iteración ( $j$ ) se realiza con las ecuaciones siguientes:

$$\delta u_j = -\left(\frac{\alpha_j}{n}\right) \left\{ b \cdot Q_j + h \cdot \left(\frac{P_j+Q_j}{k_j}\right) + \frac{f}{k_j} \cdot \left[ R_j - \left(\frac{P_j+Q_j}{k_j}\right) \right] \right\} \quad (28)$$

$$\delta \alpha_j = -\left(\frac{\alpha_j}{n}\right) \left\{ h \cdot Q_j + a \cdot \left(\frac{P_j+Q_j}{k_j}\right) + \frac{g}{k_j} \cdot \left[ R_j - \left(\frac{P_j+Q_j}{k_j}\right) \right] \right\} \quad (29)$$

$$\delta k_j = -\left(\frac{1}{n}\right) \left\{ f \cdot Q_j + g \cdot \left(\frac{P_j+Q_j}{k_j}\right) + \frac{c}{k_j} \cdot \left[ R_j - \left(\frac{P_j+Q_j}{k_j}\right) \right] \right\} \quad (30)$$

En las expresiones anteriores,  $a$ ,  $b$ ,  $c$ ,  $f$ ,  $g$  y  $h$  son los coeficientes de la matriz de varianza-covarianza, función del parámetro de forma ( $k$ ),

con la siguiente expresión general y los valores mostrados en la Tabla 1 (Campos-Aranda, 2003):

$$\text{coef.} = C0 + C1 \cdot k + C2 \cdot k^2 + C3 \cdot k^3 + C4 \cdot k^4 \quad (31)$$

**Tabla 1.** Parámetros de ajuste de la Ecuación (31).

coef.	Valores numéricos de:				
	C0	C1	C2	C3	C4
<i>a</i>	0.6524995	-0.5580180	1.0877760	-0.0617186	-0.1209662
<i>b</i>	1.2474540	-0.1987501	-0.2179270	0.0626601	0.1053422
<i>c</i>	0.4737703	-0.7670770	0.2818301	-0.1587989	0.1703716
<i>f</i>	0.2577352	-0.1293403	-0.3234050	0.0312508	0.1636221
<i>g</i>	0.1441769	0.4209507	-0.4192687	-0.0467502	-0.0986341
<i>h</i>	0.3358007	-1.1933840	-0.1146433	-0.0658476	0.0381460

El proceso inicia adoptando los parámetros de ajuste obtenidos con el método de los momentos L, como valores iniciales  $u_j$ ,  $a_j$  y  $k_j$ , para estimar sus valores en la siguiente iteración con las ecuaciones:

$$u_{j+1} = u_j + \delta u_j \quad (32)$$

$$\alpha_{j+1} = \alpha_j + \delta\alpha_j \quad (33)$$

$$k_{j+1} = k_j + \delta k_j \quad (34)$$

En cada iteración se calculan nuevos valores de los seis coeficientes  $a-h$  (Ecuación (31)), ya que  $k_j$  cambia. Los nuevos valores  $u_{j+1}$ ,  $a_{j+1}$  y  $k_{j+1}$ , se llevan a las ecuaciones (24) a (30) y el proceso concluye cuando tales incrementos son menores que la tolerancia adoptada; por ejemplo, de 0.00001.

## **$I(\theta)$ de la GVE bivariada**

La función logarítmica de verosimilitud del modelo probabilístico *GVEb* para el caso simple de variables aleatorias  $X$ ,  $Y$  con igual amplitud de registro ha sido expuesta por Escalante-Sandoval y Raynal-Villaseñor (1994), y es la siguiente:

$$\begin{aligned}
 l(\theta) = \sum_{i=1}^n \left\{ -(\ln \alpha_1 + \ln \alpha_2) + \ln \left[ 1 - k_1 \left( \frac{x_i - u_1}{\alpha_1} \right)_+^{\frac{m}{k_1} - 1} + \ln \left[ 1 - k_2 \left( \frac{y_i - u_2}{\alpha_2} \right)_+^{\frac{m}{k_2} - 1} + \right. \right. \\
 \ln \left[ \left( 1 - k_1 \left( \frac{x_i - u_1}{\alpha_1} \right)_+^{\frac{m}{k_1}} + \left( 1 - k_2 \left( \frac{y_i - u_2}{\alpha_2} \right)_+^{\frac{m}{k_2}} \right)^{\frac{1}{m} - 2} \right] + \ln \left[ (m - 1) + \left( \left( 1 - \right. \right. \right. \\
 \left. \left. \left. k_1 \left( \frac{x_i - u_1}{\alpha_1} \right)_+^{\frac{m}{k_1}} + \left( 1 - k_2 \left( \frac{y_i - u_2}{\alpha_2} \right)_+^{\frac{m}{k_2}} \right)^{\frac{1}{m}} \right) \right] - \left[ \left( 1 - k_1 \left( \frac{x_i - u_1}{\alpha_1} \right)_+^{\frac{m}{k_1}} + \left( 1 - \right. \right. \right. \\
 \left. \left. \left. k_2 \left( \frac{y_i - u_2}{\alpha_2} \right)_+^{\frac{m}{k_2}} \right)^{\frac{1}{m}} \right] \right\} \quad (35)
 \end{aligned}$$

La expresión anterior es una de las cinco partes que integran la  $l(\theta)$  del modelo trivariado GVE, con amplitudes diferentes en cada registro  $X$ ,  $Y$  y  $Z$ , desarrollado por Escalante-Sandoval y Raynal-Villaseñor (1994).

## El algoritmo Complex

La maximización de la Ecuación (35) para obtener los siete parámetros *óptimos* de ajuste  $(\theta = u_1, \alpha_1, k_1, u_2, \alpha_2, k_2, m)$  del modelo *GVEb* se debe abordar de manera *numérica* dada la complejidad de tal ecuación y de

sus derivadas parciales con respecto a  $\theta$ ; para ello, se seleccionó el algoritmo Complex de múltiples variables ( $z$ ) restringidas o acotadas. Su planteamiento teórico es el siguiente (Box, 1965):

$$\text{Minimizar } F(z_1, z_2, \dots, z_s) \quad (36)$$

Sujeta a  $w$  variables dependientes ( $y$ ), función de las variables de decisión ( $z$ ):

$$\begin{aligned} y_1 &= F(z_1, z_2, \dots, z_s) \\ &\vdots \\ y_w &= F(z_1, z_2, \dots, z_s) \end{aligned} \quad (37)$$

Ambas variables tienen límites inferiores y superiores del tipo  $\leq$ ; es decir,  $z_{inf} \leq z_i \leq z_{sup}$  y  $y_{inf} \leq y_i \leq y_{sup}$ . El *algoritmo Complex* es una técnica de exploración local, 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; para el caso analizado siete ( $u_1, \alpha_1, k_1, u_2, \alpha_2, k_2, m$ ) y  $2n$  (dos veces el número de años

del registro) debido a que las variables dependientes son las restricciones de positividad:  $[1 - k_1(x_i - u_1)/\alpha_1]$  y  $[1 - k_2(y_i - u_2)/\alpha_2]$ . MI = 500 es el número máximo de evaluaciones de la función objetivo (FO) y NQ = 25 el número de tales cálculos entre impresión de resultados. Estas variables se definen en la subrutina de lectura de datos.

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 7 y J de 1 a  $2n$  (doble del número de años del registro). Después se incluyen los criterios de convergencia FA y FR para las desviaciones absoluta y relativa de la FO. Se utilizaron los valores siguientes: 0.0002 y 0.00001, respectivamente.

La función objetivo se denomina FO en el código OPTIM y se define al final del programa; corresponde lógicamente a la Ecuación (35), con nombre FO\$ = "FLV" de función logarítmica de verosimilitud. Se asigna con signo negativo a FO debido a que el algoritmo Complex minimiza a la función (Ecuación (36)) y se desea maximizar a la FLV.

## Periodos de retorno univariados



El concepto clásico de probabilidad de un evento se define como el cociente del número de casos favorables ( $ncf$ ) a tal evento entre el número de casos posibles ( $ncp$ ) a dicho evento, por ello varía de cero a uno. Debido al manejo anual de la variable  $X$ , la probabilidad de excedencia  $F'_X(x)$  se corresponde con el recíproco del *periodo de retorno* ( $T_X$ ) en años, pues en cada año se tiene  $ncf = 1$  y  $ncp = T_X$ , esto es (Yue & Rasmussen, 2002; Shiau, 2003):

$$T_X = \frac{1}{F'_X(x)} = \frac{1}{1-F_X(x)} \quad (38)$$

$$T_Y = \frac{1}{F'_Y(y)} = \frac{1}{1-F_Y(y)} \quad (39)$$

En las expresiones anteriores,  $F_X(x)$  y  $F_Y(y)$  son la probabilidad de no excedencia que se estima con las ecuaciones (6) y (7).

## Periodos de retorno conjuntos

El primer *periodo de retorno conjunto* del evento  $(X, Y)$  se define bajo la condición OR y la Ecuación (38) de la siguiente manera (Goel *et al.*, 1998; Yue, 2000b; Shiau, 2003; Requena *et al.*, 2013; Vogel & Castellarin, 2017):

$$T(x, y) = \frac{1}{1-F(x, y)} \text{ siendo } F(x, y) = P(X \leq x \text{ ó } Y \leq y) \quad (40)$$

En la expresión anterior,  $F(x, y)$  es la probabilidad de no excedencia *conjunta* que se estima con la Ecuación (8), previa estimación de sus parámetros de ajuste  $(u_1, \alpha_1, k_1, u_2, \alpha_2, k_2, m)$  con el método de máxima verosimilitud. Este evento representa el caso en que los límites  $x$  o  $y$ , o ambos pueden ser excedidos ( $X > x$ , o  $Y > y$ ; o bien,  $X > x$  y  $Y > y$ ).

El segundo *periodo de retorno conjunto* del evento  $(X, Y)$  está asociado al caso en que ambos límites son excedidos ( $X > x$ ,  $Y > y$ ) o condición AND; su ecuación es (Goel *et al.*, 1998; Aldama, 2000; Ramírez-Orozco & Aldama, 2000; Yue, 2000b; Shiau, 2003; Requena *et al.*, 2013; Vogel & Castellarin, 2017):

$$T'(x, y) = \frac{1}{F'(x, y)} = \frac{1}{1+F(x, y)-F_X(x)-F_Y(y)} \quad (41)$$

Aldama (2000) obtiene la expresión  $F'(x, y)$  de la probabilidad conjunta de excedencia mediante un razonamiento de probabilidades

lógico y simple aplicado en el plano cartesiano. En cambio, Yue y Rasmussen (2002) recurren al plano cartesiano para definir numéricamente un evento bivariado  $(X, Y)$ , que puede ocurrir en alguno de los cuatro cuadrantes y entonces obtienen la ecuación de la probabilidad conjunta de excedencia (denominador de la Ecuación (41)).

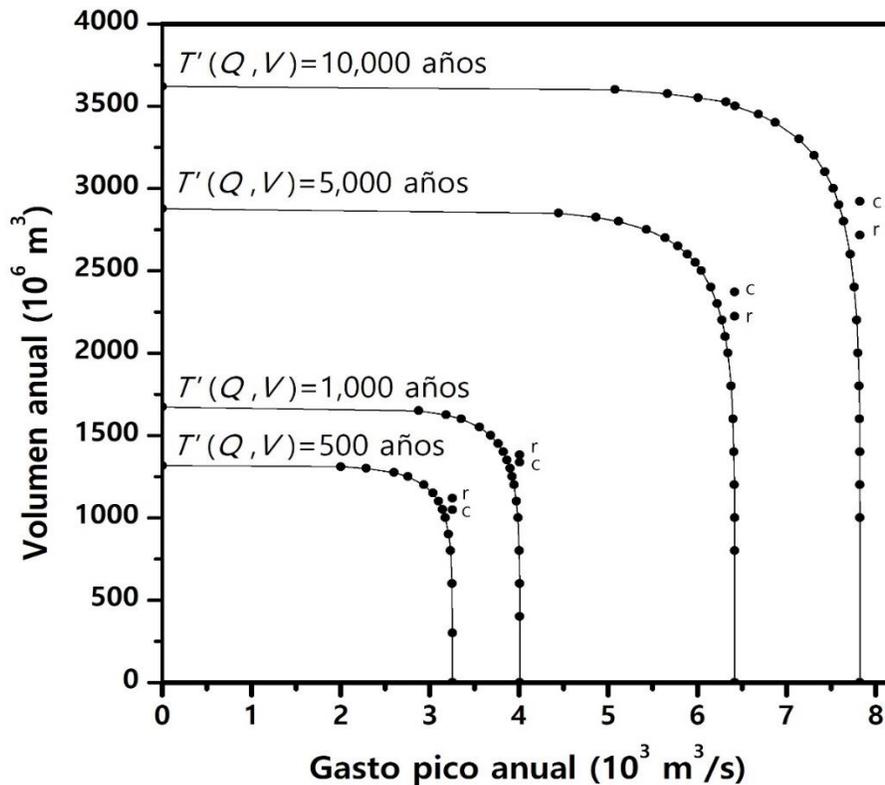
Yue y Rasmussen (2002) también establecen las siguientes relaciones entre los periodos de retorno univariados y los periodos de retorno conjuntos:

$$T(x, y) \leq \text{mínimo}(T_x, T_Y) \quad (42)$$

$$T'(x, y) \geq \text{máximo}(T_x, T_Y) \quad (43)$$

Diversos autores (Yue, 2000b; Yue & Rasmussen, 2002; Shiau, 2003; Vogel & Castellarin, 2017) han mostrado las gráficas de los dos periodos de retorno conjuntos y han discutido sus diferencias. Pero han sido Volpi y Fiori (2012) quienes han establecido un procedimiento para seleccionar sus *eventos críticos*.

En la Figura 1 se muestra la gráfica del periodo de retorno conjunto  $T'(Q, V)$ , construida con los datos de la aplicación numérica, la cual se describe posteriormente.



**Figura 1.** Gráficas de los cuatro periodos de retorno conjunto  $T'(Q, V)$  de diseño estimadas con la distribución  $GVEb$  para las crecientes anuales registradas en la estación hidrométrica La Cuña, México.

De acuerdo con Yue *et al.* (1999), Yue (2000b), y Yue y Rasmussen (2002) existe un tercer tipo de *periodo de retorno conjunto*, que tiene aplicación en la práctica hidrológica, y que se define para un evento  $X$ , dado que  $Y \leq y$  o para un evento  $Y$ , dado que  $X \leq x$  y por ello se designan

*condicionales*. Para tales eventos, sus distribuciones de probabilidad condicional se definen de manera simple con estas ecuaciones:

$$F(X|Y \leq y) = \frac{F(x,y)}{F(y)} \quad (44)$$

$$F(Y|X \leq x) = \frac{F(x,y)}{F(x)} \quad (45)$$

Sustituyendo las expresiones anteriores en las ecuaciones (38) y (39) se obtienen las fórmulas del *periodo de retorno conjunto condicional*:

$$T_{X|Y} = \frac{1}{1-F(X|Y \leq y)} \quad (46)$$

$$T_{Y|X} = \frac{1}{1-F(Y|X \leq x)} \quad (47)$$

## Eventos críticos o de diseño del $T'(x,y)$



Volpi y Fiori (2012) destacan que la gráfica del periodo de retorno conjunto de tipo AND, mostrada como Figura 1, presenta una severa inconsistencia al contener, en un contexto bivariado, umbrales críticos univariados. Debido a lo anterior, tal gráfica se considera integrada por dos porciones: las dos designadas *simples* (*naive part*) y la *correcta* (*proper part*).

Las partes rectas son las colas o rectas asíntotas a la parte curva. La probabilidad de ocurrencia de un evento o pareja de  $Qp$  y  $V$  es variable en la parte curva y decrece a lo largo de la parte recta, aunque todos los valores definen el mismo periodo de retorno conjunto.

En resumen, las parejas de valores de las rectas asíntotas tienen probabilidades de ocurrencia bajas y por ello no deben ser incluidos en los análisis de búsqueda de las crecientes ( $Qp$  y  $V$ ) críticas o severas. Para fines prácticos, los puntos extremos de la parte curva se pueden definir según su distribución empírica o cercanos al inicio de las rectas asíntotas.

## Test de Wald-Wolfowitz

Esta prueba no paramétrica ha sido utilizada por Bobée y Ashkar (1991), Rao y Hamed (2000), y Meylan *et al.* (2012) para probar *independencia*

y *estacionariedad* en registros de gastos máximos anuales ( $X_i$ ). De acuerdo con el tipo de información hidrométrica procesada, se propuso aplicar la prueba a los registros de gastos pico y volúmenes anuales que deben ser muestras de valores aleatorios. El estadístico de la prueba es:

$$R = \sum_{i=1}^{n-1} X_i \cdot X_{i+1} + X_n \cdot X_1 \quad (48)$$

Cuando el tamaño ( $n$ ) de la serie o muestra ( $x_i$ ) no es pequeño y sus datos son independientes,  $R$  procede de una distribución Normal con media y varianza, dadas por las expresiones siguientes:

$$E[R] = \bar{R} = \frac{S_1^2 - S_2}{n-1} \quad (49)$$

$$Var[R] = \frac{S_2^2 - S_4}{n-1} + \frac{S_1^4 - 4 \cdot S_1^2 \cdot S_2 + 4 \cdot S_1 \cdot S_3 + S_2^2 - 2 \cdot S_4}{(n-1)(n-2)} - \bar{R}^2 \quad (50)$$

en las cuales:

$$S_k = \sum_{i=1}^n X_i^k \quad (51)$$

Finalmente, se calcula  $U$  con la ecuación:

$$U = \frac{R - \bar{R}}{\sqrt{\text{VAR}[R]}} \quad (52)$$

El valor de  $U$  sigue una distribución Normal (0, 1) y se puede usar para probar la independencia de los datos de la serie con un nivel de significancia  $\alpha$ , comúnmente del 5 %. En una prueba de dos colas, la variable normal estandarizada es  $Z_{\alpha/2} \cong 1.96$ ; entonces, cuando el valor absoluto de  $U$  es menor que 1.96, la serie estará integrada por valores independientes (muestra aleatoria).

## Selección y prueba de los registros por procesar

El libro de Aldama, Ramírez, Aparicio, Mejía-Zermeño y Ortega-Gil (2006) es un referente en los análisis de frecuencias bivariados de crecientes. Presenta 16 registros conjuntos de gasto máximo y volumen anuales de entrada en presas importantes de México y en su página 43 indica que se obtuvieron de los registros existentes de gastos máximos anuales e hidrogramas máximos anuales. Lo anterior implica que los registros

conjuntos fueron integrados por los eventos máximos anuales no pertenecientes a cada creciente anual. En cambio, en su página 44 indica que los registros hidrométricos continuos, de naturaleza gráfica o tabular, permitieron identificar los datos de gasto y volumen de las avenidas.

En opinión del autor, la segunda aclaración del libro de Aldama *et al.* (2006) es la correcta, como indicaron desde los inicios de los análisis de frecuencias bivariados, Yue (1999) y Yue *et al.* (1999), que establecieron que los gastos máximos y su volumen deben proceder del hidrograma de la creciente anual. Además, y por lo general, tanto el gasto pico máximo como su volumen de escurrimiento respectivo incluyen el gasto base (Shiau, Wang, & Tsai, 2006).

Por otra parte, para evitar la subjetividad en la búsqueda de los registros de  $Q_p$  y  $V$ , que *acepten* a la distribución GVE como función marginal, se propone aplicar la prueba potente desarrollada por Stedinger *et al.* (1993), la cual establece que cuando el valor absoluto de la estadística  $Z$  excede a la desviación Normal estándar de 1.645, se deduce que el parámetro de forma  $k$  es estadísticamente diferente de cero, y por lo tanto se debe ajustar la distribución GVE y no la Gumbel. La estadística es:

$$Z = k\sqrt{n/0.5633} \quad (53)$$

Para la búsqueda de registros que acepten a la distribución GVE como marginal, se les ajusta tal modelo probabilístico con el método de momentos L (ecuaciones (10) a (19)) y se aplica la Ecuación (53) si el valor absoluto de  $Z$  excede a 1.645; tales datos se pueden representar por la distribución GVE.

## Estimación de probabilidades empíricas

Las probabilidades de no excedencia empíricas univariadas y bivariadas se estimaron con base en la fórmula de Cunnane, que de acuerdo con Stedinger (2017) conduce a probabilidades de no excedencia ( $p$ ) aproximadamente insesgadas para muchas FDP utilizadas en los AFC; su expresión es:

$$p = \frac{i-0.40}{n+0.20} \quad (54)$$

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 y volumen anual.

Para el caso de las probabilidades bivariadas, se trabajó en el plano bidimensional, con el gasto en los renglones y los volúmenes en las columnas (Yue *et al.*, 1999; Yue, 2000b; Yue & Rasmussen, 2002; Yue & Wang, 2004). El proceso numérico comienza guardando los registros históricos de gasto máximo ( $Q$ ) y volumen ( $V$ ) anuales en archivos  $Qh$  y  $Vh$ ; además, se ordenaron en forma progresiva de magnitud en archivos  $Qo$  y  $Vo$ . Enseguida se procesaron por parejas anuales y se comparó cada valor  $Qh$  contra  $Qo$ , y se contabilizaron las veces que el segundo resultó menor o igual, y se designó  $NQ$ . Lo mismo se hizo con  $Vh$  contra  $Vo$  para obtener  $NV$ . Lo anterior equivale a cambiar el dato original de cada pareja de valores anuales históricos por su número de orden o *rango*.

Después se compara cada pareja histórica de rangos contra todas las demás y se contabilizan las veces en que ambos rangos (condición AND) resultan menores; tal cantidad se denomina  $NQV$ ; es decir, el número de ocurrencias de combinaciones de  $q$  y de  $v$  menores en el plano bidimensional. Por último, se aplica la fórmula de posición gráfica de Cunnane; para el caso bivariado es la siguiente:

$$F_e(x, y) = P(Q \leq q, V \leq v) = \frac{NQV_i - 0.40}{n + 0.20} \quad (55)$$

## Validación de la distribución *GVEb*

Yue (2000a) indica que la relación entre las probabilidades conjuntas empíricas y teóricas del gasto pico y el volumen permiten definir la validez de la distribución conjunta propuesta. La forma más simple de representarlas consiste en llevar al eje de las abscisas la probabilidad de no excedencia empírica y en el eje de las ordenadas la teórica; lógicamente, cada pareja de datos define un punto que coincide o se aleja de la recta a 45°. La inspección de la gráfica descrita y el valor del coeficiente de correlación —en estos casos superior a 0.98— ratifican la validez del modelo probabilístico conjunto utilizado.

Yue (2000b), y 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 (dif)* entre las probabilidades conjuntas. Para evaluar la estadística ( $D_n$ ) del test, la cual es función del número de datos ( $n$ ), se utilizó la expresión que exponen Meylan *et al.* (2012) para  $\alpha = 5$  %:

$$D_n = \frac{1.358}{\sqrt{n}} \quad (56)$$

Si  $dif$  es menor que  $D_n$ , se acepta el modelo probabilístico conjunto o distribución GVE bivariada.

## Registro de datos por procesar

La estación de aforos La Cuña con clave 12054 y área de cuenca de 19097 km<sup>2</sup>, se ubica en el Río Verde de la Región Hidrológica No. 12-3 (Río Santiago), México. Gómez, Aparicio y Patiño (2010) exponen el registro de gasto máximo en m<sup>3</sup>/s y volúmenes anuales en millones de m<sup>3</sup> (Mm<sup>3</sup>), del periodo de 1947 a 2004, con 55 parejas de datos ( $n$ ), ya que los años 1983, 1984 y 1985 no tienen información de volúmenes. Tales datos se muestran en la Tabla 2.

**Tabla 2.** Gastos pico y volúmenes anuales en la estación hidrométrica La Cuña de la Región Hidrológica No. 12-3 (Río Santiago), México (Gómez *et al.*, 2010).

Año	$Q_p$ (m <sup>3</sup> /s)	V (Mm <sup>3</sup> )	Año	$Q_p$ (m <sup>3</sup> /s)	V (Mm <sup>3</sup> )
1947	784.0	146.80	1975	622.1	249.07

1948	736.8	155.12	1976	1374.0	527.96
1949	510.0	111.40	1977	439.7	111.77
1950	461.0	94.06	1978	280.2	66.23
1951	411.0	111.55	1979	267.2	45.80
1952	326.0	70.82	1980	287.3	99.60
1953	349.8	144.75	1981	280.7	28.70
1954	130.4	23.22	1982	156.5	35.37
1955	690.0	203.31	1986	698.2	193.51
1956	266.0	106.76	1987	184.7	55.39
1957	199.0	45.92	1988	595.2	242.21
1958	690.0	188.71	1989	110.2	42.49
1959	340.6	47.91	1990	523.9	248.07
1960	249.6	91.58	1991	1636.3	443.30
1961	350.0	130.68	1992	1168.0	172.49
1962	317.0	51.27	1993	295.0	96.50
1963	732.6	127.90	1994	212.8	53.55
1964	265.1	82.75	1995	367.4	114.61
1965	743.6	295.34	1996	144.6	57.43
1966	463.9	202.90	1997	78.4	16.55

1967	1 474.9	598.38	1998	261.9	66.17
1968	323.0	118.25	1999	196.3	41.15
1969	160.4	32.22	2000	46.8	18.62
1970	763.8	187.75	2001	313.8	75.78
1971	578.0	166.61	2002	319.6	153.79
1972	191.8	26.39	2003	621.1	326.28
1973	2440.0	920.30	2004	824.5	384.45
1974	238.4	66.66	Mediana	340.6	111.40

## Secuencia del análisis de frecuencias bivariado

El proceso del análisis de frecuencias conjunto se considera integrado por dos partes: la primera se denomina "Ajuste de la *GVEb*" y la segunda "Estimación de Eventos de Diseño". Integrado el registro conjunto de gastos máximos y volúmenes anuales, se prueba su aleatoriedad con el test de Wald-Wolfowitz. Si ambas series probaron no tener componentes determinísticas, se verifica si proceden de una distribución GVE. En caso afirmativo, se ajusta tal modelo con el método de los momentos L. Enseguida se estiman las probabilidades de no excedencia bivariadas

empíricas, contra las cuales se contrastan los valores de la probabilidad teórica, estimada con el modelo *GVEb*, ajustado a través del algoritmo Complex para maximizar la función de máxima verosimilitud. Se valida la *GVEb* y se ratifican las marginales GVE mediante el test de Kolmogorov-Smirnov. Esta primera parte termina verificando las restricciones de probabilidad conjunta.

La segunda parte comienza estimando con la distribución GVE las predicciones univariadas y después los eventos de diseño univariados híbridos por regresión y condicionales. Enseguida se estiman los eventos de diseño conjuntos de tipo AND, con auxilio del modelo *GVEb*, y se seleccionan sus valores críticos, ubicados en la parte curva de las gráficas de tal periodo de retorno conjunto.

## Resultados y su discusión

### Verificación de la aleatoriedad



A los registros de gasto pico y volumen anual de la Tabla 2 se les aplicaron las ecuaciones (48) a (52), para probar su independencia y estacionariedad. Se encontró que ambas series son *aleatorias*, con los valores siguientes para el estadístico  $U$ : 0.284 y 0.213.

### Aceptación de las marginales GVE

A los registros de gasto pico y volumen anuales de la Tabla 2 se les ajustaron distribuciones GVE con los métodos de momentos L (ecuaciones (10) a (19)) y máxima verosimilitud (ecuaciones (24) a (34)). En la Tabla 3 se muestran los estadísticos básicos y los valores de los parámetros de ajuste de cada distribución GVE.

**Tabla 3.** Parámetros estadísticos y de ajuste de la distribución GVE en los registros de gasto pico y volumen anuales de la estación hidrométrica La Cuña, México.

Datos (ma*)	Momentos L			Parámetros de ajuste		
	$\lambda_1$	$\lambda_2$	$\tau_3$	Ubic.	Escala	Forma
$Q_p$ (mL)	499.8746	205.4937	0.38713	296.28	201.70	-0.3131

$Q_p$ (mv)				302.40	205.04	-0.3068
$V$ (mL)	154.8391	74.48538	0.43326	79.51	66.03	-0.3734
$V$ (mv)				85.67	69.71	-0.3543

\*método de ajuste: (mL) momentos L y (mv) máxima verosimilitud.

La aplicación de la Ecuación (53) a los resultados de método de máxima verosimilitud, con  $n = 55$  conduce a un valor absoluto de  $Z$  de 3.03 para el registro del gasto pico y de 3.50 para el de volumen; por lo cual, los registros de la Tabla 2 se pueden representar por la distribución GVE.

## Búsqueda *inicial* de los parámetros óptimos de la distribución *GVEb*

De inicio se buscó maximizar la Ecuación (35) mediante el *algoritmo Complex* y para ello se asignaron valores iniciales a las primeras seis variables de decisión con magnitud similar a los parámetros de ajuste de cada función marginal GVE (Tabla 3). Los límites inferior y superior de  $u_1$ ,  $a_1$ ,  $u_2$  y  $a_2$  se hicieron variar de manera arbitraria. Los límites adoptados para el parámetro de forma ( $k$ ) fueron -0.15 y 0.75. Para el parámetro

de asociación ( $m$ ), sus límites se definieron de manera aproximada con la Ecuación (4), variando de 1.0 a 3.8; este último valor corresponde a un  $\rho = 0.9302$  o correlación calculada para los registros de la Tabla 2.

Con tal enfoque se comprobó que los parámetros de ajuste ( $u$ ,  $a$ ,  $k$ ) óptimos resultaban muy diferentes a los iniciales: en general, el parámetro de asociación  $m$  final era el valor máximo adoptado y los parámetros de forma  $k$  óptimos sus límites inferiores asignados; esto llevaría a predicciones conjuntas inferiores a las univariadas. Además, se observó que el proceso de maximización de la Ecuación (35) dependía notoriamente de los valores iniciales y de sus límites en las siete variables de decisión, no permitiendo alcanzar resultados consistentes u obtención del mínimo global.

## **Búsqueda *combinada* de los parámetros óptimos de la distribución *GVEb***

En este nuevo enfoque de maximización de la Ecuación (35), el algoritmo Complex se aplicó bajo las tres normas siguientes: (1) se asignaron valores iniciales *únicos* a los seis parámetros de ajuste ( $u_1$ ,  $a_1$ ,  $k_1$ ,  $u_2$ ,  $a_2$ ,  $k_2$ ) iguales a los obtenidos con el método de máxima verosimilitud en el

ajuste de las distribuciones marginales GVE (Tabla 3); (2) sus límites inferior y superior se obtuvieron multiplicando por 0.90 y por 1.10 los valores iniciales, por lo tanto, tampoco variaron; y (3) se definieron cinco intervalos para el parámetro de asociación ( $m$ ) y se adoptó como valor inicial el valor medio, tales límites fueron: 1.0-1.5, 1.5-2.0, 2.0-2.5, 2.5-3.0 y 3.0-3.5. En la Tabla 4 se muestran los resultados principales de las cinco aplicaciones descritas del algoritmo Complex, con excepción de su última columna, cuya estimación se detalla enseguida.

**Tabla 4.** Resultados óptimos del algoritmo Complex durante la maximización de la Ecuación (35), con la *búsqueda combinada*, en la estación hidrométrica La Cuña, México.

FO inicial	FO final	Núm. iter.	Parámetros de ajuste:			$m$ inic.	$r_{xy}$
			$u_1, u_2$	$a_1, a_2$	$k_1, k_2$	$m$ final	
630.1	612.6	187	278.269	185.005	-0.2700	1.25	0.9918
			77.014	63.000	-0.3192	1.50	
618.4	601.5	284	270.387	185.027	-0.2700	1.75	0.9942
			80.555	63.000	-0.3129	2.00	
607.8	591.6	221	274.302	185.026	-0.2700	2.25	0.9928
			90.005	63.000	-0.3181	2.50	
598.5	582.9	267	270.000	185.052	-0.2700	2.75	0.9954
			79.120	63.000	-0.3194	3.00	

590.3	575.2	232	270.464	185.008	-0.2700	3.25	0.9920
			92.831	63.010	-0.3176	3.50	

Con base en los resultados expuestos en las columnas 4 a 7 de la Tabla 4, se estimaron con la Ecuación (8) las probabilidades de no excedencia conjunta teórica  $F(x,y)$ , utilizando los datos  $x_i$ ,  $y_i$  de la Tabla 2. Por otra parte, aplicando la Ecuación (55) y su procedimiento numérico descrito, se calcularon las llamadas probabilidades de no excedencia bivariadas empíricas  $F_e(x,y)$ , contra las cuales se contrastan los valores de la  $F(x,y)$  teórica estimada.

La mejor correspondencia o similitud entre ambas probabilidades conjuntas  $F_e(x,y)$  y  $F(x,y)$  se obtuvo para la cuarta aplicación o corrida del algoritmo Complex, con un valor del coeficiente de correlación ( $r_{xy}$ ) de 0.9954, y diferencias máximas positiva y negativa de 0.0482 y -0.0969, que se indican sombreadas en la Tabla 5.

**Tabla 5.** Probabilidades de no excedencia conjuntas y sus diferencias para las crecientes anuales de la estación hidrométrica La Cuña, México.

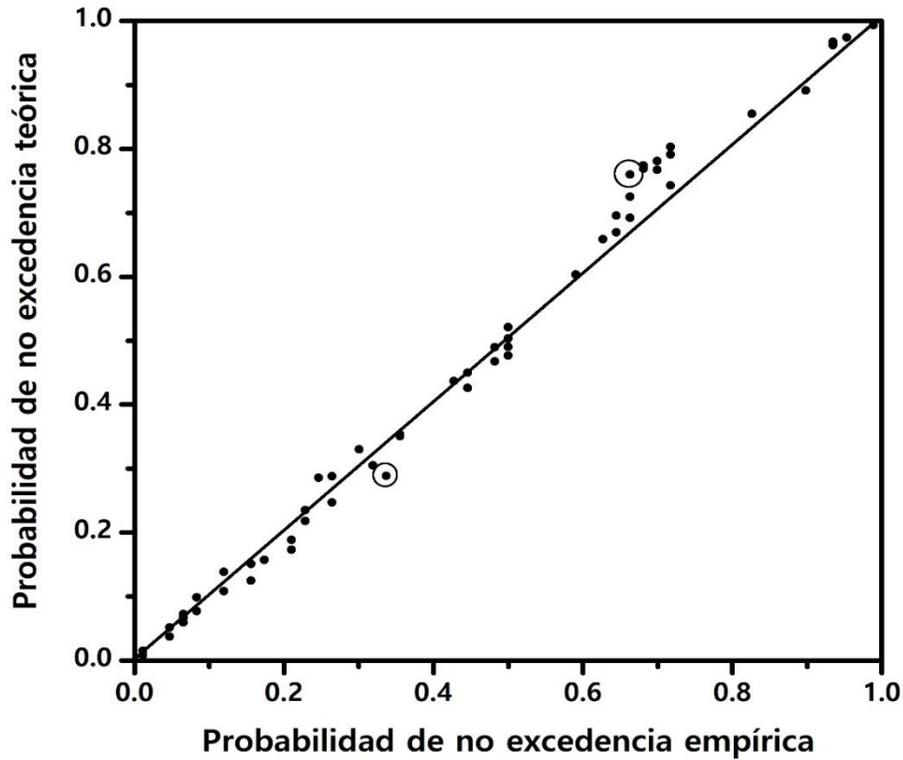
No.	$F_e(x,y)$ empírica	$F(x,y)$ teórica	Diferencias	No.	$F_e(x,y)$ empírica	$F(x,y)$ teórica	Diferencias
1	0.6449	0.6695	-0.0246	29	0.7174	0.7909	-0.0735
2	0.6630	0.6919	-0.0289	30	0.9348	0.9670	-0.0322

3	0.5000	0.5213	-0.0213	31	0.5000	0.5036	-0.0036
4	0.4275	0.4367	-0.0092	32	0.2645	0.2470	0.0175
5	0.4819	0.4895	-0.0076	33	0.1558	0.1513	0.0045
6	0.3370	0.2888	0.0482	34	0.3551	0.3504	0.0047
7	0.5000	0.4905	0.0095	35	0.0833	0.0770	0.0063
8	0.0471	0.0375	0.0096	36	0.0652	0.0727	-0.0075
9	0.6993	0.7809	-0.0816	37	0.6812	0.7688	-0.0876
10	0.3007	0.3303	-0.0296	38	0.1196	0.1387	-0.0191
11	0.1558	0.1247	0.0311	39	0.6812	0.7738	-0.0926
12	0.6630	0.7599	-0.0969	40	0.0471	0.0517	-0.0046
13	0.2101	0.1734	0.0367	41	0.6630	0.7253	-0.0623
14	0.2464	0.2860	-0.0396	42	0.9348	0.9618	-0.0270
15	0.5000	0.4768	0.0232	43	0.7174	0.7426	-0.0252
16	0.2101	0.1883	0.0218	44	0.3551	0.3535	0.0016
17	0.5906	0.6036	-0.0130	45	0.1739	0.1574	0.0165
18	0.2645	0.2882	-0.0237	46	0.4819	0.4674	0.0145
19	0.8261	0.8546	-0.0285	47	0.0833	0.0986	-0.0153
20	0.6268	0.6583	-0.0315	48	0.0109	0.0150	-0.0041
21	0.9529	0.9740	-0.0211	49	0.2283	0.2351	-0.0068
22	0.4457	0.4263	0.1964	50	0.1196	0.1082	0.0114
23	0.0652	0.0670	-0.0018	51	0.0109	0.0082	0.0027

24	0.6993	0.7671	-0.0678	52	0.3188	0.3048	0.0140
25	0.6449	0.6953	-0.0504	53	0.4457	0.4500	-0.0043
26	0.0652	0.0596	0.0056	54	0.7174	0.8029	-0.0855
27	0.9891	0.9933	-0.0042	55	0.8986	0.8911	0.0075
28	0.2283	0.2180	0.0103	-	-	-	-

## Validación del modelo probabilístico

En la Figura 2 se han dibujado ambas probabilidades de no excedencia conjuntas (empíricas y teóricas de la Tabla 5, observándose un predominio de diferencias negativas, es decir, de puntos por encima de la recta a 45°. El valor del estadístico del test de Kolmogorov-Smirnov es 0.1831 (Ecuación (56)), por lo cual, la distribución GVE bivariada es aceptada como modelo probabilístico conjunto de los datos de la Tabla 2, ya que  $dif = 0.0969 < D_n = 0.1831$ .



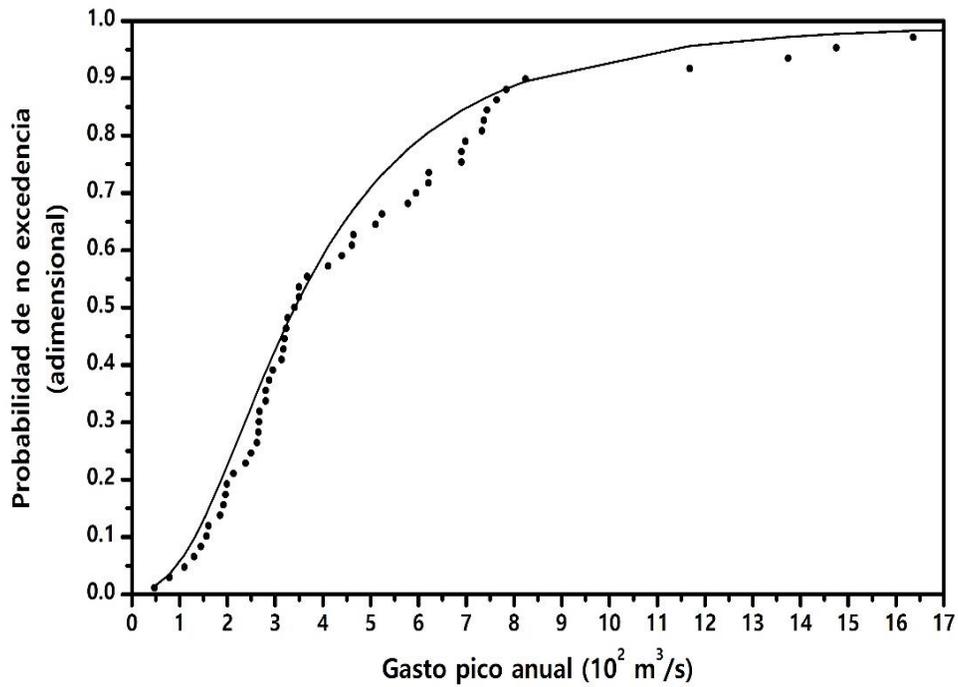
**Figura 2.** Contraste gráfico de probabilidades conjuntas estimadas con la distribución *GVEb* para las crecientes anuales de la estación hidrométrica La Cuña, México.

## Ratificación de las marginales

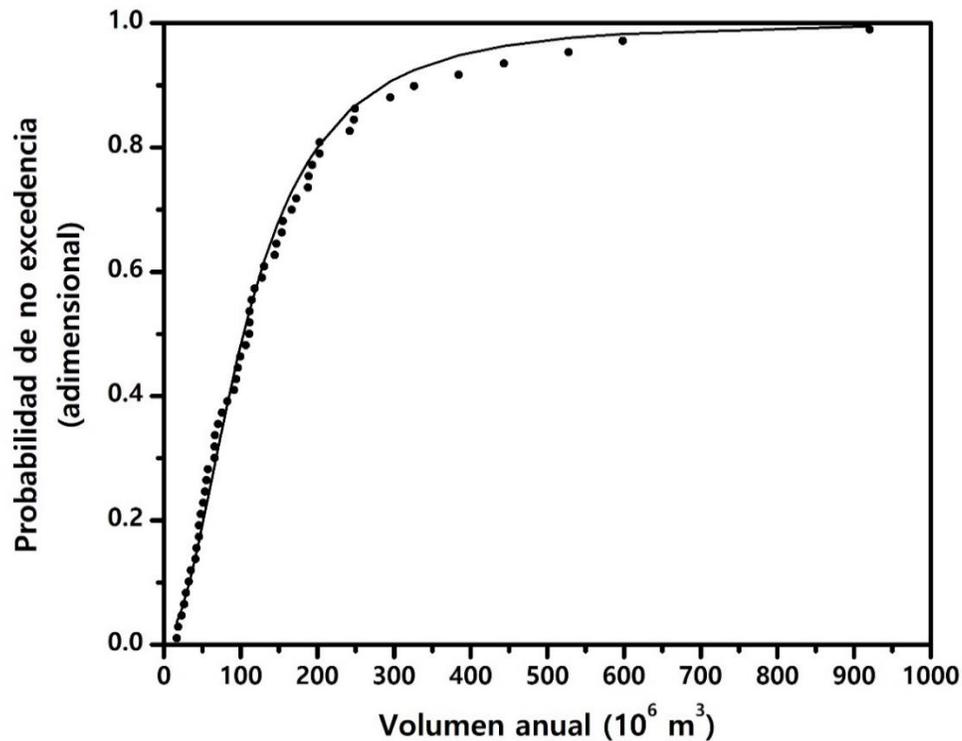
Este contraste gráfico y de ratificación con base en el test de Kolmogorov-Smirnov es necesario debido al cambio que tuvieron los parámetros de ajuste ( $u$ ,  $a$ ,  $k$ ) univariados (Tabla 3) durante la aplicación del algoritmo Complex. Es decir, permite definir si el cambio aceptado del 10 % en tales parámetros es aceptable o no.

Primeramente, los gastos pico y volúmenes de la Tabla 2 se ordenaron de menor a mayor. Después, se calcularon sus probabilidades de no excedencia *teóricas* con las ecuaciones (6) y (7), utilizando los parámetros de ajuste óptimos de la Tabla 4 para la cuarta corrida del algoritmo Complex. Las probabilidades de no excedencia *empíricas* de ambas series ( $Qp$  y  $V$ ) se estimaron con la Ecuación (54).

En la Figura 3 y Figura 4 se muestra el contraste gráfico de probabilidades para cada serie ordenada. Se observa un ajuste mucho mejor en el registro de volúmenes. Las diferencias máximas absolutas entre probabilidades empíricas y teóricas de los gastos pico y volúmenes fueron 0.0954 y 0.0457; la primera ocurrió en el dato ordenado número 38 y la segunda en el 16, como se muestra en la Tabla 6 de resultados parciales. Como ambas diferencias son menores que  $D_n = 0.1831$ , obtenido con la Ecuación (56), se acepta que los registros de  $Qp$  y  $V$  de la Tabla 2 pueden ser representados por la distribución GVE con los parámetros de ajuste de la Tabla 4 en la cuarta corrida del algoritmo Complex.



**Figura 3.** Distribución marginal *GVEb* del gasto pico anual de las crecientes de la estación hidrométrica La Cuña, México.



**Figura 4.** Distribución marginal *GVEb* del volumen anual de las crecientes de la estación hidrométrica La Cuña, México.

**Tabla 6.** Probabilidades empíricas ( $F_e$ ) y teóricas en las distribuciones marginales *GVEb*, y sus diferencias en las crecientes de la estación hidrométrica La Cuña, México.

No.	$Q_p$	$V$	$F_e$	$F(Q_p)$	$F(V)$	Dif $Q_p$	Dif $V$
1	46.8	16.55	0.0109	0.0135	0.0368	-0.0027	-0.0259
8	184.7	41.15	0.1377	0.1948	0.1418	-0.0571	-0.0042

16	265.1	57.43	0.2826	0.3581	0.2369	-0.0755	0.0457
27	326.0	106.76	0.4819	0.4735	0.5152	0.0084	-0.0333
38	578.0	155.12	0.6812	0.7765	0.6974	-0.0954	-0.0162
45	732.6	203.31	0.8080	0.8624	0.8051	-0.0544	0.0028
55	2440.0	920.30	0.9891	0.9949	0.9945	-0.0058	-0.0054

## Verificación de la restricción de probabilidades

Antes de proceder a estimar los periodos de retorno conjuntos de diseño  $T'(Q,V)$  es conveniente verificar la Ecuación (9), que establece las restricciones de la probabilidad. Lo anterior se muestra en la Tabla 7 para un número reducido de parejas de datos históricos.

**Tabla 7.** Verificación de la restricción de probabilidades del gasto pico y volumen anuales de las crecientes de la estación hidrométrica La Cuña, México.

1	2	3	4	5	6	7
Núm.	$Q_p$	$V$	$F(x)$	$F(y)$	$F(x) \cdot F(y)$	$F(x, y)$

1	784.0	146.80	0.8817	0.6723	0.5928	0.6695
10	266.0	106.76	0.3599	0.5152	0.1854	0.3303
20	463.9	202.90	0.6720	0.8044	0.5406	0.6583
30	1374.0	527.96	0.9718	0.9759	0.9484	0.9670
40	110.2	42.49	0.0690	0.1492	0.0103	0.0517
50	196.3	41.15	0.2178	0.1418	0.0309	0.1082
55	824.5	384.45	0.8947	0.9479	0.8481	0.8911

Se observa en la Tabla 7 anterior y en su versión completa, que el valor de la columna 7 es siempre mayor que el de la 6 y menor que el más pequeño de las columnas 4 o 5, por lo cual se cumple con la restricción que impone la Ecuación (9).

## Predicciones univariadas

Aplicando en la Ecuación (2) los seis parámetros de ajuste óptimos de la cuarta corrida del algoritmo Complex, mostrados en la Tabla 4, se obtuvieron las predicciones finales mostradas en la Tabla 8.

**Tabla 8.** Predicciones univariadas finales calculadas con la distribución *GVEb* en la estación hidrométrica La Cuña, México.

Datos	Periodo de retorno en años						
	25	50	100	500	1 000	5 000	10 000
$Q_p$	1 210	1 550	1 958	3 254	4 009	6 418	7 823
$V$	430	568	739	1 317	1 673	2 877	3 619

## Eventos de diseño para obras hidráulicas

Suponiendo que en las cercanías de aguas debajo de la estación hidrométrica La Cuña y sobre el Río Verde se van a construir diques para protección de planicies de inundación y un puente para cruzarlo, entonces es necesario estimar eventos de diseño con periodos de retorno conjuntos de 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. Entonces se deben estimar gastos pico y volúmenes anuales con los cuatro periodos de retorno conjuntos  $T'(Q, V)$  citados.

## Eventos de diseño obtenidos con regresión

El diagrama de dispersión de las 55 parejas de datos originales (Tabla 2) mostró una nube de puntos con tendencia lineal, con un coeficiente de correlación lineal ( $r_{xy}$ ) de 0.9302. La ecuación de regresión lineal que la representa es la siguiente (Campos-Aranda, 2003):

$$V = -20.0532 + 0.34987 \cdot Qp \quad (57)$$

De la Tabla 8 se obtienen las predicciones siguientes para el gasto pico ( $Qp$ ) y los cuatro periodos de retorno conjuntos de diseño: 3 254, 4 009, 6 418 y 7 823  $m^3/s$ , respectivamente. Con base en la Ecuación (57) se definen los volúmenes anuales siguientes: 1 118, 1 383, 2 225 y 2 717  $Mm^3$  para las crecientes de diseño conjuntas.

Al igual que Serinaldi y Grimaldi (2011), se encuentra una similitud entre los volúmenes estimados con regresión y sus predicciones de la Tabla 8, que son 1 317, 1 673, 2 877 y 3 619  $Mm^3$ ; es decir, del orden de un 18 a un 33 % mayores.

## Eventos de diseño condicionales tipo $T(V|Q)$

Están definidos por las ecuaciones (45) y (47), cuya aplicación emplea las expresiones (6) y (8). Para los cuatro periodos de retorno conjuntos de diseño definidos, se obtienen de la Tabla 8 los siguientes cuatro gastos pico: 3 254, 4 009, 6 418 y 7 823 m<sup>3</sup>/s. Adoptando tales gastos como valores *condicionantes* ( $X \leq x$ ), se procedió por tanteos del volumen ( $y$ ) a estimar, con la Ecuación (47), el periodo de retorno condicional que debe igualar al del gasto pico. Los volúmenes estimados fueron 1 049, 1 338, 2 317 y 2 920 Mm<sup>3</sup>.

## Gráficas del periodo de retorno conjunto $T'(Q,V)$

Los periodos de retorno conjuntos de tipo AND se estiman con base en la Ecuación (41). Definidos los cuatro periodos de retorno conjuntos de diseño, se seleccionan arbitrariamente gastos pico y volúmenes, para obtener sus probabilidades de no excedencia marginales y conjunta. Las primeras se estiman con las ecuaciones (6) y (7), y la segunda con la Ecuación (8). En la Tabla 9 se muestran las parejas de gasto pico y

volumen anual utilizadas para definir las cuatro gráficas de la Figura 1 relativas a los periodos de retorno conjuntos de diseño de la aplicación numérica descrita.

**Tabla 9.** Parejas de gasto pico y volumen anual utilizadas para definir las gráficas del periodo de retorno conjunto (Figura 1), con la distribución *GVEb* en las crecientes de la estación hidrométrica La Cuña, México.

$T'(Q,V)$ 500 años		$T'(Q,V)$ 1 000 años		$T'(Q,V)$ 5 000 años		$T'(Q,V)$ 10 000 años	
Vol. Mm <sup>3</sup>	Qp m <sup>3</sup> /s	Vol. Mm <sup>3</sup>	Qp m <sup>3</sup> /s	Vol. Mm <sup>3</sup>	Qp m <sup>3</sup> /s	Vol. Mm <sup>3</sup>	Qp m <sup>3</sup> /s
0	3 254	0	4 009	0	6 418	0	7 823
300	3 254	400	4 009	800	6 418	1 000	7 823
600	3 249	600	4 008	1 000	6 417	1 200	7 822
800	3 233	800	4 003	1 200	6 415	1 400	7 821
900	3 212	1 000	3 988	1 400	6 409	1 600	7 820
1 000	3 173	1 100	3 971	1 600	6 399	1 800	7 814
1 050	3 142	1 200	3 944	1 800	6 379	2 000	7 802
1 100	3 098	1 250	3 925	2 000	6 342	2 200	7 787
1 150	3 034	1 300	3 900	2 100	6 313	2 400	7 759

1 200	2 933	1 350	3 867	2 200	6 275	2 600	7 714
1 250	2 756	1 400	3 823	2 300	6 222	2 800	7 641
1 275	2 598	1 450	3 765	2 400	6 149	2 900	7 588
1 300	2 289	1 500	3 682	2 500	6 046	3 000	7 524
1 310	2 000	1 550	3 559	2 550	5 976	3 100	7 429
1 317	0	1 600	3 354	2 600	5 890	3 200	7 310
		1 625	3 181	2 650	5 781	3 300	7 140
		1 650	2 875	2 700	5 637	3 400	6 874
		1 673	0	2 750	5 431	3 450	6 685
				2 800	5 115	3 500	6 422
				2 825	4 863	3 525	6 322
				2 850	4 445	3 550	6 009
				2 877	0	3 575	5 667
						3 600	5 075
						3 619	0

Por otra parte, en la Tabla 10 se muestran los gastos pico y volúmenes anuales definidos con los periodos de retorno univariados híbridos. Tales parejas de  $Q_p$  y  $V$  también se han dibujado en la Figura 1, y se indican con las letras "r" y "c".

**Tabla 10.** Eventos de diseño obtenidos con los periodos de retorno univariados *híbridos* en las crecientes de la estación hidrométrica La Cuña, México.

Evento de diseño	Periodo de retorno conjunto de diseño en años							
	500		1 000		5 000		10 000	
	<i>Qp</i>	<i>V</i>	<i>Qp</i>	<i>V</i>	<i>Qp</i>	<i>V</i>	<i>Qp</i>	<i>V</i>
Con regresión (r)	3	1	4	1	6	2	7	2
	254	118	009	383	418	224	823	717
Condicional (c)	3	1	4	1	6	2	7	2
	254	049	009	338	418	317	823	920

En la Figura 1 o en la Tabla 9 se pueden seleccionar infinitas parejas de  $Q_p$  y  $V$  que satisfacen el periodo de retorno conjunto de diseño, y que definen como *subgrupo de parejas críticas* a las que están dentro de la porción curva de cada gráfica de  $T'(Q,V)$ , fuera de las rectas asíntotas (Volpi & Fiori, 2012).

Las combinaciones de gasto pico y volumen que tienen el mismo *periodo de retorno conjunto* definen crecientes o *hidrogramas* que producirán diferentes efectos en el embalse que se diseña o revisa, adoptando por seguridad, el que genera las condiciones más críticas, severas o desfavorables. Lo anterior está incorporando en el diseño

hidrológico, las características físicas del vertedor y almacenamiento o vaso del embalse en proyecto o bajo revisión.

Para formar cada hidrograma de diseño existen métodos teóricos y empíricos (Aldama, 2000; Aldama *et al.*, 2006; Ramírez-Orozco & Aldama, 2000; Serinaldi & Grimaldi, 2011). Campos-Aranda (2008) ha expuesto un proceso empírico que define hidrogramas tipo Gamma esbeltos y aplanados.

## Otras distribuciones bivariadas

Los primeros análisis de frecuencias bivariados utilizaron la distribución Normal bivariada y se aplicaron a las crecientes anuales (Yue, 1999) y a las tormentas de lluvias diarias (Yue, 2000a). Posteriormente, se aplicó la versión bivariada Log-normal (Yue, 2000c), y también se generalizó el uso del llamado *modelo Logístico* (Ecuación (3)), que acepta distribuciones marginales Gumbel y tiene solución explícita (Yue *et al.*, 1999; Yue, 2000b). El modelo logístico adquirió popularidad por aceptar como funciones marginales *iguales* las de valores extremos: GVE, Gumbel mixta, TCEV y GVE mixta. Actualmente se conoce como *función cópula* Gumbel-Hougaard y acepta funciones marginales *diferentes* (Shiau *et al.*,

2006), existiendo la posibilidad de procesar distribuciones marginales con una y/o dos poblaciones.

## Conclusiones

Se propuso y aplicó un enfoque combinado de maximización de la función logarítmica de verosimilitud de la *distribución GVE bivariada (GVEb)*, que intenta preservar los ajustes univariados de las funciones marginales, dejando variable el parámetro de asociación ( $m$ ), para buscar la mejor correspondencia o similitud entre las probabilidades de no excedencia observadas y las teóricas bivariadas. En este estudio se permitió una variación del 10 % en los valores de los parámetros de ajuste ( $u_1, a_1, k_1, u_2, a_2, k_2$ ) de las funciones marginales.

Se expuso la aplicación de la *GVEb* en el análisis de frecuencias conjunto de los 55 gastos pico y volúmenes anuales de las crecientes registradas en la estación hidrométrica La Cuña, México. El ajuste de la *GVEb* permite el cálculo de las probabilidades univariadas, conjuntas y condicionales. Esta parte del proceso concluye con el contraste de las funciones marginales y la validación de la *GVEb*, previa estimación numérica de las probabilidades empíricas univariadas y bivariadas.

Después se abordó el cálculo de los periodos de retorno univariados híbridos y conjuntos. Los primeros utilizan una variable de diseño dominante, por ejemplo, el gasto pico, y definen por regresión el volumen asociado o lo estiman mediante probabilidad condicional. Las parejas únicas de gasto pico y volumen estimadas se muestran en la Figura 1 del periodo de retorno conjunto de tipo AND.

En el AFC con la *GVEb* se definen decenas de *hidrogramas críticos* que producirán diferentes efectos en el embalse que se diseña o revisa, adoptando por seguridad el que genera las condiciones más severas o adversas. De esta manera se están incorporando en el diseño hidrológico las características físicas del embalse bajo estudio o en revisión.

La *GVEb* es adecuada para procesar de manera conjunta registros de gasto pico y volúmenes de crecientes que presentan valores extremos dispersos (*outliers*), pero que no están conformados por poblaciones mezcladas. Por lo anterior, la *GVEb* permite procesar crecientes de cuencas medianas y grandes de zonas o regiones con mecanismos meteorológicos únicos de formación de las crecientes.

## Agradecimientos

Se agradecen las correcciones sugeridas por los árbitros anónimos A y C, las cuales permitieron enmendar errores de redacción e incluir dos explicaciones relativas a los datos procesados. Además, ayudaron a ubicar el trabajo en el contexto bivariado, de sus antecedentes y de sus análisis futuros.



## Referencias

- Aldama, A. A. (2000). Hidrología de avenidas. Conferencia Enzo Levi 1998. *Ingeniería Hidráulica en México*, 15(3), 5-46.
- 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.
- 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 optimization* (pp. 98-106). London, UK: Edward Arnold Publishers, Ltd.
- Campos-Aranda, D. F. (2003). Capítulo 5: Ajuste de curvas y Capítulo 9: Optimización numérica. En: *Introducción a los métodos numéricos: software en Basic y aplicaciones en hidrología superficial* (pp. 93-127, 72-211). San Luis Potosí, México: Editorial Universitaria Potosina.

- Campos-Aranda, D. F. (2008). Procedimiento para revisión (sin hidrometría) de la seguridad hidrológica de presas pequeñas para riego. *Agrociencia*, 42(5), 551-563.
- Coles, S. (2001). Chapter 2: Basics of statistical modeling and chapter 3: Classical extreme value theory and models. In: *An introduction to statistical modeling of extreme values* (pp. 18-44, 45-73). London, UKA: Springer-Verlag.
- Chow, V. T. (1964). Statistical and probability analysis of hydrologic data. Section 8–I: Frequency Analysis. In: Chow, V. T. (ed.). *Handbook of Applied Hydrology* (pp. 8.1-8.42). New York, USA: McGraw-Hill Book Co.
- Davis, P. J. (1972). Gamma function and related functions. Chapter 6 (pp 253–296). In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions*. New York, USA: Dover Publications.
- 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. A., & Reyes-Chávez, L. (2002). Capítulo 9: Análisis conjunto de eventos hidrológicos. En: *Técnicas estadísticas en hidrología* (pp. 203-246). México, DF, México: Facultad de Ingeniería, Universidad Nacional Autónoma de México.

- Escalante-Sandoval, C. A., & Reyes-Chávez, L. (2004). Análisis bivariado de gastos máximos anuales con distribuciones marginales TCEV. *XVIII Congreso Nacional de Hidráulica* (pp. 523-529), 10-12 de noviembre, San Luis Potosí, México.
- Escalante-Sandoval, C. A. (2005). Análisis de eficiencia de la distribución Bi-Gumbel. *Ingeniería. Investigación y Tecnología*, 6(1), 13-17. DOI: 10.22201/fi.25940732e.2005.06.n1.002
- 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.
- Greenwood, J. A., Landwehr, J. M., Matalas, N. C., & Wallis, J. R. (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*, 15(5), 1049-1054. DOI: 10.1029/WR015i005p01049
- 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.

- Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P., & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 329(3-4), 534-552. DOI: 10.1016/j.jhydrol.2006.03.004
- Kite, G. W. (1977). Chapter 3: Frequency distributions. In: *Frequency and risk analyses in Hydrology* (pp. 27-39). Fort Collins, USA: Water Resources Publications.
- Kottegoda, N. T., & Rosso, R. (2008). Chapter 3: Random variables and their properties. In: *Applied Statistics for Civil and Environmental Engineers* (pp. 83-164). Oxford, UK: Blackwell Publishing Ltd.
- Metcalfe, A. V. (1997). Chapter 2: Probability distributions and Monte Carlo simulations. *Statistics in Civil Engineering* (pp. 7-38). London, UK: Arnold Publishers.
- Meylan, P., Favre, A. C., & Musy, A. (2012). Chapter 3: Selecting and checking data series and Chapter 5: Estimation of model parameters. In: *Predictive Hydrology. A Frequency Analysis Approach* (pp. 9-102, 29-70). Boca Raton, Florida, USA: CRC Press.
- Mujumdar, P. P., & Nagesh-Kumar, D. N. (2012). *Floods in a changing climate. Hydrologic Modeling*. Cambridge, UK. International Hydrology Series (UNESCO) and Cambridge University Press.
- Ramírez-Orozco, A. I., & Aldama, A. A. (2000). Capítulo 1: Teoría estadística y análisis de frecuencias conjunto. En: *Análisis de frecuencias conjunto para estimación de avenidas de diseño* (pp.

- 25-58). Avances en Hidráulica No. 7. México, DF, México: Asociación Mexicana de Hidráulica, Instituto Mexicano de Tecnología del Agua.
- Rao, A. R., & Hamed, K. H. (2000). Theme 1.8: Tests on hydrologic data and Chapter 7: Extreme Value Distributions (pp. 12-21, 207-257). In: *Flood frequency analysis*. 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
- Serinaldi, F., & Grimaldi, S. (2011). Synthetic design hydrographs based on distribution functions with finite support. *Journal of Hydrologic Engineering*, 16(5), 434-446. DOI: 10.1061/(ASCE)HE.1943-5584.0000339
- Shiau, J. T. (2003). Return period of bivariate distributed extreme hydrological events. *Stochastic Environmental Research and Risk Assessment*, 17(1-2), 42-57. DOI: 10.1007/s00477-003-0125-9
- 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

- Stedinger, J. R. (2017). Flood frequency analysis. In: Singh, V. P. (ed.). *Handbook of Applied Hydrology* (2<sup>nd</sup> ed.). (pp. 76.1–76.8). New York, USA: McGraw-Hill Education.
- Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Chapter 18: Frequency Analysis of Extreme Events. In: Maidment, D. R. (ed.). *Handbook of Hydrology* (pp. 18.1–18.66). New York, USA: McGraw-Hill, Inc.
- Teegavarapu, R. S. V. (2012). *Floods in a changing climate. Extreme precipitation*. Cambridge, UK: International Hydrology Series (UNESCO) and Cambridge University Press.
- Teegavarapu, R. S. V., Salas, J. D., & Stedinger, J. R. (2019). *Statistical analysis of hydrologic variables*. Reston, USA: American Society of Civil Engineers.
- Vogel, R. M., & Castellarin, A. (2017). Risk, reliability, and return periods and hydrologic design. In: Singh, V. P. (ed.). *Handbook of applied hydrology* (2<sup>nd</sup> ed.) (pp. 78.1-78.10). New York, USA: McGraw-Hill Education.
- Volpi, E., & Fiori, A. (2012). Design event selection in bivariate hydrological frequency analysis. *Hydrological Sciences Journal*, 57(8), 1506-1515. DOI: 10.1080/02626667.2012.726357
- Yue, S. (1999). Applying bivariate Normal distribution to flood frequency analysis. *Water International*, 24(3), 248-254.

- 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. (2000a). 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. (2000b). The Gumbel mixed model applied to storm frequency analysis. *Water Resources Management*, 14(5), 377-389.
- Yue, S. (2000c). The bivariate lognormal distribution to model a multivariate flood episode. *Hydrological Processes*, 14(14), 2575-2588.
- 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
- Yue, S., & Wang, C. Y. (2004). A comparison of two bivariate extreme value distributions. *Stochastic Environmental Research and Risk Assessment*, 18(2), 61-66. DOI: 10.1007/s00477-003-0124-x