DOI: 10.24850/j-tyca-2022-03-03

Artículos

Análisis regional de frecuencias de crecientes con base en la distribución TCEV en la Región Hidrológica No. 10 (Sinaloa), México

Regional flood frequencies analysis based on the TCEV distribution in Hydrological Region No. 10 (Sinaloa),

Mexico

Daniel Francisco Campos-Aranda<sup>1</sup>

<sup>1</sup>Profesor jubilado de la Universidad Autónoma de San Luis Potosí, San Luis Potosí, México, campos aranda@hotmail.com

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

#### Resumen

Las obras hidráulicas como embalses, diques, puentes y el drenaje urbano se planean, diseñan, construyen y revisan (las existentes) bajo una seguridad hidrológica, con base en las *Crecientes de Diseño*. Tales



estimaciones del gasto máximo anual asociadas a bajas probabilidades de excedencia, se llaman *Predicciones* y se obtienen a través del llamado análisis de frecuencia de crecientes. Técnica que consiste en seleccionar una función de distribución de probabilidades (FDP), que represente a la muestra de gastos disponible y con base en tal modelo, realizar las predicciones buscadas. Cuando el registro de crecientes es reducido, poco confiable o no existe en el sitio de interés, se recurre al análisis regional; enfoque que amplía la longitud de los registros disponibles para reducir los errores en las estimaciones y además, permite estimar predicciones en localidades sin datos. En este estudio se realizó un contraste, entre las predicciones locales y las regionales; obtenidas estas últimas, según dos métodos diferentes en la Región Hidrológica No. 10 (Sinaloa), México. En el primero, se procesaron sus registros de gasto máximo anual y sus respectivas fechas de ocurrencia, para definir las regiones de influencia de cada uno y en la segunda, la cual se expone con detalle, se utiliza la misma información hidrométrica y se aplica el método del índice de crecientes, utilizando como FDP regional a la distribución TCEV (twocomponent extreme value), sugerida para registros de gastos que fueron generados por dos mecanismos físicamente diferentes y que no es posible separarlos. El método regional expuesto es mucho más simple y conduce a predicciones bastante aproximadas por déficit y más exactas por exceso, según los errores relativos evaluados. Por lo anterior, se recomienda su aplicación sistemática, en otras regiones hidrológicas del país.



**Palabras clave**: distribución TCEV, índice de estacionalidad, índice de crecientes, detección de valores dispersos, prueba del rango regional normalizado, predicciones locales y regionales, error relativo, creciente media anual.

#### **Abstract**

Hydraulic works such as reservoirs, dikes, bridges, and urban drainage, are planned, designed, built, and reviewed (the existing) under a hydrological safety, based on the *Design Floods*. Such estimates of the maximum annual flow associated with low probabilities of exceedance, are called *Predictions and* are obtained through the so-called *frequency* analysis of floods. This technique consists of selecting a probability distribution function (PDF), which represents the sample of available flows, and based on such a model prediction is sought. When the record of floods is reduced, unreliable, or does not exist at the site of interest, regional analysis is used; an approach that extends the length of available records to reduce errors in estimates and allows estimating predictions in locations without data. In this study a *contrast* was made between local and regional predictions; the latter was obtained according to two different methods in the Hydrological Region No. 10 (Sinaloa), Mexico. In the first method, the records of maximum annual flows and their respective dates of occurrence were processed, to define the regions of influence each. For the second method, which is detailed, the same hydrometric information is used and the flood index method, using the



TCEV (two-component extreme value) distribution as a regional PDF, is suggested for flow records that were generated by two physically different mechanisms and cannot be separated. The exposed regional method is much simpler and leads to fairly approximate predictions by deficit and more accurate by excess, according to the relative errors evaluated. Therefore, its systematic application is recommended in other hydrological regions of the country.

**Keywords**: TCEV distribution, seasonality index, flood index, detection of outlier's values, normalized regional range test, local and regional predictions, relative error, average annual flood.

Recibido: 07/11/2019

Aceptado: 30/03/2021

### Introducción

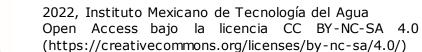
### Análisis de frecuencia de crecientes



Todas las obras hidráulicas que requiere la sociedad humana para el uso razonable de los recursos hídricos y para la protección contra sus excesos y deficiencias se planean, diseñan y construyen con base en las llamadas *crecientes de diseño*. Además, las obras hidráulicas que ya están operando, periódicamente deben ser revisadas desde un punto de vista hidrológico, pues es común que ocurran cambios físicos en sus cuencas y en las condiciones meteorológicas debido a los impactos del cambio climático (Mujumdar & Nagesh-Kumar, 2012).

Las crecientes de diseño son estimaciones del gasto máximo anual asociadas con bajas probabilidades de excedencia, cuyo recíproco es el periodo de retorno (*Tr*) o intervalo *promedio* de recurrencia en años. Las estimaciones más exactas de tales *predicciones* se obtienen a través del llamado *análisis de frecuencia de crecientes* (AFC), cuyo objetivo consiste en interpretar una muestra de gastos máximos anuales en términos de probabilidades de ocurrencia futura (Bobée & Ashkar, 1991; Hosking & Wallis, 1997).

Para que los resultados del AFC sean confiables, se deben cumplir los tres requisitos siguientes: (1) que los gastos máximos anuales de la muestra hayan sido generados por un proceso aleatorio estacionario, lo cual implica que sean independientes y que estén libres de componentes determinísticas; (2) que el modelo probabilístico o función de distribución de probabilidades (FDP) utilizada para representar a los datos y realizar las estimaciones sea la adecuada, y (3) que el método de estimación de





los parámetros de ajuste de la FDP sea exacto (Meylan, Favre, & Musy, 2012).

Para cumplir el primer requisito se verifica la calidad estadística de la muestra por medio de pruebas generales y específicas, como el test de Wald-Wolfowitz; y las pruebas de persistencia, tendencia y de cambio en la media o en la variabilidad (WMO, 1971; Bobée & Ashkar, 1991; Rao & Hamed, 2000; Machiwal & Jha, 2012). Respecto al segundo requisito existen dos enfoques: (a) usar las FDP que han sido establecidas bajo precepto y (b) emplear las FDP que tienen base teórica (parent distribution) para ser aceptadas como modelo fuente o de origen (Houghton, 1978). Para abordar el tercer requisito se pueden aplicar los métodos disponibles de estimación de parámetros de ajuste y seleccionar resultados con base en indicadores cuantitativos de calidad del ajuste, por ejemplo, el error estándar de ajuste (Kite, 1977; Meylan *et al.*, 2012).

# Análisis regional de frecuencia de crecientes

Con objeto de reducir la incertidumbre que se genera al emplear muestras pequeñas en el AFC tanto en la estimación de los parámetros de ajuste de la FDP como de las predicciones, se recurre al análisis *regional* de



frecuencia de crecientes (ARFC). Este procedimiento o *regionalización* es una herramienta que permite ampliar la longitud de los registros históricos disponibles y por ello reduce los errores en las estimaciones. Además, hace posible la estimación de predicciones en sitios o localidades dentro de la región que no cuentan con datos de gasto máximo anual. En el ARFC se juntan o concatenan los registros estandarizados de gasto máximo anual de varios sitios, para formar series amplias, siempre y cuando muestren una similitud u homogeneidad hidrológica (Cunnane, 1988; Hosking & Wallis, 1997; Rahman, Haddad, & Eslamian, 2014; Ouarda, 2017).

#### La distribución TCEV

El modelo probabilístico TCEV (*two-component extreme value*), sugerido por Rossi, Fiorentino y Versace (1984), permite una interpretación razonable de los fenómenos físicos que generan las crecientes y además es capaz de reproducir las características reales observadas en los registros de gastos máximos anuales, como son el llamado por Potter (1958) "efecto de pata de perro" (*dog leg effect*) y la designada por Matalas, Slack y Wallis (1975) como "condición de separación".



La distribución TCEV considera que las crecientes de un registro fueron generadas por dos mecanismos físicamente diferentes, las cuales no es posible separar. Al ser una FDP con cuatro parámetros de ajuste, su aplicación a registros individuales genera gran incertidumbre en el cálculo de las predicciones (Fiorentino, Versace, & Rossi, 1985). Por ello, Francés-García (1995) ha señalado que el campo habitual de aplicación del modelo TCEV es dentro del ARFC.

### Índice de estacionalidad

Desde mediados de la década de 1990 (Zrinji & Burn, 1996; Burn, 1997; Cunderlik & Burn, 2002) se ha sugerido que la fecha de ocurrencia media y la regularidad estacional de las crecientes anuales se pueden utilizar como medidas de similitud de la respuesta hidrológica de las cuencas. Es un hecho que la similitud en la distribución temporal y la regularidad de las crecientes de dos cuencas implica una semejanza en características fisiográficas y meteorológicas. De manera que tales cuencas se pueden integrar en una región con propósitos del ARFC. Este enfoque de regionalización, con base en las características estacionales de las crecientes, tiene la ventaja de reservar el uso de la información





hidrométrica (crecientes anuales) para ratificar tal homogeneidad regional.

Campos-Aranda (2014a) expone con detalle cómo se transforma la fecha de ocurrencia de cada creciente anual en un estadístico direccional para obtener el día medio de las crecientes, cuya medida adimensional de su dispersión es el llamado *índice de estacionalidad* ( $\bar{r}$ ), el cual toma valores entre cero y uno. Un valor unitario indica que todas las crecientes ocurren en la misma fecha; en cambio, un valor cercano a cero implica ocurrencias a lo largo de todo el año.

Ramírez-Orozco, Gutiérrez-López y Ruiz-Silva (2009) establecen los siguientes cinco grados de estacionalidad: (1) muy fuerte, cuando  $\bar{r} > 0.90$ ; (2) fuerte, cuando  $\bar{r}$  fluctúa entre 0.70 y 0.90; (3) media, cuando  $\bar{r}$  varía de 0.50 a 0.70; (4) baja, cuando  $\bar{r}$  cambia de 0.10 a 0.50, y (5) muy baja o débil, cuando  $\bar{r} < 0.10$ .

# **Objetivo**

El objetivo básico de este estudio consistió en realizar un *contraste* entre las predicciones locales y las regionales; estas últimas obtenidas según dos técnicas diferentes en la Región Hidrológica No. 10 (Sinaloa), México.



En la primera se procesaron sus registros de gasto máximo anual y sus respectivas fechas de ocurrencia, por medio del enfoque regional de la región de influencia (Campos-Aranda, 2014a). En la segunda, la cual se describe con detalle, se utiliza la misma información hidrométrica y se aplica el método clásico de los ARFC, es decir, el método del índice de crecientes, empleando como modelo probabilístico poblacional la distribución TCEV.

# Teoría operativa y datos procesados

#### Procedimiento del método del índice de crecientes

Las técnicas que integran el llamado *método del índice de crecientes* (*index flood method*) constituyen una manera conveniente de obtener un resumen de parámetros estadísticos procedentes de la agrupación de varias muestras o registros diferentes (Hosking & Wallis, 1997). Fue propuesto por Dalrymple (1960) para procesar datos de crecientes,



también llamadas avenidas máximas o riadas, y por ello su nombre, pero se puede aplicar a cualquier tipo de datos.

Cuando existen registros de crecientes en ns sitios y en cada localidad k, el número de datos es  $n_i$  de gastos  $Q_i^k$ . Entonces, se tiene que  $Q^k(F)$  con 0 < F < 1 es la función de cuantiles de la distribución de frecuencias de cada sitio k. La hipótesis básica del método de índice de crecientes establece que los sitios k forman una  $región\ homogénea$ , lo cual implica que las distribuciones de frecuencias de los ns sitios son idénticas, excepto por el factor de escalamiento, que es precisamente el  $indice\ de\ crecientes\ (IC^k)$ . La ecuación del método es:

$$Q^{k}(F) = IC^{k} \cdot q(F) \operatorname{con} k = 1, 2, \dots, ns \tag{1}$$

donde q(F) es la curva de crecimiento regional (regional growth curve), es decir, una función de cuantiles adimensional y común a todos los sitios; q(F) es una distribución de frecuencias regional o función común de los valores  $Q_i^k/IC^k$ ; por lo general, el  $IC^k$  es igual al gasto medio anual  $(\bar{Q}^k)$  de los datos de cada sitio, pero se puede emplear la mediana o una media truncada (Hosking & Wallis, 1997).

En el método de índice de crecientes se aceptan las cinco hipótesis siguientes: (1) las crecientes de cada sitio se distribuyen de manera idéntica; (2) las crecientes de cada sitio son independientes; (3) las crecientes de sitios diferentes son independientes; (4) las distribuciones



de frecuencias en sitios diferentes son idénticas, excepto por el factor de escalamiento, y (5) la forma matemática de la curva regional de crecimiento está correctamente especificada (Hosking & Wallis, 1997).

Todos los métodos del ARFC —incluido el del índice de crecientes— tienen dos aplicaciones fundamentales dentro de la región homogénea: la primera se realiza en sitios con registros cortos o no confiables, y la segunda en localidades sin datos hidrométricos, utilizando las ecuaciones previamente encontradas, para la relación entre el  $Q^k$  y el área de cuenca, u otras características fisiográficas de la cuenca y/o del cauce principal (Campos-Aranda, 2013).

#### Fórmulas de la distribución TCEV

La base teórica de la distribución TCEV la establecen Rossi  $et\ al$ . (1984), considerando que la variable aleatoria X es el máximo anual de una variable aleatoria Z no negativa, cuyo número K de ocurrencias en el año también es una variable aleatoria, y ambas cumplen las hipótesis siguientes: (1) Z es una variable aleatoria independiente idénticamente distribuida, condición designada "iid"; (2) K es una variable "iid" que sigue el proceso de Poisson, y (3) Z y K son independientes una de otra.



Por otra parte, el efecto pata de perro de un registro de crecientes se observa al dibujar en el papel de probabilidad Gumbel-Powell (Chow, 1964) sus datos, y encontrar que un grupo numeroso de gastos definen una tendencia y en un punto final; otro grupo reducido de crecientes de mayor magnitud comienzan a definir otra tendencia con mayor pendiente. Lo anterior indica la presencia de dos poblaciones: una básica con el subíndice 1 y otra de crecientes extraordinarias con el subíndice 2.

Al considerar que las crecientes de ciertos registros han sido generadas por dos mecanismos meteorológicos diferentes, como son las tormentas convectivas locales y las tormentas de origen ciclónico, Rossi *et al.* (1984) formulan un modelo probabilístico mixto, con lo cual establecen la distribución TCEV con cuatro parámetros de ajuste:

$$F_X(x) = \exp\left(-\lambda_1 \cdot e^{-x/\theta_1} - \lambda_2 \cdot e^{-x/\theta_2}\right) \tag{2}$$

haciendo  $\varepsilon_j = \theta_j \cdot \ln(\lambda_j)$  para j=1, 2, se obtiene que la Ecuación (2) equivale al producto de dos distribuciones Gumbel (Fiorentino *et al.*, 1985):

$$F_{x}(x) = \exp\left[-e^{-(x-\varepsilon_{1})/\theta_{1}}\right] \cdot \exp\left[-e^{-(x-\varepsilon_{2})/\theta_{2}}\right]$$
(3)

$$F_X(x) = F_{X_1}(x) \cdot F_{X_2}(x) \text{ para } x \ge 0$$
 (4)

 $\lambda_j$  es el número medio de gastos máximos independientes y  $\theta_j$  es la magnitud media de tales gastos máximos, con j=1 para las crecientes básicas y j=2 para las extraordinarias. Cuando no existen las crecientes dispersas, la distribución TCEV se reduce al modelo Gumbel simple. La función de densidad de probabilidad de X, según la Ecuación (2), es (Rossi et al., 1984):

$$f_X(x) = F_X(x) \cdot \psi(x) \tag{5}$$

en la cual:

$$\psi(x) = (\lambda_1/\theta_1) \cdot \exp(-x/\theta_1) + (\lambda_2/\theta_2) \cdot \exp(-x/\theta_2)$$
 (6)

# Cálculo de los parámetros de ajuste

Rossi *et al*. (1984) exponen el logaritmo natural de la *función de máxima* verosimilitud y lo designan por  $L_{fmv}$ ; su expresión es:



$$L_{fmv} = \sum_{i=1}^{n} \ln f_X(x_i) = \sum_{i=1}^{n} \ln F_X(x_i) + \sum_{i=1}^{n} \ln \psi(x_i)$$
 (7)

Igualando a cero las derivadas parciales de  $L_{fmv}$  con respecto a los cuatro parámetros de ajuste y realizando varias operaciones algebraicas, Rossi *et al.* (1984) obtienen las cuatro ecuaciones siguientes, las cuales se resuelven mediante una técnica numérica de sustitución sucesiva:

$$\lambda_{j} = \lambda_{j} \frac{\sum_{i=1}^{n} \frac{\exp(-x_{i}/\theta_{j})}{\psi(x_{i})}}{\theta_{j} \sum_{i=1}^{n} \exp(-x_{i}/\theta_{j})} \text{ para } j = 1,2$$
(8)

$$\theta_{j} = \frac{\sum_{i=1}^{n} \frac{x_{i} \cdot \exp\left(-x_{i}/\theta_{j}\right)}{\psi(x_{i})}}{\sum_{i=1}^{n} x_{i} \cdot \exp\left(-\frac{x_{i}}{\theta_{j}}\right) + \sum_{i=1}^{n} \frac{\exp\left(-x_{i}/\theta_{j}\right)}{\psi(x_{i})}} \text{ para } j = 1, 2$$

$$(9)$$

Para asegurar una convergencia rápida, conviene iniciar con unos valores de  $\lambda$  y  $\theta$  lo más aproximados posible. Para la búsqueda de los valores iniciales, primero se dibujan en el papel Gumbel-Powell (Chow, 1964) los datos del registro de crecientes, asignándoles una posición gráfica o probabilidad de no excedencia por medio de la fórmula de Weibull (Benson, 1962):

$$P(X < x) = \frac{m}{n+1} \tag{10}$$

donde m es el número de orden del dato cuando éstos han sido ordenados de menor a mayor; n, el número de observaciones, crecientes o datos. En seguida se identifica la serie básica y se representa por una línea recta o distribución Gumbel, definiendo un punto F1, X1 en su inicio y otro hacia el final de los datos que se designa (F2, X2). Después se traza otra recta o modelo Gumbel con mayor pendiente para representar los valores dispersos o crecientes extraordinarias, usando como punto inicial F2, X2 y definiendo un punto final en F3, X3, hacia el último de los datos. Estas tres parejas de valores se llevan a las fórmulas siguientes que proceden de la Ecuación (3) (Campos-Aranda, 2002) para definir los valores iniciales  $\lambda_j$  y  $\theta_j$ :

$$\theta_1 = \frac{(X2 - X1)}{\{-\ln[-\ln(F2)]\}} - \{-\ln[-\ln(F1)]\}$$
 (11)

$$\varepsilon_1 = X1 - \theta_1 \{-\ln(F1)\}$$
 (12)

$$\lambda_1 = \exp(\varepsilon_1/\theta_1) \tag{13}$$

$$\theta_2 = \frac{(X3 - X2)}{\{-\ln[-\ln(F3)]\}} - \{-\ln[-\ln(F2)]\}$$
 (14)



$$\varepsilon_2 = X2 - \theta_2 \{ -\ln[-\ln(F2)] \}$$
 (15)

$$\lambda_2 = \exp(\varepsilon_2/\theta_2) \tag{16}$$

Con base en los valores iniciales se aplican por primera vez las ecuaciones (8) y (9), y se evalúa la función logarítmica de máxima verosimilitud ( $L_{fmv}$ ) con la ecuación (7) y las número (5) y (6) auxiliares. En seguida, los nuevos valores de  $\lambda_j$  y  $\theta_j$  se convierten en los iniciales y se vuelven a aplicar las ecuaciones (8) y (9). El proceso se repite hasta que el valor de  $L_{fmv}$  se estabiliza, es decir, que ya no disminuye.

Al concluir este proceso de sustituciones sucesivas se tienen los valores óptimos de los parámetros de ajuste y con ellos se aplica de manera repetida la Ecuación (1), para obtener parejas de valores de x y  $F_X(x)$  para construir el modelo TCEV en el papel Gumbel-Powell y obtener las predicciones buscadas asociadas con periodos de retorno (Tr) o intervalos promedio de recurrencia de 5, 10, 25, 50, 100 y 500 años. Al ser Tr el recíproco de la probabilidad de excedencia, entonces los Tr citados corresponden a probabilidades de no excedencia  $[F_X(x)]$  de 0.80, 0.90, 0.96, 0.98, 0.99 y 0.998, respectivamente.

# Análisis regional jerárquico con la distribución TCEV



Para comparar datos de ríos o registros diferentes, Rossi *et al*. (1984) y Fiorentino *et al*. (1985) introducen la variable adimensional *Y*, que estandariza las crecientes anuales, la cual es:

$$Y = \frac{(X - \varepsilon_1)}{\theta_1} = \frac{X}{\theta_1} - \ln(\lambda_1)$$
 (17)

siendo  $\theta_1$  y  $\varepsilon_1$  o  $\lambda_1$  los parámetros de la componente básica o de una distribución Gumbel sin las crecientes extraordinarias. La FDP de Y se obtiene con la Ecuación (2), esto es:

$$F_{Y}(y) = F_{X}(\theta_{1} \cdot y + \varepsilon_{1}) = \exp(-\lambda_{1}^{*} \cdot e^{-y/\theta_{1}^{*}} - \lambda_{2}^{*} \cdot e^{-y/\theta_{2}^{*}})$$
(18)

 $\lambda_1^*$ ,  $\theta_1^*$ ,  $\lambda_2^*$  y  $\theta_2^*$  son los llamados parámetros de ajuste *regionales*, los cuales permiten asumir que la distribución TCEV definida por la Ecuación (18) representa todos los registros de una región aceptada homogénea hidrológicamente; registros que lógicamente pueden tener inicios y amplitudes diferentes.

El análisis regional *jerárquico* de crecientes fue propuesto inicialmente por Fiorentino, Gabriele, Rossi y Versace (1987). Se basa en



la Ecuación (18) y asume tres etapas: (1) identificación de una región homogénea, donde el coeficiente de asimetría (*Cs*) no varía de un sitio a otro o entre los registros disponibles; (2) identificación de subregiones homogéneas, donde el coeficiente de variación (*Cv*) también se puede considerar constante, y (3) deducción de modelos de regresión para estimar el gasto medio anual en localidades sin datos hidrométricos. Este enfoque regional ha sido descrito y aplicado por Cannarozzo, D'Asaro y Ferro (1995); Ferro y Porto (2006), y Boni, Parodi y Rudari (2006), para procesar crecientes y precipitaciones extremas.

## Detección de datos dispersos máximos

Ya sea que se aplique la distribución TCEV a un solo registro amplio (ajuste local), o que se utilice a nivel regional, la aplicación de las ecuaciones (11) a (16) para definir los valores iniciales de los parámetros de ajuste ( $\lambda_1$ ,  $\theta_1$ ,  $\lambda_2$ ,  $\theta_2$ ) requiere de la definición de las crecientes extraordinarias o valores dispersos (*outliers*).

Kottegoda y Rosso (2008) indican que una manera simple y práctica de detección de los valores dispersos (anómalos, discordantes o atípicos) consiste en ordenar los gastos máximos anuales en forma progresiva de



magnitud y luego revisar su evolución para detectar dónde ocurre un cambio sustancial de incremento entre valores y con ello establecer el inicio de los valores dispersos.

Por otra parte, si se acepta que el registro de crecientes de un río sigue un modelo probabilístico TCEV, se deberán detectar las crecientes extraordinarias de una distribución Gumbel. Entre las pruebas estadísticas que existen para detectar valores dispersos, la que se basa en el sesgo o asimetría muestral ha probado ser potente. Tal prueba, denominada *Test Skew* por Francés-García (1995), también la emplearon Rossi *et al.* (1984) y Fiorentino *et al.* (1985), y en ella se comienza por calcular el estadístico  $G_i$  con la ecuación:

$$G_{j} = \sqrt{n - j + 1} \frac{\sum_{i=1}^{n-j+1} (x_{i} - \bar{x}(n-j+1))^{3}}{\left\{\sum_{i=1}^{n-j+1} [x_{i} - \bar{x}(n-j+1)]^{2}\right\}^{3/2}}$$
(19)

en la cual  $x_i$  son los datos ordenados de menor a mayor y  $\bar{x}(k)$  es la media aritmética de las primeras k observaciones mayores. Según la ecuación (19),  $G_j$  es un estimador de la asimetría muestral que no considera los j-1 valores más grandes.

Matalas et al. (1975) obtuvieron mediante simulación Monte Carlo la media y desviación estándar del sesgo muestral de la distribución Gumbel para muestras que variaron de 10 a 90 datos. Con base en tales valores, Francés-García (1995) obtiene los valores críticos  $G_a$  en una



prueba de una cola, con un nivel de significancia (a) del 5 %, mostrados en la Tabla 1.

**Tabla 1**. Valores críticos del estadístico  $G_j$  con un nivel de significancia (a) del 5 %.

Tamaño	Ga	Tamaño	<b>G</b> a	Tamaño	<b>G</b> a
muestra (n)		muestra (n)		muestra (n)	
10	1.690	40	1.879	70	1.859
20	1.833	50	1.877	80	1.845
30	1.877	60	1.869	90	1.831

# Pruebas de homogeneidad subregional

Antes de ajustar la distribución TCEV a un conjunto de datos de crecientes que pertenecen a una cierta región o subregión que se sospecha es homogénea desde un punto de vista hidrológico, se debe probar o verificar tal homogeneidad. A fin de comprobar numéricamente la conveniencia de la subdivisión de una región hidrológica en dos o más subgrupos de cuencas, se puede aplicar una prueba estadística muy



simple que mide la heterogeneidad de un área o zona específica, denominada  $rango\ regional\ normalizado\ de los coeficientes de variación de las crecientes [RRN(<math>Cv$ )], definido por Burn (1990) como:

$$RRN(Cv) = \frac{R(Cv)}{M(Cv)}$$
 (20)

donde R(Cv) y M(Cv) son el rango y la mediana de los valores del Cv de la subregión o zona analizada. Entonces, cuando el RRN(Cv) aplicado por subregiones resulta menor que el de la región total, la subdivisión es conveniente o procedente.

Otras pruebas estadísticas que permiten verificar la homogeneidad hidrológica de una región y/o de cada subregión son, por ejemplo: (1) el test *R* de Wiltshire (1986), que contrasta la forma de la FDP regional con las locales; (2) el test de momentos de Gupta, Mesa y Dawdy (1994), que verifica la homogeneidad a través de las pendientes de los momentos de orden uno a cuatro de las crecientes locales, según el tamaño de cuenca, y (3) el test de momentos de probabilidad pesada de Kumar, Guttarp y Foufoula-Georgiu (1994), similar al anterior. Estas pruebas las han aplicado Ferro y Porto (2006) en Sicilia, Italia.

# Registros de crecientes por procesar

Corresponden a los 21 que procesó Campos-Aranda (2014b) de la Región Hidrológica No. 10 (Sinaloa), cuyas características estadísticas e índice de estacionalidad se exponen en la Tabla 2, por orden decreciente de tamaño de cuenca. En la columna 4 se indican las amplitudes de cada registro, abarcando un total de 777 datos. En la Figura 1 se muestra la ubicación de la Región Hidrológica No. 10 en la república mexicana.

**Tabla 2**. Parámetros estadísticos de los registros de gasto máximo anual de las 21 estaciones hidrométricas indicadas de la Región Hidrológica No. 10 (Sinaloa), México.

1	2	3	4	5	7	8	9	10
No.	Estación	A	n	Qm	Cv	Cs	Ck	$\bar{r}$
	hidrométrica	(km²)		(m³/s)				
1	Huites	26 057	51	3 328.333	0.995	2.210	7.7	0.3739
2	San Francisco	17 531	33	1 724.636	0.841	1.885	6.6	0.4133
3	Santa Cruz	8 919	52	1 037.615	1.120	3.231	16.3	0.5507
4	Jaina	8 179	56	1 020.786	1.114	3.358	17.2	0.5281
5	Palo Dulce	6 439	21	1 129.238	1.218	3.844	18.7	0.3405



6	Ixpalino	6 166	45	1 198.978	0.905	2.969	13.7	0.6214
7	La Huerta	6 149	28	945.107	0.593	0.186	2.1	0.4219
8	Chinipas	5 098	24	883.083	0.666	1.403	5.6	0.4457
Valor	promedio	-	-	-	-	-	-	0.4619
9	Tamazula	2 241	32	596.875	0.655	2.846	13.8	0.5928
10	Naranjo	2 064	45	633.311	1.053	1.816	6.4	0.7047
11	Acatitán	1 884	43	813.256	1.065	2.331	10.7	0.7497
12	Guamúchil	1 645	32	702.344	0.928	2.826	13.5	0.7299
13	Choix	1 403	38	348.974	0.880	2.903	13.3	0.6889
14	Badiraguato	1 018	26	1 224.346	1.555	3.420	15.9	0.5989
Valor	promedio	-	1	-	-	-	-	0.6775
15	El Quelite	835	33	479.091	0.869	1.558	5.6	0.8035
16	Zopilote	666	56	351.857	0.806	0.774	2.8	0.8322
17	Chico Ruiz	391	19	205.737	0.717	0.339	2.3	0.8161
18	El Bledal	371	56	289.000	0.915	2.693	12.9	0.8458
19	Pericos	270	30	250.800	0.625	1.133	4.1	0.8158
20	La Tina	254	24	104.958	1.471	3.860	19.6	0.7207



21	Bamícori	223	33	189.182	0.933	1.415	4.3	0.8810
Valor	promedio	-	-	-	-	-	-	0.8164

#### Simbología:

A: área de cuenca.

n: número de datos del registro.

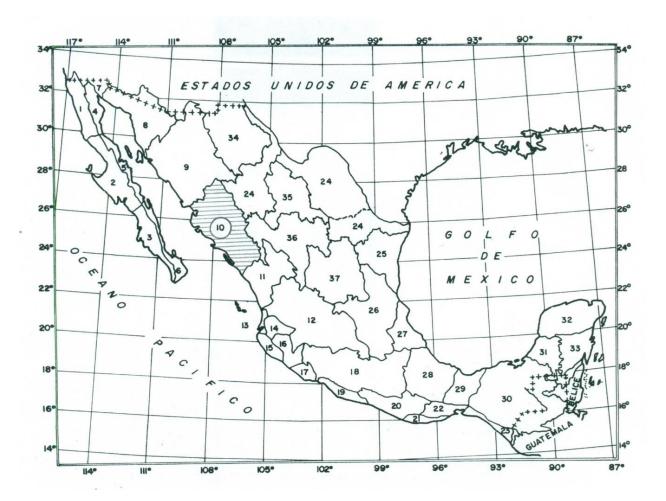
Qm: gasto promedio anual (antes  $\bar{Q}^k$ ),  $m^3/s$ .

Cv: coeficiente de variación, adimensional.

Cs: coeficiente de asimetría, adimensional.

Ck: coeficiente de curtosis, adimensional.

 $\bar{r}$ : índice de estacionalidad, adimensional.



**Figura 1**. Ubicación de la Región Hidrológica No. 10 (Sinaloa) en la república mexicana.

# Resultados y discusión



### Subregiones por procesar

El estudio de los valores mostrados en la última columna de la Tabla 2 orientó a subdividir la región hidrológica No. 10 (Sinaloa) en tres *subregiones* para la aplicación del método del índice de crecientes, tales subregiones se denominan: (**A**) de cuencas mayores a los cinco mil km², cuyo índice de estacionalidad varía de 0.3405 en Palo Dulce a 0.6214 en Ixpalino, con 0.4619 como valor medio; (**B**) de cuencas mayores de los mil km², con índices de estacionalidad que fluctúan entre 0.5928 y 0.7497, con 0.6775 como valor promedio, y (**C**) cuencas menores de los mil km², con índices de estacionalidad que oscilan entre 0.7207 y 0.8810, con un valor medio de 0.8164.

En la Tabla 3 se presentan las 8, 6 y 7 estaciones hidrométricas que integran cada *subregión*; los resultados respectivos de la prueba del *RRN(Cv)* se muestran en cursivas. Se observa que la subdivisión resulta muy conveniente para la subregión **A** de las grandes cuencas, pero no implica gran diferencia en las subregiones **B** y **C** de cuencas menores. Sin embargo, en lo general, tal subdivisión es conveniente (Campos-Aranda, 2013).

**Tabla 3**. Aplicación de la prueba del rango regional normalizado en las tres subregiones adoptadas de la Región Hidrológica No. 10 (Sinaloa), México y en ella misma.

Subregión	Α	Subregión	В	Subregió	on C			
Estación hidrométrica	Cv	Estación hidrométrica	Cv	Estación hidrométrica	Cv			
Huites	0.995	Tamazula	0.655	El Quelite	0.869			
San Francisco	0.841	Naranjo	1.053	Zopilote	0.806			
Santa Cruz	1.120	Acatitán	1.065	Chico Ruiz	0.717			
Jaina	1.114	Guamúchil	0.928	El Bledal	0.915			
Palo Dulce	1.218	Choix	Choix 0.880		0.625			
Ixpalino	0.905	Badiraguato	1.555	La Tina	1.471			
La Huerta	0.593	Valor mínimo	0.655	Bamícori	0.933			
Chinipas	0.666	Valor máximo	1.555	Valor mínimo	0.625			
Valor mínimo	0.593	Valor mediano	0.990	Valor máximo	1.471			
Valor máximo	1.218			Valor mediano	0.869			
Valor mediano	0.950	RRN(Cv)	0.909	RRN(Cv)	0.974			
RRN(Cv)	0.658							
Región Hidrológica No. 10 (Sinaloa)								
Valor mínimo			0.593	Valor mediano	0.915			
\	/alor máx	kimo	1.555	RRN(Cv)	1.051			



# Eliminación de registros anómalos

Definidas las subregiones donde se aplicará el método del índice de crecientes, se procede a revisar de manera conjunta los parámetros estadísticos para detectar registros discordantes o anómalos; lo anterior equivale a la primera etapa del análisis regional jerárquico. En la subregión  $\bf A$ , el registro de La Huerta tiene un Cs y un Ck que difieren notablemente del resto y por ello se eliminó. En la subregión  $\bf B$  no se eliminó ningún registro, pues los discordantes —Naranjo y Badiraguato—se puede considerar extremos. Finalmente, en la subregión  $\bf C$ , el registro de La Tina tiene valores del Cs, Ck y  $\bar{r}$ , plenamente discordantes con el resto y por ello se eliminó. Con base en lo anterior, el registro *conjunto* de la subregión  $\bf A$  tendrá 282 datos de crecientes ( $Q_i$ ); el de la subregión  $\bf B$  abarcará 216 valores, y el de la  $\bf C$  incluirá 227 datos.

#### Detección de las crecientes extraordinarias



Una inspección rutinaria a los valores críticos de la Tabla 1 destaca que para el tamaño común (de 30 a 60 datos) de los registros de crecientes disponibles actualmente se puede adoptar un valor crítico general de  $G_a$  de 1.87, el cual coincide con el usado por Fiorentino  $et\ al$ . (1985) para los ríos del norte de Italia. Si se utiliza tal valor crítico, se observa en la Tabla 4, con base en los valores sombreados, que en la mayoría de los registros sólo se detecta una creciente extraordinaria y como máximo tres en Huites, y dos en Jaina, Ixpalino, Choix y Badiraguato.

**Tabla 4**. Valores del test Skew para ratificación del número de crecientes extraordinarias adoptadas en cada estación hidrométrica de la Región Hidrológica No. 10 (Sinaloa), México.

No.	Estación	<b>G</b> <sub>1</sub>	<b>G</b> <sub>2</sub>	<b>G</b> <sub>3</sub>	<b>G</b> <sub>4</sub>	<b>G</b> <sub>5</sub>	<b>G</b> <sub>6</sub>	CE
	hidrométrica							
	Subregión A							
1	Huites	2.144	2.180	1.966	1.844	1.696	1.659	6
2	San Francisco	1.798	1.511	1.537	1.233	1.037	0.963	4
3	Santa Cruz	3.137	1.846	1.785	1.660	1.324	0.564	6
4	Jaina	3.268	2.176	1.458	1.403	1.332	1.295	8
5	Palo Dulce	3.564	1.435	0.544	0.622	0.617	0.514	2
6	Ixpalino	2.869	2.137	1.259	1.085	0.840	0.690	2



7	La Huerta	0.176	0.173	0.125	0.100	0.089	0.061	0
8	Chinipas	1.314	0.623	0.615	0.639	0.620	0.543	1
			Subre	gión B				
9	Tamazula	2.711	0.912	0.498	0.427	0.371	0.207	2
10	Naranjo	1.755	1.358	1.408	1.412	1.456	1.443	3
11	Acatitán	2.249	1.026	0.928	0.776	0.777	0.767	3
12	Guamúchil	2.692	1.011	0.997	0.867	0.734	0.698	1
13	Choix	2.787	1.955	0.956	0.686	0.668	0.578	3
14	Badiraguato	3.219	2.268	1.624	1.811	1.766	0.773	2
			Subre	gión C	1	1		
15	El Quelite	1.486	1.341	0.768	0.709	0.607	0.387	2
16	Zopilote	0.753	0.727	0.665	0.636	0.639	0.636	0
17	Chico Ruiz	0.312	0.290	0.330	0.255	0.308	0.388	0
18	El Bledal	2.620	1.546	1.191	1.099	0.953	0.772	2
19	Pericos	1.075	1.010	0.769	0.479	0.216	0.237	0
20	La Tina	3.614	0.494	0.469	0.435	0.427	0.290	1
21	Bamícori	1.350	1.400	1.214	1.256	1.356	1.488	2

Simbología

 $G_k$ : coeficiente de asimetría de orden k.

CE: número de crecientes extraordinarias detectadas.



En teoría, el valor de  $G_a$  debe ir decreciendo conforme j aumenta (Ecuación (19)) y cuando no lo hace, como es en el caso de los registros de Naranjo, Zopilote, Chico Ruiz y Bamícori, se debe a la presencia de valores extremos bajos.

Por otra parte, en la última columna de la Tabla 4 se indica el número de crecientes extraordinarias detectadas en cada registro de acuerdo con los valores dispersos, es decir, aquellos que se alejan de la tendencia general. Del análisis de los valores de la Tabla 4, se concluye que el test Skew es bastante *restrictivo* para la definición de las crecientes extraordinarias y por ello se adoptan los valores de la última columna. Por lo anterior, en la subregión **A** se tienen 29 crecientes extraordinarias; en la **B** ocurren 14, y en la **C** se adoptan siete. Estos valores auxilian en la adopción del punto *F*2, *X*2 de las ecuaciones (11) a (16).

# Ajuste de la distribución TCEV

Definidos los registros de crecientes por procesar en cada subregión, se estandarizan sus datos  $(Q_i)$  con base en los valores de la columna 5 de la Tabla 2, que son los gastos medios anuales de cada registro (Qm). En relación con lo anterior, en el segundo renglón de la Tabla 5 se indican el



número de estaciones procesadas y de datos adimensionales ( $Q_i/Qm$ ) procesados; en el tercer renglón se citan los respectivos valores mínimo y máximo de tales registros concatenados.

**Tabla 5**. Resultados del ajuste de la distribución TCEV según método numérico de sustitución en las tres subregiones hidrológicas indicadas de la Región Hidrológica No. 10 (Sinaloa), México.

Subregión	Α	В	С
Núm. (est.) datos	(7) 282	(6) 216	(6) 227
Mínimo y máximo	0.1008-6.8486	0.0086-7.5510	0.0284-5.4533
F1 - X1	0.021-0.1214	0.060-0.1037	0.026-0.0995
F2 - X2	0.890-1.9881	0.908-2.1543	0.961-2.6505
F3 - X3	0.990-6.5000	0.990-5.6000	0.996-5.4533
(λ <sub>1</sub> )inicial	4.851	3.337	4.353
$( heta_1)$ inicial	0.533	0.608	0.565
( $\lambda_2$ )inicial	0.343	0.397	0.349
$( heta_2)$ inicial	1.841	1.523	1.221
FO(ln L) <sub>inicial</sub>	-315.304	-246.828	-259.935
Núm. de iteraciones	46	50	106
(\lambda_1)final	5.693	3.816	4.023



2022, Instituto Mexicano de Tecnología del Agua Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$( heta_1)_{final}$	0.267	0.299	0.106
( $\lambda_2$ ) <sub>final</sub>	0.451	0.551	2.238
$( heta_2)_{final}$	1.386	1.277	0.678
FO(ln L) <sub>final</sub>	-238.226	-207.944	-217.299

Para establecer los puntos de inicio y fin de las crecientes básicas y de las extraordinarias, en cada registro conjunto se ordenaron sus datos en orden progresivo de magnitud y se les calculó una probabilidad empírica con base en la Ecuación (10). En seguida, con base en los resultados del inciso anterior, se define el punto F2, X2; posteriormente, hacia el final del registro, el punto F3, X3. El punto F1, X1 se escoge en el comienzo del registro. Los valores seleccionados para tales puntos se indican en los renglones cuarto a sexto de la Tabla 5. Estos valores permitieron el inicio y desarrollo del método numérico de sustituciones sucesivas.

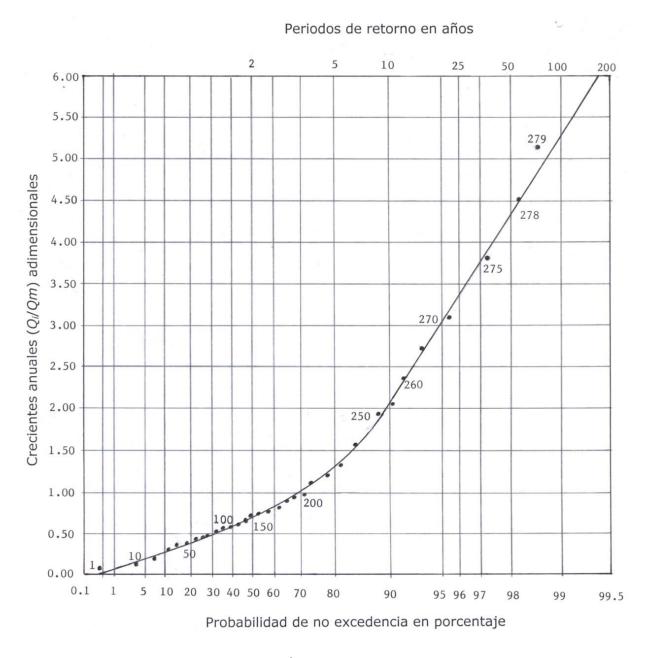
En los siguientes cinco renglones de la Tabla 5 se muestran los resultados de las ecuaciones (11) a (16) y de la Ecuación (7). Después del número de iteraciones indicadas, se estabilizó el valor mínimo de la función logarítmica de máxima verosimilitud, y con ello se tienen los valores finales u óptimos de los parámetros de ajuste.

En seguida se aplica la Ecuación (2) para diversos valores de los cocientes adimensionales  $Q_i/Qm$  de cada registro concatenado, para definir las distribuciones TCEV subregionales, las cuales se han dibujado



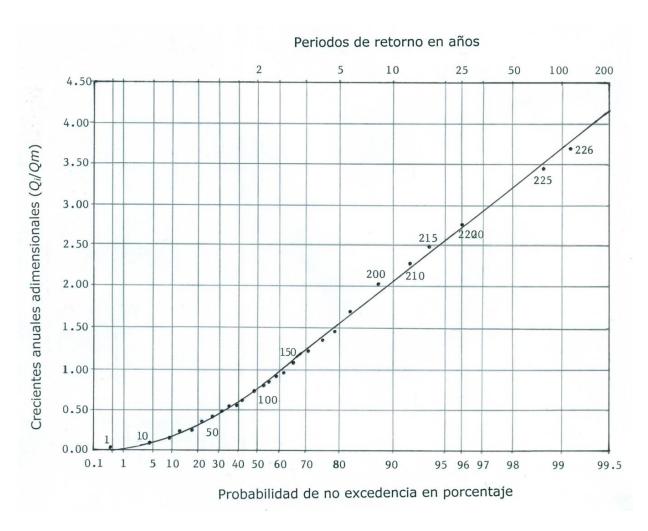
en el papel Gumbel-Powell en la Figura 2, Figura 3 y Figura 4 con igual escala. Los datos de cada registro concatenado se han dibujado en intervalos de diez valores para las crecientes básicas y de manera aleatoria para el grupo de crecientes extraordinarias.





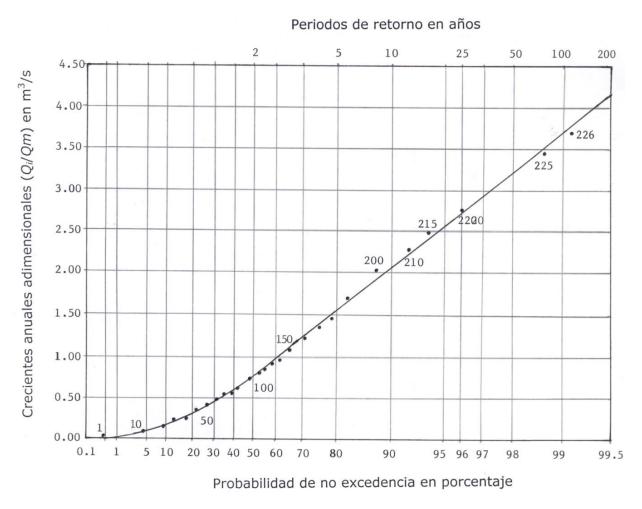
**Figura 2**. Ajuste de la distribución TCEV al registro concatenado de 282 crecientes anuales adimensionales  $(Q_i/Qm)$  integrado en la subregión A de la Región Hidrológica No. 10 (Sinaloa), México.





**Figura 3**. Ajuste de la distribución TCEV al registro concatenado de 216 crecientes anuales adimensionales  $(Q_i/Q_m)$  integrado en la subregión B de la Región Hidrológica No. 10 (Sinaloa), México.





**Figura 4**. Ajuste de la distribución TCEV al registro concatenado de 227 crecientes anuales adimensionales  $(Q_i/Q_m)$  integrado en la subregión C de la Región Hidrológica No. 10 (Sinaloa), México.

También se aplica la Ecuación (2) para definir por tanteos las predicciones ( $Q^{Tr}/Qm$ ) con probabilidades de no excedencia de 0.80, 0.90, 0.96, 0.98, 0.99 y 0.998 correspondientes a cada subregión. Tales



predicciones se citan en cursivas en la Tabla 6 en el renglón de cada subregión.

**Tabla 6**. Errores relativos del contraste entre predicciones (m³/s) locales y las del método regional del índice de crecientes en la Región Hidrológica No. 10 (Sinaloa), México.

Estación	FDP	Periodos de retorno en años FDP							
hidrométrica		5	10	25	50	100	500		
Subregión <b>A</b>	TCEV <sub>adim</sub>	1.29	2.05	3.33	4.31	5.27	7.51		
	LP3	4 368	6 679	11 009	15 613	21 776	45 130		
Huites	TCEV	4 294	6 823	11 083	14 345	17 540	24 996		
	ER	-1.7	2.2	0.7	-8.1	-19.5	-44.6		
	LP3	2 316	3 390	5 286	7 195	9 635	18186		
San Francisco	TCEV	2 225	3 536	5 743	7 433	9 089	12 952		
	ER	-3.9	4.3	8.6	3.3	-5.7	-28.8		
	LP3	1 525	2 274	3 434	4 447	5 583	8 721		
Santa Cruz	TCEV	1 339	2 127	3 455	4 472	5 468	7 792		
	ER	-12.2	-6.5	0.6	0.6	-2.1	-10.7		
	GVE	1 347	2 029	3 309	4 692	6 589	14 199		
Jaina	TCEV	1 317	2 093	3 399	4 400	5 380	7 666		
	ER	-2.2	3.2	2.7	-6.2	-18.3	-46.0		
Palo Dulce	LP3	1 674	2 543	3 906	5 107	6 460	10 215		



	TCEV	1 457	2 315	3 760	4 867	5 951	8 481
	ER	-13.0	-9.0	-3.7	-4.7	-7.9	-17.0
	LP3	1 635	2 311	3 410	4 431	5 651	9 450
Ixpalino	TCEV	1 547	2 458	3 993	5 168	6 319	9 004
	ER	-5.4	6.4	17.1	16.6	11.8	-4.7
	GVE	1 431	1 718	2 027	2 223	2 393	2 709
La Huerta	TCEV	1 219	1 937	3 147	4 073	4 981	7 098
	ER	-14.8	12.7	55.3	83.2	108.1	162.0
	GVE	1 231	1 626	2 209	2 714	3 283	4 919
Chinipas	TCEV	1 139	1 810	2 941	3 806	4 654	6 632
	ER	-7.5	11.3	33.1	40.2	41.8	34.8
Subregión <b>B</b>	TCEV <sub>adim</sub>	1.39	2.15	3.32	4.22	5.11	7.17
	LP3	779	1 029	1 421	1 777	2 195	3 476
Tamazula	TCEV	830	1 283	1 982	2 519	3 050	4 280
Tamazula	TCEV ER	830 6.5	1 283 24.7	1 982 39.5	2 519 41.8	3 050 39.0	4 280
Tamazula							
Tamazula Naranjo	ER	6.5	24.7	39.5	41.8	39.0	23.1
	ER LP3	6.5 947	24.7 1 473	39.5 2 323	41.8	39.0 3 973	23.1
	ER LP3 TCEV	6.5 947 880	24.7 1 473 1 362	39.5 2 323 2 103	41.8 3091 2 673	39.0 3 973 3 236	23.1 6 499 4 541
	ER LP3 TCEV ER	6.5 947 880 -7.1	24.7 1 473 1 362 -7.5	39.5 2 323 2 103 -9.5	41.8 3091 2 673 -13.5	39.0 3 973 3 236 -18.6	23.1 6 499 4 541 -30.1
Naranjo	ER LP3 TCEV ER GVE	6.5 947 880 -7.1 1 227	24.7 1 473 1 362 -7.5 1 779	39.5 2 323 2 103 -9.5 2 654	41.8 3091 2 673 -13.5 3 459	39.0 3 973 3 236 -18.6 4 422	23.1 6 499 4 541 -30.1 7 458



	TCEV	976	1 510	2 332	2 964	3 589	5 036
	ER	3.0	11.8	14.4	9.1	0.3	-23.7
	LP3	470	668	999	1 317	1 706	2 976
Choix	TCEV	485	750	1 159	1 473	1 783	2 502
	ER	3.2	12.3	16.0	11.8	4.5	-15.9
	LP3	1 405	2 556	5 330	9 061	15 191	48 733
Badiraguato	TCEV	1 702	2 632	4 065	5 167	6 256	8 779
	ER	21.1	3.0	-23.7	-43.0	-58.8	-82.0
Subregión <b>C</b>	TCEV <sub>adim</sub>	1.56	2.07	2.72	3.19	3.67	4.76
	LP3	722	1 043	1 519	1 919	2 355	3 506
El Quelite	TCEV	747	992	1 303	1 528	1 758	2 280
	ER	3.5	-4.9	-14.2	-20.4	-25.4	-35.0
	LP3	556	730	926	1 051	1 158	1 352
Zopilote	TCEV	549	728	957	1 122	1 291	1 675
	ER	-1.3	-0.3	3.3	6.8	11.5	23.9
	LP3	341	43-1	515	558	589	629
Chico Ruiz	TCEV	321	426	560	656	755	979
	ER	-5.9	-1.2	8.7	17.6	28.2	55.6
	GVE	401	565	834	1 090	1 404	2 446
El Bledal	TCEV	451	598	786	922	1 061	1 376
	ER	12.5	5.8	-5.8	-15.3	-24.4	-43.7
Pericos	LP3	354	465	620	745	878	1 222



	TCEV	391	519	682	800	920	1 194
	ER	10.5	11.6	10.0	7.4	4.8	-2.3
	GVE	144	229	393	573	824	1 860
La Tina	TCEV	164	217	285	335	385	500
	ER	13.9	-5.2	-27.5	-41.5	-53.3	-73.1
	GVE	275	416	649	866	1 125	1 913
Bamícori	TCEV	295	392	515	603	694	901
	ER	7.3	-5.8	-20.6	-30.4	-38.3	-52.9
No. ER neg	ativos	12	7	7	10	12	15
<i>ER</i> mínimo n	ER mínimo negativo		-9.0	-27.5	-43.0	-58.8	-82.0
Suma <i>ER</i> ne	Suma <i>ER</i> negativos		-42.1	-105.0	-183.9	-278.3	-532.3
No. ER pos	No. ER positivos		14	14	11	9	6
ER máximo į	ER máximo positivo		24.7	55.3	83.2	108.1	162.0
Suma <i>ER</i> po	sitivos	81.5	109.3	211.1	238.4	250.0	299.4
Indic	Indicadores de los <i>ER</i> del cuadro 4 de Campos-Aranda (2014b)						
No. ER neg	No. <i>ER</i> negativos		13	10	9	7	7
<i>ER</i> mínimo n	ER mínimo negativo		-12.2	-34.1	-49.4	-61.1	-78.9
Suma <i>ER</i> ne	gativos	-71.2	-77.8	-128.0	-152.0	-171.7	-219.2
No. ER pos	sitivos	9	8	11	12	14	14
ER máximo į	positivo	20.6	21.3	41.3	76.7	122.6	289.1
Suma <i>ER</i> po	sitivos	63.7	70.8	176.4	283.7	423.6	915.5

Simbología:

FDP: función de distribución de probabilidades.



TCEV<sub>adim</sub>: distribución TCEV subregional, adimensional.

TCEV: predicción del método del índice de crecientes, m³/s.

LP3: predicción de la FDP Log-Pearson tipo III, m³/s.

GVE: predicción de la FDP General de Valores Extremos, m³/s.

ER: error relativo en porcentaje.

# Contraste entre predicciones locales y regionales

En la Tabla 6, las predicciones locales proceden del cuadro 2 de Campos-Aranda (2014b) y se obtuvieron con una de las tres FDP que ajustó, la que condujo al menor error estándar de ajuste (Kite, 1977). Tales FDP aplican bajo precepto y son la Log-Pearson tipo III (LP3), la General de Valores Extremos (GVE) y la Logística Generalizada (LOG). En la Tabla 6, las predicciones regionales se designan TCEV y se obtienen multiplicando las predicciones adimensionales ( $TCEV_{adim}$ ) de cada subregión por el gasto medio anual (Qm) tomado de la Tabla 2. Los errores relativos (ER) de las predicciones calculadas con el método del índice de crecientes se obtienen con la expresión:



$$ER = \frac{(q^{Tr})_{estim} - (q^{Tr})_{obser}}{(q^{Tr})_{obser}} 100$$
 (21)

en donde el error relativo se expresa en porcentaje y presenta un valor negativo cuando el gasto regional estimado  $(Q^{Tr})_{estim}$  resultó menor que el calculado como predicción local u observado  $(Q^{Tr})_{obser}$ ; cuando conduce a un valor positivo, el gasto regional estimado fue superior al observado.

En los últimos 13 renglones de la Tabla 6 se confrontan los indicadores de los *ER* calculados entre las predicciones locales y las regionales de la distribución TCEV, y los que obtuvo Campos-Aranda (2014b) con el método regional de las estaciones-años y la distribución GVE, aplicado según el enfoque de la región de influencia, que es un procedimiento mucho más complicado. Se detectan las siguientes diferencias; las dos primeras no son preponderantes:

- Respecto al número de los ER negativos, los cuales son más importantes, ya que implican una subestimación de predicciones, se observa que es menor en los Tr bajos y mayor en los Tr altos. Lo anterior es señal de un comportamiento no adecuado.
- En cuatro de los seis Tranalizados el ER mínimo extremo fue menor, no así su suma, que nuevamente en los Tr bajos es menor, pero resulta mayor en los Tr altos.

Donde sí hay diferencias apreciables es en los ER positivos, principalmente en los Tr altos, ya que tanto su número como sus valores



extremos y de manera importante las sumas son menores con el método de índice de crecientes.

## Estimación de predicciones locales

Esta aplicación del método del índice de crecientes desarrollado se realizará en las cinco estaciones hidrométricas que Campos-Aranda (2013) utilizó para contraste de su gasto medio anual (*Qm*), ya que sus registros, además de ser cortos, son no confiables. Tales estaciones son San Ignacio, Piaxtla, Urique II, Cazanate y Los Molinos.

En la Tabla 7 se exponen sus tamaños de cuenca (A) en km², su amplitud de registro (n) en años y el valor estimado de Qm en m³/s con las ecuaciones empíricas de la referencia citada. Con base en el tamaño de cuencas, las tres primeras estaciones pertenecen a la subregión  $\bf A$ , Cazanate a la subregión  $\bf B$  y Los Molinos a la subregión  $\bf C$ .

**Tabla 7**. Errores relativos del contraste entre predicciones (m³/s) locales y las del método regional del índice de crecientes en cinco registros no confiables de la Región Hidrológica No. 10 (Sinaloa), México.



Estación	Periodos de retorno, en años FDP						
hidrométrica	. 5.	5	10	25	50	100	500
Subregión <b>A</b>	TCEV <sub>adim</sub>	1.29	2.05	3.33	4.31	5.27	7.51
San Ignacio	LP3	2 068	2 640	3 525	4 319	5 244	8 040
$A = 10 920 \text{ km}^2$ ( $n = 19$ )	TCEV	1 975	3 138	5 098	6 598	8 067	11 496
$Qm = 1 530.8$ $m^3/s$	ER	-52.9	18.9	44.6	52.8	53.8	43.0
Piaxtla	LOG	1 518	2 327	4 076	6 250	9 625	26 565
$A = 5 307 \text{ km}^2$ (n = 16)	TCEV	1 314	2 089	3 393	4 391	5 369	7 651
$Qm = 1 \ 018.8$ $m^3/s$	ER	-13.4	-10.2	-16.8	-29.7	-44.2	-71.2
Urique II	LOG	398	475	583	672	769	1035
A=4000 km <sup>2</sup> (n=19)	TCEV	1 136	1 805	2 932	3 795	4 640	6 613
Qm = 880.5 m <sup>3</sup> /s	ER	185.4	280.0	402.9	464.7	503.4	538.9
Subregión <b>B</b>	TCEV <sub>adim</sub>	1.39	2.15	3.32	4.22	5.11	7.17
Cazanate	LP3	848	1 298	1 956	2 492	3 052	4 418
$A = 1 813 \text{ km}^2$ $(n = 19)$	TCEV	827	1 219	1 975	2 564	3 039	4 265
Qm = 594.8 m <sup>3</sup> /s	ER	-2.5	-6.1	1.0	2.9	-0.40	-3.5



Subregión <b>C</b>	TCEV <sub>adim</sub>	1.56	2.07	2.72	3.19	3.67	4.76
Los Molinos	LOG	219	255	297	328	358	428
$A = 501 \text{ km}^2$ (n = 13)	TCEV	468	621	815	956	1 100	1 427
Qm = 299.8 m <sup>3</sup> /s	ER	113.7	143.5	174.4	191.5	207.3	233.4

Simbología

FDP: función de distribución de probabilidades.

TCEV<sub>adim</sub>: distribución TCEV subregional, adimensional.

TCEV: predicción del método del índice de crecientes, m³/s.

LP3: predicción de la FDP Log-Pearson tipo III, m³/s.

LOG: predicción de la FDP Logística Generalizada, m³/s.

ER: error relativo, en porcentaje.

A: área de cuenca, en km<sup>2</sup>.

n: amplitud del registro, en años.

Qm: gasto medio anual, en m<sup>3</sup>/s.

Los registros de las cinco estaciones hidrométricas citadas proceden de Escalante-Sandoval y Reyes-Chávez (2002) y en efecto tienen errores de muestreo evidentes, por ejemplo: (1) el registro de Piaxtla tiene gastos máximos de 6 399 y 3 527 m³/s, que son superiores a los de San Ignacio de 4 020 y 3 075 m³/s. Lo anterior se supone incorrecto, pues la estación Piaxtla tiene un área de cuenca menor de la mitad de San Ignacio; (2) en la estación Urique II, con un tamaño de cuenca de 4 000 km², su gasto



máximo es de 716 m $^3$ /s y su Qm de 302.6 m $^3$ /s, ambos datos se consideran bastante bajos; (3) algo similar ocurre en Los Molinos, con un máximo de 273 m $^3$ /s y un Qm de 159.2 m $^3$ /s.

Antes del análisis probabilístico de los registros de las cinco estaciones hidrométricas seleccionadas, se les aplicó el test de Wald-Wolfowitz (Bobée & Ashkar, 1991; Rao & Hamed, 2000; Meylan *et al.*, 2012), para verificar que están integrados por datos independientes y que son estacionarios. La estimación de las predicciones locales se llevó a cabo ajustando las distribuciones Log-Pearson tipo III (Bobée & Ashkar, 1991), General de Valores Extremos y Logística Generalizada (Hosking & Wallis, 1997). La selección de predicciones locales en cada registro se hizo de manera subjetiva, aceptando las representativas en los periodos de retorno de 100 y 500 años. Tales predicciones se indican en la Tabla 7, en el renglón donde aparece el nombre de cada estación hidrométrica.

Las predicciones regionales se tienen en el renglón de la distribución TCEV y proceden de multiplicar Qm por los valores adimensionales de cada subregión ( $TCEV_{adim}$ ). Finalmente, en el último renglón de cada estación hidrométrica se tienen los errores relativos (ER) evaluados según la Ecuación (21).

En la estación Cazanate, las predicciones locales y las regionales coinciden. En la Estación Piaxtla, las predicciones regionales tienen *ER* por déficit, pero posiblemente las predicciones locales están elevadas por la presencia de sus valores dispersos máximos. Lo contrario ocurre en las estaciones Urique II, Los Molinos y San Ignacio, cuyos *ER* por exceso en



el periodo de retorno de 500 años ascienden al 538.9, 233.4 y 43.0 %, respectivamente. Al tomar en cuenta que los registros procesados son cortos y no confiables, las predicciones regionales se consideran más exactas.

## **Conclusiones**

El procedimiento del método regional del índice de crecientes, expuesto y aplicado usando la distribución TCEV (two-component extreme value) como modelo probabilístico regional es mucho más simple que el enfoque de la región de influencia, y conduce a predicciones bastante aproximadas por déficit y más exactas por exceso, según los errores relativos evaluados contra las predicciones locales, los cuales se contrastaron en el final de la Tabla 6.

Además, el procedimiento expuesto permite el contraste entre las predicciones regionales y las locales de registros cortos o no confiables para disponer de *crecientes de diseño* más exactas, como se mostró en la Tabla 7.



El ajuste de la distribución TCEV con el método numérico de sustituciones sucesivas no ofreció dificultades computacionales ni al inicio ni para converger, por lo cual se recomienda su aplicación con el esquema expuesto, con base en las ecuaciones (11) a (16).

Expuestas las tres ventajas anteriores, se recomienda la aplicación sistemática del método del índice de crecientes con la distribución TCEV como modelo poblacional en otras regiones hidrológicas del país, para obtener las predicciones adimensionales  $(Q^{Tr}/Qm)$  que permiten estimar crecientes de diseño en cuencas sin datos hidrométricos, haciendo uso de relaciones regionales Qm = f(A).

### **Agradecimientos**

Se agradece al árbitro anónimo D la revisión exhaustiva del texto, cuyas correcciones permitieron eliminar errores ortográficos y mejorar la redacción. Por recomendación del árbitro anónimo F, se incluyó la Figura 1, que muestra la ubicación de la Región Hidrológica No. 10 en México.

#### Referencias

- Benson, M. A. (1962). Plotting positions and economics of engineering planning. *Journal of Hydraulics Division*, 88(6), 57-71.
- Bobée, B., & Ashkar, F. (1991). *The Gamma Family and derived distributions applied in Hydrology*. Littleton, USA: Water Resources Publications.



- Boni, G., Parodi, A., & Rudari, R. (2006). Extreme rainfall events: Learning from rain gauge time series. *Journal of Hydrology*, 327(3-4), 304-314.
- Burn, D. H. (1997). Catchment similarity for regional flood frequency analysis using seasonality measures. *Journal of Hydrology*, 202(1-4), 212-230.
- Burn, D. H. (1990). Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Research*, 26(10), 2257-2265.
- Campos-Aranda, D. F. (2014a). Análisis regional de frecuencia de crecientes en la Región Hidrológica No. 10 (Sinaloa), México. 1: índices de estacionalidad y regiones de influencia. *Agrociencia*, 48(2), 147-158.
- Campos-Aranda, D. F. (2014b). Análisis regional de frecuencia de crecientes en la Región Hidrológica No. 10 (Sinaloa), México. 2: contraste de predicciones locales y regionales. *Agrociencia*, 48(3), 255-270.
- Campos-Aranda, D. F. (2013). Estimación de la creciente media anual en la Región Hidrológica No. 10 de México con base en ecuaciones de regresión potencial. *Ingeniería, Investigación y Tecnología*, 14(4), 563-574.
- Campos-Aranda, D. F. (2002). Ajuste de la distribución TCEV por medio de optimización numérica no restringida. En: XVII Congreso Nacional de Hidráulica. Monterrey, N. L. (pp. 527-532). *Avances en*



- Hidráulica (9). México, DF, México: Asociación Mexicana de Hidráulica.
- Cannarozzo, M., D'Asaro, F., & Ferro, V. (1995). Regional rainfall and flood frequency analysis for Sicily using the two-component extreme value distribution. *Hydrological Sciences Journal*, 40(1), 19-40.
- Cunderlik, J. M., & Burn, D. H. (2002). The use of flood regime information in regional flood frequency analysis. *Hydrological Sciences Journal*, 47(1), 77-92.
- Cunnane, C. (1988). Methods and merits of regional flood frequency analysis. *Journal of Hydrology*, 100(1-3), 269-290.
- 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.
- Dalrymple, T. (1960). Flood frequency analysis. *Water supply paper* 1543-A. Reston, USA: U. S. Geological Survey.
- Escalante-Sandoval, C. A., & Reyes-Chávez, L. (2002). Apéndice A. En: *Técnicas estadísticas en hidrología* (pp. 281-298). México, DF,

  México: Facultad de Ingeniería de la Universidad Nacional

  Autónoma de México.
- Ferro, V., & Porto, P. (2006). Flood frequency analysis for Sicily, Italy. *Journal of Hydrologic Engineering*, 11(2), 110-122.



- Fiorentino, M., Gabriele, S., Rossi, F., & Versace, P. (1987). Hierarchical approach for regional flood frequency analysis (pp. 35-49). In: Singh, V. P. (ed.). *Regional flood frequency analysis*. Dordrecht, The Netherlands: Reidel Publishing Company.
- Fiorentino, M., Versace, P., & Rossi, F. (1985). Regional flood frequency estimation using the two-component extreme value distribution. *Hydrological Sciences Journal*, 30(1), 51-64.
- Francés-García, F. (1995). *Utilización de la información histórica en el análisis regional de las avenidas*. Nomografía No. 27 del CIMNE. Barcelona, España: Centro Internacional de Métodos Numéricos en Ingeniería.
- Gupta, V. K., Mesa, O. J., & Dawdy, D. R. (1994). Multiscaling theory of flood peaks: Regional quantile analysis. *Water Resources Research*, 30(12), 3405-3421.
- Hosking, J. R., & Wallis, J. R. (1997). Chapter 1: Regional frequency analysis. In: *Regional frequency analysis*. *An approach based on L-moments* (pp. 1-13). Cambridge, UK: Cambridge University Press.
- Houghton, J. C. (1978). Birth of a parent: The Wakeby distribution for modeling flood flows. *Water Resources Research*, 14(6), 1105-1109.
- Kite, G. W. (1977). Chapter 12: Comparison of frequency distributions.

  In: Frequency and risk analyses in hydrology (pp. 156-168). Fort
  Collins, USA: Water Resources Publications.



- Kottegoda, N. T., & Rosso, R. (2008). Theme 5.9: Identification and accommodation of outliers. In: *Applied statistics for civil and environmental engineers* (pp. 305-312) (2<sup>nd</sup> ed.). Oxford, UK: Blackwell Publishing Ltd.
- Kumar, P., Guttarp, P., & Foufoula-Georgiu, E. (1994). A probability weighted moment test to assess simple scaling. *Stochastic Hydrology and Hydraulics*, 8(3), 173-183.
- Machiwal, D., & Jha, M. K. (2012). Chapter 4: Methods for time series analysis. In: *Hydrologic time series analysis: Theory and practice* (pp. 51-84). Dordrecht, The Netherlands: Springer.
- Matalas, N. C., Slack, J. R., & Wallis, J. R. (1975). Regional skew in search of a parent. *Water Resources Research*, 11(6), 815-826.
- Meylan, P., Favre, A. C., & Musy, A. (2012). *Predictive hydrology. A frequency analysis approach*. Boca Raton, USA: CRC Press.
- Mujumdar, P. P., & Nagesh-Kumar, D. (2012). Floods in a changing climate. Hydrologic modeling. Cambridge, UK: International Hydrology Series (UNESCO) and Cambridge University Press.
- Ouarda, T. B. M. J. (2017). Regional flood frequency modeling. In: Singh, V. P. (ed.). *Handbook of applied hydrology* (pp. 77.1-77.8) (2<sup>nd</sup> ed.). New York, USA: McGraw-Hill Education.
- Potter, W. D. (1958). Upper and lower frequency curves for peak rates of runoff. *EOS. Transactions of AGU*, 39(1), 100-105.
- Rahman, A., Haddad, K., & Eslamian, S. (2014). Regional flood frequency analysis. In: Eslamian, S. (ed). *Handbook of engineering hydrology:*



- Modeling, climate change and variability (pp. 451-469). Boca Raton, USA: CRC Press.
- Ramírez-Orozco, A. I., Gutiérrez-López, A., & Ruiz-Silva, H. L. (2009).

  Análisis de la ocurrencia en el tiempo de los gastos máximos en

  México. *Ingeniería Hidráulica en México*, 24(1), 115-124.
- 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.
- Rossi, F., Fiorentino, M., & Versace, P. (1984). Two-component extreme value distribution for flood frequency analysis. *Water Resources Research*, 20(7), 847-856.
- Wiltshire, S. E. (1986). Regional flood frequency analysis I: Homogeneity statistics. *Hydrological Sciences Journal*, 31(3), 321-333.
- WMO, World Meteorological Organization. (1971). Annexed III: Standard tests of significance to be recommended in routine analysis of climatic fluctuations. In: *Climatic Change* (pp. 58-71) (Technical Note No. 79). Geneva, Switzerland: World Meteorological Organization. Reprinted edition.
- Zrinji, Z., & Burn, D. H. (1996). Regional flood frequency with hierarchical region of influence. *Journal of Water Resources Planning and Management*, 122(4), 245-252.