DOI: 10.24850/j-tyca-2022-02-06

Articles

Flood frequency analysis based on a mixed GEV distribution with upper limit applied to the Hydrological Region No. 10 (Sinaloa), Mexico

Análisis de frecuencias de crecientes basado en una distribución GVE mixta con límite superior en la Región Hidrológica No. 10 (Sinaloa), México

Daniel Francisco Campos-Aranda[1]

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

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

Abstract

Flood Frequency Analysis (FFA) processes the available record of annual maximum flows of a river, to estimate predictions associated with low probabilities of exceedance, its reciprocal is the return period (*Tr*) in years. These predictions are the *design floods*, with which all hydraulic works are planned, designed, and reviewed hydrologically, such as reservoirs, protective embankments, channel rectifications, bridges, and urban drainage works. In this work, a novel method of the FFA is described and applied, which incorporates additional hydrometric information in a mixed General Extreme Values (GEV) distribution; this procedure is aimed to accurately define the flood of *Tr* = 1 000 years. Initially, the information on average annual flow and an annual maximum flow of all the hydrometric stations that make up the homogeneous region under study are processed. The above, with the approach of the enveloping curves and to define an enveloping curve with zero probability of exceedance; whereby, defines the extreme maximum flow that is approached as asymptote by the upper part of the mixed GEV, which avoids an unreal increase in predictions. Lastly, based on each record of floods, synthetic sequences of 1 500 values are generated and one with similarity to the available data and more than ten floods of *Tr* > 150 years is chosen. To such a synthetic sequence a GEV distribution is adjusted, with the methods of moments L and LH, to select the one with the lowest standard error of fit. This distribution forms the lower part of the mixed GEV, until the point of inflection of *Tr* = 500 years. The described method was applied to the seven largest flood records of the Hydrological Region No. 10 (Sinaloa), Mexico, and based on its results, the Conclusions were formulated, which highlight its advantages and suggest its application to estimate predictions of high return periods (100 ≤ *Tr* ≤ 1 000 years) more accurately, by incorporating regional hydrometric information.

Keywords: GEV distribution, L moments, LH moments, enveloping curves, synthetic sequences, standard error of fit, inflection point, mixed GEV distribution.

Resumen

El análisis de frecuencias de crecientes (AFC) procesa el registro disponible de gastos máximos anuales de un río para estimar predicciones asociadas con bajas probabilidades de excedencia, cuyo recíproco es el periodo de retorno (*Tr*) en años. Estas predicciones son las *crecientes de diseño*, con las cuales se planean, diseñan y revisan hidrológicamente todas las obras hidráulicas, como embalses, diques protectores, rectificación de cauces, puentes y obras de drenaje urbano. En este trabajo se describe y aplica un método novedoso del AFC, que incorpora información hidrométrica adicional en una distribución General de Valores Extremos (GVE) mixta; tal procedimiento está orientado a definir con exactitud la creciente de *Tr* = 1 000 años. Inicialmente se procesa la información de gasto medio anual y de gasto máximo anual de todas las estaciones hidrométricas que integran la región homogénea bajo estudio. Lo anterior, con el enfoque de las curvas envolventes y teniendo como objetivo definir una curva envolvente cuya probabilidad de excedencia sea nula, por lo cual define el gasto máximo extremo al que se aproxima como asíntota la parte superior de la GVE mixta, la cual evita un incremento irreal de las predicciones. Posteriormente, con base en cada registro de crecientes, se generan secuencias sintéticas de 1 500 valores y se escoge una con similitud con los datos disponibles y más de diez crecientes de *Tr*

> 150 años. A tal secuencia sintética se le ajusta una distribución GVE, con los métodos de momentos L y LH, para seleccionar la de menor error estándar de ajuste. Esta distribución forma la parte inferior de la GVE mixta hasta el punto de inflexión de *Tr* = 500 años. El método se aplicó a los siete registros de crecientes más amplios de la Región Hidrológica No. 10 (Sinaloa), México y con base en sus resultados se formularon las conclusiones, las cuales destacan sus ventajas y sugieren su aplicación para estimar predicciones de altos periodos de retorno (100 ≤ *Tr* ≤ 1 000 años) con mayor exactitud, al incorporar información hidrométrica regional.

Palabras clave: distribución GVE, momentos L, momentos LH, curvas envolventes, secuencias sintéticas, error estándar de ajuste, punto de inflexión, distribución GVE mixta.

Received: 08/01/2020

Accepted: 02/25/2021

Introduction

Generalities

In all hydraulic infrastructure works used for protection against floods, such as reservoirs, dams, rectifications, canalizations, bridges, and urban sewage, their dimensioning and hydrological safety is conducted based on the *Design Floods* (DF). The most reliable hydrological estimation of the DFs is developed through the so-called *Flood Frequency Analysis* (FFA), which is the statistical processing of the maximum annual flows recorded in the river or channel, in the site selected for the construction of any hydraulic work. The FFA is used to estimate the DFs, which are maximum flows in the river associated with low probabilities of exceedance, whose reciprocal is the average recurrence interval or *return period* (*Tr*) in years. In particular, the estimates for *Tr* > 100 years are quite uncertain, due to the limited length of the maximum annual flows records and the small number of extraordinary floods that they include (Guse, Hofherr & Merz, 2010a).

The FFC consists of the following four steps: (1) verification of the statistical quality of the available flood register; (2) selection of a *probability distribution function* (PDF), or probabilistic model, from which extrapolations or predictions will be carried out on its right tail; (3) adoption of a method for estimating the fit parameters of the PDF, mainly variations of the method of moments, of maximum likelihood or the moments L; and (4) contrast of the various adjusted PDFs and their

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

methods for estimating parameters, to select the most convenient one for the available data. The latter is commonly done through the standard error of fit and mean absolute error (Kite, 1977; Stedinger, Vogel & Foufoula–Georgiou, 1993; Rao & Hamed, 2000; Meylan, Favre, & Musy, 2012; Stedinger, 2017).

The FFC holds several intrinsic weaknesses, perhaps the most important one is that it accepts that the available record of maximum annual flows becomes representative of the floods that will occur in the future. Furthermore, as already indicated, the maximum annual flows series are generally short enough to extrapolate to high *Trs*. The latter is exacerbated by the physical changes that occur in the basin and its main channel, due to deforestation and the construction of reservoirs and levees, as well as global or regional climate change (Merz & Blöschl, 2008).

To reduce the uncertainty of each DF estimate, it is suggested to use more hydrometric information in the FFA. Such information comes from three kinds of groups: (1) cause-effect, (2) temporal or historical floods, and (3) spatial or regional (Merz & Blöschl, 2008). The first group refers to a better understanding of the flood formation processes, to make their estimation more precise. For example, it is known that the mechanisms of formation of ordinary floods are different from those that originate extraordinary floods; For this reason, mixed PDFs are recommended, in the case of the TCEV developed by Rossi, Fiorentino, and Versace (1984). Within the second group, there is the expansion of the available records by integrating historical floods as non–systematic

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

data (Francés, 1998; Botero & Francés, 2010). Finally, the regionalization of floods, defining homogeneous areas or regions, allows the integration of various registers into a broad one, where predictions will be more exact and DF can also be estimated in places without measurements, located within such a homogeneous area (Hosking & Wallis, 1997; Rahman, Haddad, & Eslamian, 2014; Ouarda, 2017).

Objective

The purpose of this work was to describe and apply a novel approach to the FFA, incorporating regional information to estimate high *Tr*, on the order of one thousand years, with greater accuracy. The method combines extraordinary floods generated by random simulation and an extreme maximum flood, in a mixed PDF, with an upper limit. The method developed in Germany by Guse *et al*. (2010a), was applied to seven large records from Hydrological Region No. 10 (Sinaloa), Mexico, with slight modifications to expose a simpler FFA. Based on the analysis of the results, the conclusions of the study were formulated.

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

Materials and methods

Original approach

The approach of Guse *et al*. (2010a) incorporates additional spatial information to the FFA, in the form of maximum flows and their corresponding *Trs*, as well as an extreme maximum flow (Q_{max}) in a mixed PDF, in which the inflection point corresponds to the *Tr* of 500 years. The maximum additional flows are obtained based on probabilistic regional envelope curves (PREC), constructed with each of the homogeneous group's records that constitute such hydrometric stations. The original method consists of two fundamental steps: (1) integration of flows (Q_{PREC}) in the observed records, in a lower PDF, showing concavity upwards, and (2) definition from the inflection point of a mixed PDF with upper bound on Q_{max}, and therefore concave down. The PDFs used are the *General Extreme Values* (GVE, by its acronym in Spanish) and the Q_{PREC}*s* are generated in a random process with *Tr* which ranged from 150 to 1 500 years. Logically, the new series of floods obtained by synthetic simulation must show similarity with the observed record. A diagram of the proposed method is shown in Figure 1.

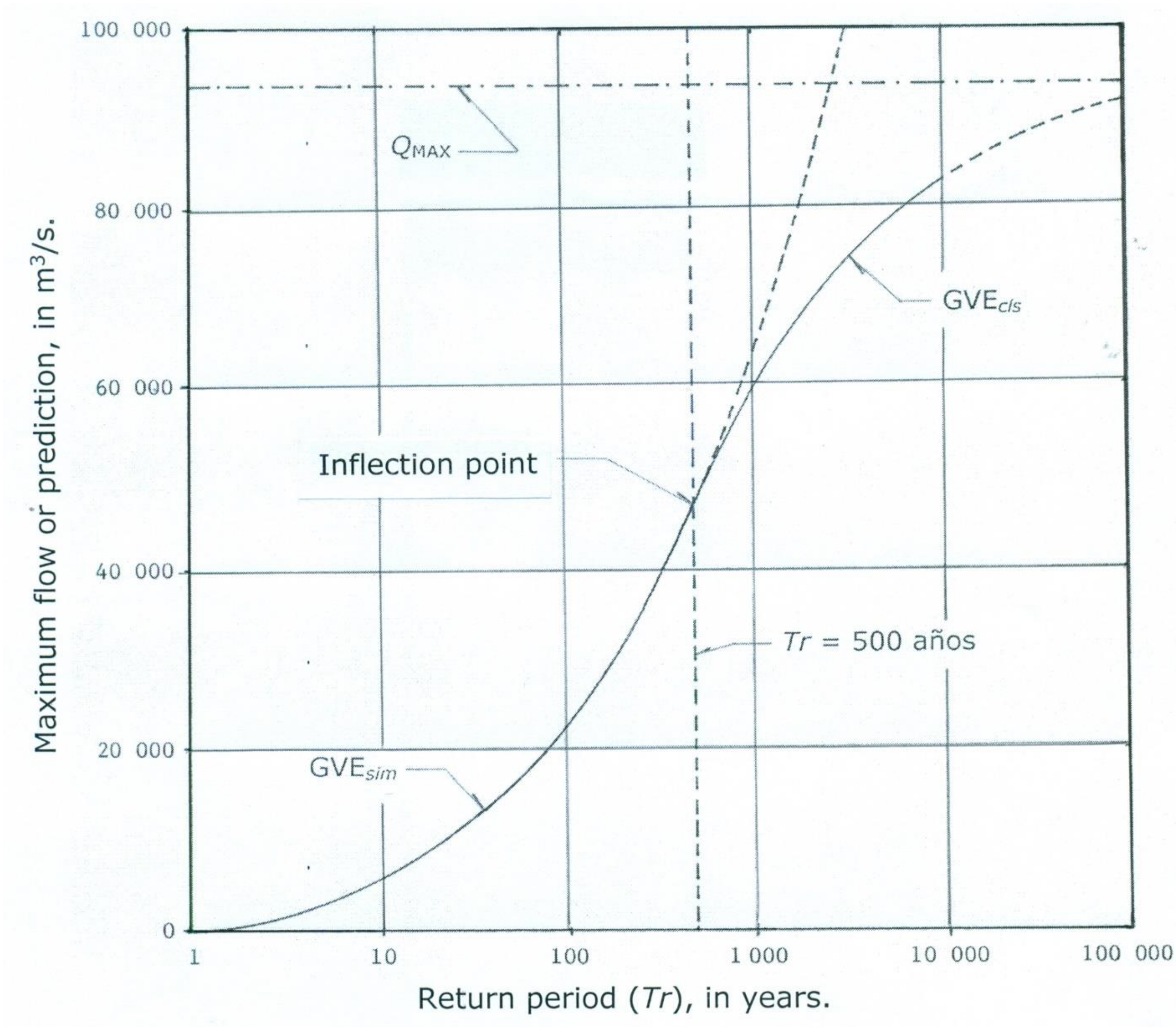

Figure 1. Scheme of the FFA method of the mixed GVE on semi-logarithmic paper, with numerical values for the Huites hydrometric station, of the Hydrological Region No. 10 (Sinaloa), Mexico.

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

Adopted modification

In the original method of Guse *et al*. (2010a) the selection and inclusion of the Q_{PREC} flows in each observed record was done randomly, generating 100 synthetic series of 1 500 values each, from the GVE distribution, adjusted to the historical or observed data (GVE_{obs}). In the modified procedure, synthetic series of 1 500 flows are also generated, but only one series is selected and with it, the mixed GVE is defined, according to the procedure and restrictions of Guse *et al*. (2010a). The synthetic series adopted contains more than ten maximum flows with $Tr > 150$ years and its adjustment is made with L moments of a higher order to give more importance to such values.

Fit with L moments for the GVE

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

The theory of extreme values justifies and establishes that extreme data follow asymptotically one of the three following types of distributions called: Gumbel, Fréchet, and Weibull (Clarke, 1973; Stedinger *et al*., 1993; Coles, 2001). These three probabilistic models can be represented in a single one, called the *General of Extreme Values distribution* (GVE, by its acronym in Spanish), which has been widely recommended to model maximum annual flows (Q) and other extreme data (Hosking & Wallis, 1997; Papalexiou & Koutsoyiannis, 2013). The PDF of the GVE with a probability of non–exceedance [$F(Q) = p$] is:

$$F(Q) = \exp\left\{-\left[1 - \frac{k(Q-u)}{\alpha}\right]^{1/k}\right\} \quad \text{when } k \neq 0 \quad (1)$$

In the above expression, u, α and k are the locations, scale, and shape parameters of the GVE distribution. When $k = 0$ the Gumbel distribution is obtained, which is a straight line on the Gumbel–Powell probability paper (Chow, 1964), thus, the interval of the variable is: $-\infty < Q < \infty$. When $k > 0$ the distribution is Weibull which is a curve with concavity downwards and an upper limit, therefore: $-\infty < Q \leq u + \alpha/k$. Finally, if $k < 0$ the distribution is Fréchet, which is also a curve but with an upward concavity and a lower boundary, so that: $u + \alpha/k \leq Q < \infty$. The sought *predictions* (Q_{Tr}) are obtained with the inverse solution of Equation (1):

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

$$Q_{Tr} = u + \frac{\alpha}{k}\{1 - [-\ln(p)]^k\} \qquad \text{when } k \neq 0 \tag{2}$$

in which:

$$p = 1 - \frac{1}{Tr} \tag{3}$$

On the other hand, to estimate the fit parameters of PDFs used in hydrology, the L–moment method is perhaps the simplest and has become one of the most reliable procedures. This is because the L moments, designated λ, are linear combinations (Hosking & Wallis, 1997) of the weighted probability moments (β_r), which are not significantly affected by the dispersed values (*outliers*) of the sample. The first three L moments of a *sample* (l_1, l_2, l_3) and the coefficient L of skewness (t_3), are estimated through the unbiased estimator (b_r) of the β_r, as follows:

$$l_1 = b_0 \tag{4}$$

$$l_2 = 2 \cdot b_1 - b_0 \tag{5}$$

$$l_3 = 6 \cdot b_2 - 6 \cdot b_1 + b_0 \tag{6}$$

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

$$t_3 = \frac{l_3}{l_2} \quad (7)$$

The unbiased estimator of the β_r is (Hosking & Wallis, 1997):

$$b_r = \frac{1}{n}\sum_{j=r+1}^{n} \frac{(j-1)(j-2)\cdots(j-r)}{(n-1)(n-2)\cdots(n-r)} Q_j \quad (8)$$

where r = 0, 1, 2,… and Q_j are the sample data or records of available floods of size n, ordered from smallest to largest $(Q_1 \leq Q_2 \leq \cdots \leq Q_n)$. The following equations allow estimating the three fit parameters of the GVE (Stedinger *et al*., 1993; Hosking & Wallis, 1997; Rao & Hamed, 2000):

$$k \cong 7.8590 \cdot c + 2.9554 \cdot c^2 \quad (9)$$

being:

$$c = \frac{2}{3+t_3} - 0.63093 \quad (10)$$

$$\alpha = \frac{l_2 \cdot k}{(1-2^{-k}) \cdot \Gamma(1+k)} \quad (11)$$

$$u = l_1 - \frac{\alpha}{k}[1 - \Gamma(1+k)] \quad (12)$$

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

For the estimation of the Gamma function Γ(ω) the Stirling formula (Davis, 1972) was used:

$$\Gamma(\omega) \cong e^{-\omega} \cdot \omega^{\omega-\frac{1}{2}} \cdot (2\pi)^{1/2} \cdot F1 \tag{13}$$

being:

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

Fit with LH moments of the GVE

Wang (1997b) proposes the L moments of higher-order (λ^η), known as "LH-moments" of *higher* which means higher order, as a generalization of the L moments, which allow a better characterization of the right tail at the PDF, and of big events in the data. Wang (Wang, 1997a; Wang, 1997b) developed two procedures to adjust the GVE distribution with the method of LH moments, the first is similar to that of L moments and is the following:

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

$$c^{\eta} = \frac{(\eta+2)\cdot\beta_{\eta+1} - (\eta+1)\cdot\beta_{\eta}}{(\eta+3)\cdot\beta_{\eta+2} - (\eta+1)\cdot\beta_{\eta}} - \frac{\ln(\eta+2) - \ln(\eta+1)}{\ln(\eta+3) - \ln(\eta+1)} \quad (14)$$

$$k^{\eta} \cong a_1 \cdot c^{\eta} + a_2 \cdot (c^{\eta})^2 \quad (15)$$

The coefficients a_1 and a_2 vary with the order η of the LH moments, have the values of Table 1, and are applicable in the interval: $-0.50 \leq k^{\eta} \leq 0.50$.

Table 1. Coefficients a_1 and a_2 of Equation (15) and maximum absolute error (δ) of the shape parameter of the GVE distribution (Wang, 1997a).

η	a_1	a_2	$\lvert\delta\rvert$
1	11.9082	2.7787	$3.4\cdot10^{-4}$
2	15.9316	2.7301	$1.8\cdot10^{-4}$
3	19.9455	2.7072	$1.1\cdot10^{-4}$
4	23.9546	2.6936	$7.7\cdot10^{-5}$

$$\alpha^{\eta} = \frac{k^{\eta}\left[(\eta+2)\,\beta_{\eta+1} - (\eta+1)\cdot\beta_{\eta}\right]}{\Gamma(1+k^{\eta})\cdot\left[(\eta+1)^{-k^{\eta}} - (\eta+2)^{-k^{\eta}}\right]} \quad (16)$$

$$u^{\eta} = (\eta+1)\cdot\beta_{\eta} - \frac{\alpha^{\eta}}{k^{\eta}}\left[1 - \Gamma(1+k^{\eta})\cdot(\eta+1)^{-k^{\eta}}\right] \quad (17)$$

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

The probability-weighted moments β_1 to β_6 are calculated with Equation (8) and the Gamma functions with Equation (13). Campos–Aranda (2016) exposes the application of the LH moments in the observed floods of Hydrological Region No. 10 (Sinaloa), Mexico.

Empirical envelope curves

The extreme maximum flow (Q_{max}) that is used as the upper limit of the Weibull distribution that is part of the mixed GVE, is estimated based on the *empirical envelope curves* that are constructed for the homogeneous region analyzed and in theory, the third envelope curve constructed has a probability of exceedance equal to zero. It is indicated "in theory" because such a curve comes from observed data, and therefore, it cannot be guaranteed that it will not be exceeded. This approach is much simpler and more practical than estimating the maximum probable precipitation in each basin and with it its Maximum Probable Flood.

The procedure suggested by Guse *et al*. (2010a) is based on the concept of enveloping curves and therefore consists of drawing on a logarithmic paper the specific flows ($q_m = Qm/A$) of each basin in the ordinate and their respective basin areas (A) in km^2 on the abscissa.

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

Logically, Qm is the average annual flow in m^3/s. Such data are represented by linear regression of the following type:

$$\log(q_m) = b + m \cdot \log(A) \tag{18}$$

Where b is the ordinate to the origin and m is the slope of the line. Then, the specific flows (q_{max}) of the maximum annual flows observed (Q_{max}) in each record or basin are drawn on the same graph and a line is drawn parallel to the one defined by Equation (18), at the upper point to define the curve regional envelope. Later, Guse *et al*. (2010a) drew two upper envelope curves previously calculated, the one for Germany and the one for Europe. As both extreme curves showed slopes similar to that defined by Equation (18), the second one was accepted to define the Q_{max} in the studied region of southeastern Germany. Other details of the procedure can be found in Guse, Thieken, Castellarin and Merz (2010b).

Generation of synthetic series

The GVE$_{obs}$ distribution was adjusted to each available record in the analyzed homogeneous region, based on equations (9) to (12) and with these parameters, synthetic sequences of 1 500 values each were generated, since it is intended to estimate with greater reliability the prediction of *Tr* equal to a thousand years. The above is based on Equation (2), making *p* equal to a random number (u_{m+1}) with a uniform distribution in the interval from zero to one.

Taking into account that many random series must be generated from each available or observed record and that each one will have 1 500 values, it is necessary to use an efficient generation algorithm, whose cycle length is quite large. Therefore, the one exposed by Metcalfe (1997) was used, with an elementary restriction to avoid the occurrence of a zero due to rounding when applying the "mod" instruction of the *congruential mixed pseudo-random generator* (Wichmann & Hill, 2006). This algorithm begins by defining values in the range of one to thirty thousand for Io, *Jo*, and *Ko*, then the following recursive equations are applied:

$$I_{m+1} = 171 \cdot I_m + 100 \cdot (\text{mod } 30269) \tag{19}$$

$$J_{m+1} = 172 \cdot J_m + 150 \cdot (\text{mod } 30307) \tag{20}$$

$$K_{m+1} = 170 \cdot K_m + 200 \cdot (\text{mod } 30323) \tag{21}$$

$$S_{m+1} = I_{m+1}/30269 + J_{m+1}/30307 + K_{m+1}/30323 \tag{22}$$

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

$$u_{m+1} = S_{m+1} - \text{parte entera de } S_{m+1} \quad (23)$$

This algorithm has a cycle length of the product order the modules, that is, $2.78 \cdot 10^{13}$. To show how the instruction "mod = [(*i*/*k*) – integer part of (*i*/*k*)]·*k*" works, it is clarified that *i* is the quantity to the left of the parentheses and *k* its modulus. For the following example, *k* = 5095, Io = 193 and the recursive equation is: $I_{m+1} = 128 \cdot I_m + 1569 \cdot (\text{mod } 5095)$, having thus:

I_1 = 128·193 + 1 569 (mod 5095) = 2 6273·(mod 5095) = 798

I_2 = 128·798 + 1 569 (mod 5095) = 10 3713·(mod 5095) = 1 813

I_3 = 128·1813 + 1 569·(mod 5095) = 23 3633·(mod 5095) = 4 358

The sequence of random numbers u_i that has been generated corresponds to the value in rectangular parentheses of the "mod" instruction, which are: 0.156624141, 0.355839057, and 0.85534838. Various aspects of random number generators can be found in Wichmann and Hill (2006).

Synthetic sequence selection

To ensure similarity between the GVE distributions of each available record (GVE_{obs}) and from the synthetic record of 1 500 values (GVE_{sim}) to be selected, the two restrictions established by Guse *et al*. (2010a) were respected. The first is related to your maximum flows of Tr = 1 500 years ($Q_{1\ 500}$), which should be approximately equal; preferably higher in the synthetic sequence. The second defines a minimum of ten values with Tr greater than 150 years, so its probability of non–exceedance (p) must be greater than 0.993333, according to Equation (3).

Due to the importance that Tr = 1 500 years' flow has in estimating the design floods or predictions of Tr > 500 years (selected inflection point), it was considered necessary to verify its magnitude by contrast with the *median value* obtained by applying the PDFs established under precept and those of almost universal application. The first is The Log–Pearson type III (LP3), the GVE, and the Generalized Logistics (LOG); the second were: The Log-Normal of 3 fit parameters (LN3), the Generalized Pareto (PAG), and the Wakeby (WAK), of five fit parameters. Except for the LP3 distribution, which was adjusted with two methods of moments, one in the logarithmic domain and the other in the real one (Bobée & Ashkar, 1991), the rest were adjusted with the method of L moments (Hosking & Wallis, 1997). Once the synthetic sequence which meets both restrictions is selected, the GVE distribution is adjusted with the method of L moments (equations (9) to (12)) and with LH moments (equations (14) to (17)); the one which leads to the lowest standard error of fit (EEA) is adopted, where the parameters will be called u_{sim}, a_{sim} and k_{sim}. These values define the lower portion of the mixed GVE, which is concave upward.

Standard error of fit

It is the most common index (Chai & Draxler, 2014) for contrasting PDFs to real data, it was established in the mid–the 1970s (Kite, 1977) and has been applied in Mexico using the empirical Weibull formula (Benson, 1962). Now it will be applied using Cunnane's formula (Equation (25)), which according to Stedinger (2017) leads to approximately unbiased probabilities of non–exceedance (*p*) for many PDFs. The expression for the standard error of fit (*EEA*) is:

$$EEA = \left[\frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{(n-np)}\right]^{1/2} \quad (24)$$

Q_i is the maximum annual flows ordered from lowest to highest, whose number is *n* and $\hat{Q}_i$ the maximum estimated flows, for the probability estimated with Equation (25) and the PDF that is contrasted. *Np* is the number of fit parameters of the PDF, with five for the Wakeby and three for the rest of which will be applied.

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

$$p = \frac{i-0.40}{n+0.20} \qquad (25)$$

Mixed GVE fit parameters

The lower part of the mixed GVE reaches the inflection point with Tr = 500 years and corresponds to the GVE_{sim} adjusted with L or LH moments. The upper part is designated GVE_{cls} because it is a curve *with an upper limit* that begins at the inflection point and becomes asymptotic to the $Q_{\max}$ value; its fit parameters (u_{cls}, α_{cls}, and k_{cls}) are determined based on the equations that are established with the following three restrictions (Guse *et al*., 2010a). The first indicates that $Q_{\max}$ is obtained with the GVE_{cls} when p = 1 in Equation (2), obtaining this way:

$$Q_{\max} = u_{cls} + \frac{\alpha_{cls}}{k_{cls}} \qquad (26)$$

The second restriction establishes that the predictions of Tr = 500 years (p = 0.998) are the same with the GVE_{sim} and GVE_{cls} distributions, which, according to Equation (2), results in:

$$u_{sim} + \frac{\alpha_{sim}}{k_{sim}}\{1—\} = u_{cls} + \frac{\alpha_{cls}}{k_{cls}}\{1——\} \quad (27)$$

Finally, the third constraint indicates that at the inflection point the slopes of the two GVE_{sim} and GVE_{cls} distributions are equal. The expression for the slope of Equation (2) is:

$$\frac{dQ}{dp} = \frac{\alpha}{p}[-\ln(p)]^{k-1} \quad (28)$$

obtaining this way:

$$\alpha_{sim}[-\ln(0.998)]^{k_{sim}-1} = \alpha_{cls}[-\ln(0.998)]^{k_{cls}-1} \quad (29)$$

The numerical solution procedure consisted of clearing u_{cls} from Equation (26) and replacing it in 27, to clear a_{cls} in k_{cls} function only; this expression was designated a_1. On the other hand, on Equation (29), a_{cls} was cleared and called a_2. Then, by trials of k_{cls}, a_1 and a_2 were equalized, accepting a difference of less than 0.05. Finally, based on Equation (26), u_{cls} is evaluated.

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

Records of floods processed

In this study, the hydrometric information processed by Campos–Aranda (2014) from Hydrological Region No. 10 (Sinaloa), Mexico was used. Its general data on the extent of maximum annual flows record (*n*), basin areas (*A*), average annual flows (*Qm*), and maximum annual flows (Q_{max}) are shown in Table 2, for the 21 hydrometric stations analyzed.

Table 2. General data of the annual flood records of the 21 indicated hydrometric stations of the Hydrological Region No. 10 (Sinaloa), Mexico.

1	2	3	4	5	6	7	8
No.	**Name:**	***n***	***A*** (km²)	***Qm*** (m³/s)	q_m (l/s/km²)	Q_{max} (m³/s)	q_{max} (l/s/km²)
1	Huites	51	26 057	3 328.333	127.7	15 000.0	575.7
2	San Francisco	33	17 531	1 724.636	98.4	6 640.0	378.8
3	Santa Cruz	52	8 919	1 037.615	116.3	7 000.0	784.8
4	Jaina	56	8 179	1 020.786	124.8	6 991.0	854.7
5	Palo Dulce	21	6 439	1 129.238	175.4	6 800.0	1 056.1
6	Ixpalino	45	6 166	1 198.978	194.4	6 200.0	1 005.5
7	La Huerta	28	6 149	945.107	153.7	1 931.0	314.0

8	Chinipas	24	5 098	883.083	173.2	2 683.0	526.3
9	Tamazula	32	2 241	596.875	266.3	2 289.0	1 021.4
10	Naranjo	45	2 064	633.311	306.8	3 093.0	1 498.5
11	Acatitán	43	1 884	813.256	431.7	4 600.0	2 441.6
12	Guamuchil	32	1 645	702.344	427.0	3 507.0	2 131.9
13	Choix	38	1 403	348.974	248.7	1 700.0	1 211.7
14	Badiraguato	26	1 018	1 224.346	1 202.7	9 245.0	9 081.5
15	El Quelite	33	835	479.091	573.8	1 743.0	2 087.4
16	Zopilote	56	666	351.857	528.3	1 030.0	1 546.5
17	Chico Ruiz	19	391	205.737	526.2	476.0	1 217.4
18	El Bledal	56	371	289.000	779.0.	1 576.0	4 248.0
19	Pericos	30	270	250.800	928.9	654.0	2 422.2
20	La tina	24	254	104.958	413.2	778.0	3 063.0
21	Bamícori	33	223	189.182	848.3	650.0	2 914.8

Symbology:

n number of data in the record.

A basin area.

Qm average annual flow.

q_m average specific flow (Qm/A).

Q_{max} maximum flow of the record.

q_{max} maximum specific flow (Q_{max}/A).

Initially, it was decided to apply the FFA with the mixed GVE method, exclusively in the seven largest registers (43 ≤ n ≤ 56), which correspond to the hydrometric stations: Huites, Santa Cruz, Jaina, Naranjo, Acatitán, Zopilote, and El Bledal.

Verification of the randomness of the records

For the results of the FFA to be reliable, the maximum annual flows' records to be processed must be generated by a stationary random process, which implies that it has not changed over time. Therefore, the flood registers must be integrated by independent data, free of deterministic components.

To verify the above, seven statistical tests were applied, one general, the Von Neumann Test, and six specifics: two for persistence (Anderson and Sneyers), two for trend (Kendall and Spearman), one for the change in the mean (Cramer) and the last one related to variability (Bartlett). These tests can be consulted in WMO (1971), and Machiwal and Jha (2012).

All the cited tests were applied with a significance level (α) of 5% and six of them indicate that the selected records are random. The Bartlett

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

test detects excess variability, due to the presence of scattered values, in four of the records.

Results and their analysis

Empirical envelope Curves

On a logarithmic paper, the average specific flows (q_m) of column 6 in Table 2 were drawn, as well as the maximum specific flows (q_{max}) from column 8. The above is shown in Figure 2. Afterwards, the data were charted with a line of least squares (Equation (18)) from the residuals, or simple linear regression, the q_m values, and the following expression was obtained:

$$\log(q_m) = 4.0449940 - 0.4703735 \cdot \log(A) \quad (30)$$

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

Without eliminating any data, it was obtained that its correlation coefficient was 0.89417, with 0.1511 as the standard error of the estimate.

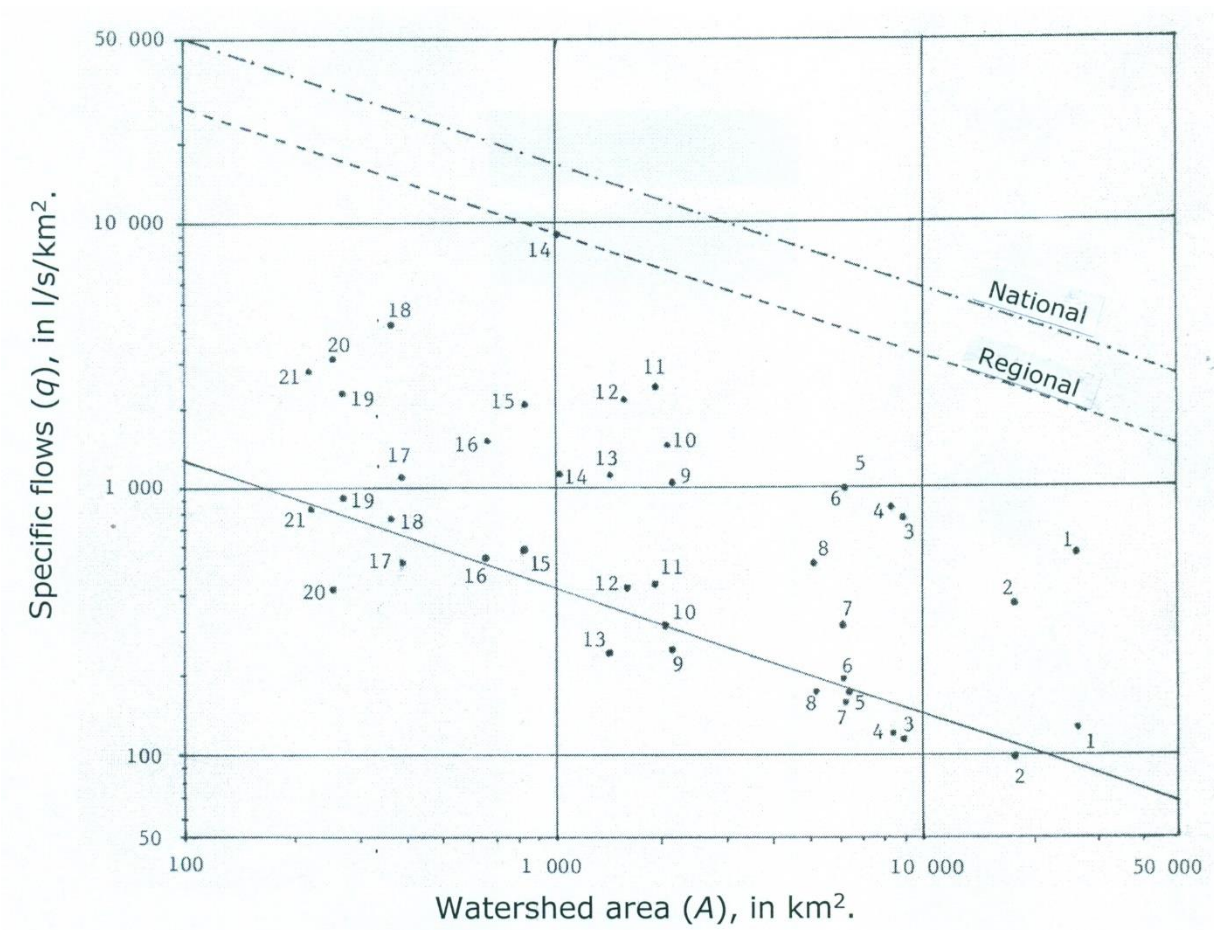

Figure 2. Empirical envelope curves of mean annual flow, maximum regional flow, and extreme maximum or national curve in Hydrological Region No. 10 (Sinaloa), Mexico.

Next, the most dispersed point upwards of the q_{max} values was sought; it was number 14 corresponding to the Badiraguato hydrometric station with q_{max} = 9081.5 l/s/km^2 and A = 1018 km^2. These two values were taken to Equation (30), keeping their slope (m) to solve for the new ordinate to the origin (b), which was: 5.3729224; then the equation of the regional line is:

$$\log(q_{max}) = 5.3729224 - 0.4703735 \cdot \log(A) \quad (31)$$

Ramírez–Orozco, Gómez–Martínez and Campos-Aranda (2005) defined that the coefficient (C_L) of the national Lowry envelope was of 7200. So for the Badiraguato station a national q_{max} of:

$$q_{max} = \frac{C_L}{(A+259)^{0.85}} = \frac{7200}{(1018+259)^{0.85}} = 16.48429 \cdot \mathrm{m}^3/\mathrm{s}/\mathrm{km}^2 \quad (32)$$

Taking again this value of q_{max} = 16484.3 l/s/km^2 and A = 1018 km^2 to Equation (30), to solve for the new b while preserving the slope (m), the envelope that will define the maximum specific extreme flows (q_{max}) was obtained for of Hydrological Region No. 10 (Sinaloa), Mexico, this is:

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

$$\log(q_{\max}) = 5.6318354 - 0.4703735 \cdot \log(A) \quad (33)$$

Extreme maximum flows

Applying Equation (33) to the basin area data (A) in km^2 (column 4 in Table 2) of the seven selected hydrometric stations, the respective value of q_{max} in l/s/km^2 was obtained, where the product by A and between thousand defines the Q_{max} values in m^3/s; These corresponded to Huites 93460.0, Santa Cruz 52969.7, Jaina 50594.7, Naranjo 24400.2, Acatitán 23249.0, Zopilote 13403.6 and El Bledal 9832.0.

Review of flows of *Tr* = 1 500 years

Table 3 concentrated the predictions of six *Tr* that varied from 50 to 1 500 years, obtained with each of the six applied PDFs. The *Tr* of 50 and 100 years allowed to verify the similarity between predictions of each PDF.

It was observed that all the predictions occurred within their order of magnitude, but those of Jaina and El Bledal stood out. The standard errors of fit (*EEA*) also showed a similarity in orders of magnitude, in the seven hydrometric stations.

Table 3. Predictions (m^3/s) of the indicated return periods, in the seven hydrometric stations processed in Hydrological Region No. 10 (Sinaloa), Mexico.

Hydrometric station	**FDP**	***EEA*** **(m^3/s)**	**Return periods (*Tr*) in years**				
			50	**100**	**500**	**1 000**	**1 500**
Huites	LN3	784.0	14640	19409	34574	43235	48999
Median Value	LP3	957.1	15613	21776	45130	60942	72442
of *Tr* = 1 500	GVE	930.0	14166	19665	41312	56610	*68008*
years = 59 744	LOG	984.3	13837	19452	42767	60040	73225
m^3/s	PAG	781.4	14355	19056	34919	44673	51455
	WAK	798.4	14355	19058	34930	44692	51480
Santa Cruz	LN3	283.6	4454	5707	9428	11433	12733
Median Value	LP3	299.3	4447	5583	8721	10299	11288
of *Tr* = 1 500	GVE	270.8	4410	5883	11118	14497	*16903*
years = 14 818	LOG	277.5	4335	5891	11843	15948	18973
m^3/s	PAG	306.5	4387	5534	8864	10648	11807
	WAK	309.4	4160	5963	13991	20289	25236
Jaina	LN3	188.9	4547	5999	10562	13143	14852

Median Value of *Tr* = 1 500 years = 16 943 m^3/s	LP3	194.2	4504	5969	10785	13644	15589
	GVE	194.3	4 419	6 101	12 614	17 148	*20 501*
	LOG	213.4	4 320	6 046	13 117	18 294	22 222
	PAG	205.6	4 466	5 884	10 555	13 364	15 296
	WAK	190.3	4 495	6 102	11 870	15 618	18 297
Naranjo Median Value of *Tr* = 1 500 years = 8 964 m^3/s.	LN3	120.4	2 860	3 652	5 968	7 201	7 996
	LP3	121.4	2 491	2 954	4 010	4 450	4 702
	GVE	140.0	2 843	3 775	7 009	9 053	*10 494*
	LOG	152.4	2 798	3 792	7 522	10 054	11 904
	PAG	107.3	2 815	3 522	5 506	6 534	7 191
	WAK	112.8	2 820	3 548	5 629	6 726	7 433
Acatitán Median Value of *Tr* = 1 500 years = 8 847 m^3/s	LN3	150.5	3 440	4 263	6 537	7 689	8 416
	LP3	176.5	3 300	4 018	5 826	6 657	7 157
	GVE	142.2	3 459	4 422	7 458	9 218	*10 406*
	LOG	145.0	3 432	4 502	8 240	10 619	12 303
	PAG	173.9	3 353	4 005	5 588	6 301	6 727
	WAK	148.5	3 450	4 341	6 950	8 359	9 278
Zopilote Median Value of *Tr* = 1 500 years = 2 115 m^3/s	LN3	55.2	1 149	1 328	1 761	19 55	2 072
	LP3	42.0	1 051	1 158	1 352	1 415	1 448
	GVE	58.8	1 161	1 351	1 817	2 030	*2 157*
	LOG	73.9	1 181	1 428	2 152	2 544	2 800
	PAG	27.2	1 068	1 153	1 291	1 332	1 351
	WAK	122.9	1 246	1 648	3 135	4 128	4 846

El Bledal Median Value of *Tr* = 1 500 years = 3 139 m^3/s	LN3	36.8	1 089	1 357	2 117	2 512	2 764
	LP3	40.2	1 077	1 330	2 033	2 390	2 615
	GVE	31.9	1 090	1 404	2 446	3 077	*3 513*
	LOG	33.0	1 077	1 419	2 657	3 472	4 058
	PAG	46.0	1 069	1 295	1 887	2 174	2 351
	WAK	31.1	1 095	1 418	2 483	3 125	3 566

For the 1 500 years *Tr*, the *median value* of each hydrometric station was obtained and is cited after its name. Since each median flow of *Tr* = 1 500 years, resulted lower than the one obtained with the GVE distribution, the estimate with that probabilistic model is accepted (in italics in Table 3), which were the following in m^3/s: Huites 68 008, Santa Cruz 16 903, Jaina 20 501, Naranjo 10 494, Acatitán 10 406, Zopilote 21 57 and El Bledal 3 513.

Table 3 establishes a comparison between the predictions of the return period 1 500 years of the GVE and those of five distributions of conventional application (by rule or universality). Logically, due to space limitations, a comparison could not be made against estimates from regional methods.

GVE fits processed records

In Table 4, on the first line after the name of each hydrometric station, the values of the adjustment parameters of the GVE_{obs} distribution are presented, as well as its *EEA* and the *Tr* predictions for 100, 500, 1 000, and 1 500 years.

Table 4. Results of the adjustment of the GVE distribution with L moments to the available (*obs*) and simulated (*sim*) records and with LH moments to the latter, in the seven hydrometric stations processed in Hydrological Region No. 10 (Sinaloa), Mexico.

Fit parameters			***EEA***	**Return periods in years**			
Location	**Scale**	**Shape**	**(m³/s)**	**100**	**500**	**1 000**	**1 500**
Hydrometric station: Huites ($Q_{1\,500}$ = 68 498 m³/s; *nvs* = 19; LH1)							
1755.521	1 178.243	−0.445918	930.0	19 665	41 312	56 610	68 008
1564.128	1 735.921	−0.316563	1 276.5	19 604	35 284	44 911	51 601
1784.694	1 183.132	−0.475632	1 131.4	21 478	47 089	65 757	79 898
Hydrometric station: Santa Cruz ($Q_{1\,500}$ = 17 005 m³/s; *nvs* = 19; LH1)							
534.3951	451.5092	−0.362125	270.8	5 883	11 118	14 497	16 903
486.8800	613.6693	−0.252475	317.7	5 821	9 725	11 959	13 458
543.1981	449.7197	−0.394945	260.0.	6 410	12 653	16 829	19 856
Hydrometric station: Jaina ($Q_{1\,500}$ = 20 645 m³/s; *nvs* = 19; LH1)							
525.7384	382.4736	−0.432097	194.3	6 101	12 614	17 148	20 501
467.7769	555.4508	−0.306168	385.7	6 073	10 812	13 689	15 677

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

534.7980	383.3342	−0.462616	337.8	6 666	14 387	19 941	24 118
Hydrometric station: Naranjo (*$Q_{1\ 500}$* = 10 554 m³/s; *nvs* = 19; LH1)							
301.0952	310.0954	−0.342825	140.0	3 775	7 009	9 053	10 494
271.7218	414.7160	−0.237191	197.9	3 729	6 157	7 522	14 060
306.8684	308.5978	−0.375798	158.5	4 112	7 968	10 495	12 307
Hydrometric station: Acatitán (*$Q_{1\ 500}$* = 10 455 m³/s; *nvs* = 19; LH1)							
402.2712	447.9312	−0.264082	142.2	4 422	7 458	9 218	10 406
375.4133	565.8641	−0.172710	193.9	4 351	6 681	7 901	8 684
409.4722	445.2446	−0.296061	141.3	4 776	8 371	10 529	12 012
Hydrometric station: Zopilote (*$Q_{1\ 500}$* = 2 162 m³/s; *nvs* = 19; LH1)							
216.7843	218.5603	−0.051462	58.8	1 351	1 817	2 030	2 157
208.7049	256.2496	0.03286	42.1	1 303	1 649	1 792	1 875
220.6736	218.7255	−0.072403	21.4	1 415	1 937	2 181	2 329
Hydrometric station: El Bledal (*$Q_{1\ 500}$* = 3 531 m³/s; *nvs* = 19; LH1)							
166.6999	123.2785	−0.304373	31.9	1 404	2 446	3 077	3 513
157.2938	160.0594	−0.206169	65.0	1 385	2 176	2 606	2 887
168.8195	122.5516	−0.337195	49.7	1 520	2 759	3 537	4 084

Synthetic sequences and their GVE fits

With the algorithm defined by equations (19) to (23), synthetic sequences of 1 500 values were generated, to select one which had a maximum flow ($Q_{1\ 500}$) in m^3/s slightly higher than those defined by the GVE$_{obs}$ distribution, indicated in italics in Table 3 and whose number of values greater than (*nvs*) the probability $p = u_{m+1} = 0.993333$ was greater than ten. The aforementioned algorithm started with the following values: $Io = 1225$, $Jo = 4550$ and $Ko = 17840$ and in each new sequence of 1 500 values designated with the counter *ijk*, the initial values increased as follows: $Io = Io + 5 \cdot ijk$, $Jo = \text{Jo} + 10 \cdot ijk$, $Ko = Ko + 50 \cdot ijk$; Therefore, the first sequence generated ($ijk = 1$) of 1 500 values, starts with $Io = 1230$, $Jo = 4560$ and $Ko = 17890$.

In the second sequence generated, its initial values were $Io = 16517.52$, $Jo = 22873.42$, and $Ko = 16716.9$ and in it $nvs = 19$ random numbers (u_{m+1}) with probability greater than 0.993333 are obtained, whose maximum flow of such sequence of 1 500 values are cited in Table 4. This was the selected sequence and the GVE$_{sim}$ distribution was adjusted to it with the L moments method (equations (9) to (12)), where the adjustment parameters are shown in the second row after the name of the hydrometric station in Table 4. The GVE$_{sim}$ was also adjusted to this sequence with the LH moment method (equations (14) to (17)); The results are shown in the third line after the station name in Table 4.

Except for the Zopilote hydrometric station (whose k_{sim} changed from negative to positive), in the rest of the GVE$_{sim}$ adjustments to the synthetic sequences, the standard error of adjustment (*EEA*) increased concerning that of the observed record. On the other hand, all *EEA* of the

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

GVE_{sim} adjustments with LH moments were always lower in all synthetic sequences and, therefore, will be used to define the lower curve of the mixed GVEs, up to the point of inflection with Tr = 500 years.

The fit of the mixed GVE

Based on equations (26) to (29) and the numerical solution procedure described in the respective subsection, solved with the bisection method with initial values of k_{cls} of 0.0001 and 0.50. The fit parameters of the mixed GVE shown in Table 5 were obtained, making use of the extreme maximum flows ($Q_{\max}$) cited in said Table 5 and the fit parameters of the GVE_{sim} of the LH moment method, shown in Table 4 in the third row of each hydrometric station processed.

Table 5. Results of the fit of the mixed GVE distribution with upper limit (k > 0) using a numerical process at the inflection point of Tr = 500 years, in the seven hydrometric stations processed in Hydrological Region No. 10 (Sinaloa), Mexico.

Fit parameters			Return periods in years		
Location	Scale	Shape	500	1 000	1 500

Hydrometric station: Huites (Q_{max} = 93 460.0 m³/s)					
−880760.3	47 7383.7	0.490016	47 079	60 445	66 396
Hydrometric station: Santa Cruz (Q_{max} = 52 969.7 m³/s)					
−37336.8	11 720.7	0.129788	12 653	16 124	18 014
Hydrometric station: Jaina (Q_{max} = 50 594.7 m³/s)					
−65540.36	21 783.25	0.187568	14 387	18 804	21 133
Hydrometric station: Naranjo (Q_{max} = 24 400.2 m³/s)					
−30454.46	10 642.14	0.194006	7 968	10 038	11 124
Hydrometric station: Acatitán (Q_{max} = 23 249.0 m³/s)					
−24706.54	9 032.978	0.188362	8 371	10 194	11 154
Hydrometric station: Zopilote (Q_{max} = 13 403.6m³/s)					
−405.4854	413.0258	0.029910	1 937	2 172	2 307
Hydrometric station: El Bledal (Q_{max} = 9 832.0m³/s)					
−7134.746	2 389.239	0.140819	2 759	3 417	3 774

GVE_{obs} y GVE_{sim} predictions

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

Table 6 shows the predictions of both GVE distributions. It is indicated that due to how the 1 500 synthetic data of the GVE_{sim} fit and its complement with the GVE_{cls} were generated and selected, it is quite probable that, in general, the predictions of the mixed GVE are higher than those of the GVE_{obs} fit. Inspection of the values in Table 6 verifies the results numerically since only a prediction of *Tr* = 1 500 years at the Huites station exceeded the one obtained with the mixed GVE. Another important verification is the similarity that all the predictions of the low *Trs*, under 50 years of age, had.

Table 6. Predictions (m^3/s) of the GVE fit to the observed flows and the mixed GVE, in the seven hydrometric stations processed in Hydrological Region No. 10 (Sinaloa), Mexico.

Station	GVE	Return periods (*Tr*) in years							
		5	10	25	50	100	500	1 000	1 500
Huites	obs	4 271	6 321	10 114	14 166	19 665	41 312	56 610	68 008
	mixed	4 374	6 552	10 686	15 211	21 478	47 079	60 445	66 396
Santa Cruz	obs	1 434	2 104	3 258	4 410	5 883	11 118	14 497	16 903
	mixed	1 464	2 174	3 432	4 722	6 410	12 653	16 124	18 014
Jaina	obs	1 333	1 981	3 166	4 419	6 101	12 614	17 148	20 501
	mixed	1 365	2 053	3 345	4 745	6 666	14 387	18 804	21 133
Naranjo	obs	909	1 353	2 104	2 843	3 775	7 009	9 053	10 494
	mixed	929	1 399	2 218	3 044	4 112	7 968	10 037	11 124

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

Acatitán	obs	1 227	1 779	2 653	3 459	4 422	7 458	9 218	10 406
	mixed	1 250	1 833	2 782	3 680	4 776	8 371	10 193	11 154
Zopilote	obs	558	738	977	1 161	1 351	1 817	2 030	2 157
	mixed	567	755	1 008	1 207	1 415	1 937	2 172	2 307
El Bledal	obs	401	565	834	1 090	1 404	2 446	3 077	3 513
	mixed	408	582	874	1 160	1 520	2 759	3 417	3 773

On the other hand, in Table 7 a contrast is established, in the *Tr* older than 100 years, between the predictions of the mixed GVE and those of the PDF that led to the lowest standard error of fit (*EEA*), in Table 3; as well as with the maximum prediction (Q_{max}), in each *Tr* analyzed, obtained with any of the six applied PDFs, also shown in Table 3.

Table 7. The contrast of predictions (m^3/s) of high *Tr* of the mixed GVE with those of the PDF of lower *EEA* and the maximum prediction of the six applied PDFs, in the seven hydrometric stations processed in Hydrological Region No. 10 (Sinaloa), Mexico.

Station:	**Type**	**Return period (*Tr*) in years**			
	FDP	**100**	**500**	**1 000**	**1 500**
Huites	GVE_{mix}	21 478	47 079	60 445	66 396
	PAG	19 056	34 919	44 673	51 455
	Q_{max}	21 776	45 130	60 942	72 442

Santa Cruz	GVE_{mix}	6 410	12 653	16 124	18 014
	GVE	5 883	11 118	14 497	16 903
	Q_{max}	5 963	13 991	20 289	25 236
Jaina	GVE_{mix}	6 410	12 653	16 124	18 014
	LN3	5 999	10 562	13 143	14 852
	Q_{max}	6 102	13 117	18 294	22 222
Naranjo	GVE_{mix}	4 112	7 968	10 037	11 124
	PAG	3 522	5 506	6 534	7 191
	Q_{max}	3 792	7 522	10 054	11 904
Acatitán	GVE_{mix}	4 776	8 371	10 193	11 154
	GVE	4 422	7 458	9218	10 406
	Q_{max}	4 502	8 240	10 619	12 303
Zopilote	GVE_{mix}	1 415	1 937	2 172	2 307
	PAG	1 153	1 291	1 332	1 351
	Q_{max}	1 648	3 135	4 128	4 846
El Bledal	GVE_{mix}	1 520	2 759	3 417	3 773
	WAK	1 418	2 483	3 125	3 566
	Q_{max}	1 419	2 657	3 472	4 058

At the Huites station, the only prediction of the mixed GVE that exceeded the Q_{max} values was that of Tr = 500 years, but all of them

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

exceed the best fit. The same happened at the Santa Cruz, Jaina, and El Bledal stations, but at *Tr* = 100 years. In the Naranjo and Acatitán stations, the predictions of the mixed GVE exceeded the Q_{max} in the first two *Tr* analyzed and were higher than all the predictions of PDF with the lower *EEA*. Finally, at the Zopilote station, no mixed GVE prediction exceeded Q_{max}; but all exceed those of the best–fit PDF.

Finally, an extrapolation contrast of the exposed method was carried out, which allowed to verify again that the upper part of the mixed GVE prevents the unrealistic increase of the predictions. It is observed in Table 8 that the prediction of *Tr* = 5 000 years of the mixed GVE was close to the lower value in the Huites and Santa Cruz stations, it was close to the upper value in Naranjo and Acatitan and was intermediate between the lower and upper values in the rest of the hydrometric stations. Also shown in Table 8, is the extreme maximum flow, for comparison purposes.

Table 8. The contrast of predictions (m^3/s) of *Tr* = 5 000 years was obtained as the lower and upper values of the six adjusted PDFs (Table 2) and the mixed GVE, in the seven hydrometric stations processed in Hydrological Region No. 10 (Sinaloa), Mexico.

Type of Prediction	Huites	Santa Cruz	Jaina	Naranjo	Acatitán	Zopilote	El Bledal
lower	69 606	14 521	20 925	5 425	8 031	1 397	3 361

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

higher	132 090	48 342	39 591	19 611	18 978	7 796	6 469
GVE_{mix}	78 459	23 072	27 090	13 890	13 608	2 700	4 718
$Q_{\max}$	93 460	52 970	50 595	24 400	23 249	13 404	9 832

Conclusions

A novel method of Flood Frequency Analysis, suggested by Guse *et al*. (2010a), was exposed in detail, which theoretically increases the accuracy of the predictions of high return periods (500 and 1 000 years), due to the incorporation, in a mixed GVE distribution, of additional regional information and synthetic generated with the available records.

The method works initially at the regional level based on the concept of envelope curves, to define a line of specific average annual flow and subsequently establish two regional parallel curves, one of the maximum annual flows and the other of extreme maximum flows; in the latter, the probability of exceedance is null. The extreme maximum flow that is estimated with that envelope curve is used as the upper limit of a GVE distribution, which begins at an inflection point of Tr = 500 years. This

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

upper part of the mixed GVE avoids an unreal increase in predictions, with some PDF applied.

The latter was verified in Table 7, where it was observed that no 1 000-year prediction exceeded the maximum flow ($Q_{\max}$) estimated with any of the six PDFs applied and shown in Table 3. At the Huites and Acatitán stations, the estimated flood of Tr = 1 000 years approached the $Q_{\max}$ and in the Zopilote and Santa Cruz stations it was much lower.

On the other hand, the method obtains additional information from the available flood register, by generating synthetic sequences of 1 500 values and selecting one of them, to form the lower part of the mixed GVE. Designated GVE_{sim} reaches the inflection point and is fitting with the L and LH moment methods, to adopt the one with the smallest standard error of fit.

Table 6 shows seven predictions (5 ≤ Tr ≤ 1 000 years) obtained with the GVE adjusted to the available flood register and with the mixed GVE. All estimates of the mixed GVE were higher and therefore more critical or dangerous; but at the same time more reliable, due to the conceptual approaches of the described method.

Due to the advantages stated, the described method is recommended to be applied extensively in other regions of the country in estimating the Design Floods of the various hydraulic works that are built or revised in Mexico.

Acknowledgment

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

We are grateful for the observations and corrections suggested by the anonymous referee H, which made it possible to make the described procedure more explicit and to correct some involuntary but important omissions.

References

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.

Botero, B. A., & Francés, F. (2010). Estimation of high return period flood quantiles using additional non-systematic information with upper bounded statistical models. *Hydrology and Earth System Sciences*, 14(12), 2617-2628.

Campos-Aranda, D. F. (2016). Ajuste de las distribuciones GVE, LOG y PAG con momentos L de orden mayor. *Ingeniería. Investigación y Tecnología*, 17(1), 131-142.

Campos-Aranda, D. F. (2014). 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.

Clarke, R. T. (1973). Chapter 5: The estimation of floods with given return period. In: *Mathematical models in hydrology* (pp. 130-146) (Irrigation and Drainage Paper 19). Rome, Italy: FAO.

Coles, S. (2001). Chapter 3: Classical Extreme Value Theory and Models. In: *An introduction to statistical modeling of extreme values* (pp. 45-73). London, England: Springer-Verlag.

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

Chow, V. T. (1964). 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. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (9th printing) (pp. 253-296). New York, USA: Dover Publications.

Francés, F. (1998). Using the TCEV distribution function with systematic and non-systematic data in a regional flood frequency analysis. *Stochastic Hydrology and Hydraulics*, 12(4), 267-283.

Guse, B., Hofherr, Th., & Merz, B. (2010a). Introducing empirical and probabilistic regional envelope curves into a mixed bounded distribution function. *Hydrology and Earth System Sciences*, 14(12), 2465-2478.

Guse, B., Thieken, A. H., Castellarin, A., & Merz, B. (2010b). Deriving probabilistic regional envelope curves with two pooling methods. *Journal of Hydrology*, 380(1-2), 14-26.

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

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.

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.

Merz, R., & Blöschl, G. (2008). Flood frequency hydrology: 1. Temporal, spatial and causal expansion of information. *Water Resources Research*, 44(8), 1-17.

Metcalfe, A. V. (1997). Chapter 2: Probability distributions and Monte Carlo simulation. Appendix 2: Random number generation. In: *Statistics in Civil Engineering* (pp. 7-38, 319-320). London, England: Arnold Publishers.

Meylan, P., Favre, A. C., & Musy, A. (2012). Chapter 1: Introduction. In: *Predictive hydrology. A frequency analysis approach* (pp. 1-13). Boca Raton, USA: CRC Press.

Ouarda, T. B. M. J. (2017). Regional flood frequency modeling. In: Singh, V. P. (ed.). *Handbook of applied hydrology* (2nd ed.) (pp. 77.1-77.8). New York, USA: McGraw-Hill Education.

Papalexiou, S. M., & Koutsoyiannis, D. (2013). Battle of extreme value distributions: A global survey on extreme daily rainfall. *Water Resources Research*, 49(1), 187-201.

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., Gómez-Martínez, J. F., & Campos-Aranda, D. F. (2005). Actualización de las envolventes regionales de gastos máximos para la república mexicana. *Ingeniería Hidráulica en México*, 20(1), 99-108.

Rao, A. R., & Hamed, K. H. (2000). Chapter 1: Introduction. In: *Flood frequency analysis* (pp. 1-21). 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.

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

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Chapter 18: Frequency Analysis of Extreme Events. In: Maidment, D. R.

Tecnología y ciencias del agua, ISSN 2007-2422, 13(2), DOI: 10.24850/j-tyca-2022-02-06

(ed.). *Handbook of hydrology* (pp. 18.1-18.66). New York, USA: McGraw-Hill, Inc.

Wang, Q. J. (1997a). Using higher probability weighted moments for flood frequency analysis. *Journal of Hydrology*, 194(1-4), 95-106.

Wang, Q. J. (1997b). LH moments for statistical analysis of extreme events. *Water Resources Research*, 33(12), 2841-2848.

Wichmann, B. A., & Hill, I. D. (2006). Generating good pseudo-random numbers. *Computational Statistics & Data Analysis*, 51(3), 1614-1622.

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.