2021, Instituto Mexicano de Tecnología del Agua

DOI: 10.24850/j-tyca-2021-02-10

Articles

Fit of the two-component extreme value (TCEV) distribution through of maximum likelihood

Ajuste de la distribución de valores extremos de dos componentes (TCEV) por medio de máxima verosimilitud

Daniel Francisco Campos-Aranda[1]

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

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

Abstract

The annual record of floods, in many medium and large basins of our country and the world, is made up of events generated by phenomena which are physically different. For example, many floods originate from local storms, and a small portion is generated by cyclonic rains with a wide coverage and long duration, which generate extraordinary floods. The TCEV (two-component extreme value) distribution with four fitting parameters has been proposed for this type of records. TCEV has a

2021, Instituto Mexicano de Tecnología del Agua

theoretical basis that allows an approximate interpretation for two flood generation mechanisms, and it is also capable of reproducing the real variability of the asymmetry coefficient. This paper details its genesis and the fitting method by maximum likelihood, according to two numerical versions: (1) successive substitution and (2) objective function maximization. Six flood records were processed, the amplitude of which varied from 31 to 72 data, with three to six outliers or floods values that depart from the general trend. The predictions of the TCEV model, fitted with both numerical methods, are compared with those obtained using the standard application distributions (LP3, GVE and LOG) and the Wakeby distribution. Accepting the standard error of fit as a selection criterion, it follows that TCEV distribution is the best option in two of the six records processed. Lastly, as a conclusion, the systematic application of the TCEV distribution is suggested, using both numerical methods, in records with two mixed populations.

Keywords: TCEV distribution, Poisson process, maximum likelihood, Rosenbrock algorithm, standard error of fit, homogeneity and stationarity, prediction contrast.

Resumen

En muchas cuencas medianas y grandes de nuestro país y del mundo, su registro anual de crecientes está integrado por eventos que fueron generados por fenómenos físicamente diferentes. Por ejemplo, muchas crecientes se originan con tormentas locales y una porción pequeña se debe a lluvias ciclónicas de amplia cobertura y duración, que generan crecientes extraordinarias. Para este tipo de registros se propone la

2021, Instituto Mexicano de Tecnología del Agua

distribución TCEV (*two-component extreme value*) con cuatro parámetros de ajuste, la cual tiene una base teórica que permite una interpretación aproximada para dos mecanismos de generación de las crecientes y que además es capaz de reproducir la variabilidad real del coeficiente de asimetría. En este trabajo se detalla su génesis y el método de ajuste por máxima verosimilitud, según dos versiones numéricas: (1) sustitución sucesiva y (2) maximización de la función objetivo. Se procesaron seis registros de crecientes, cuya amplitud varió de 31 a 72 datos, con tres a seis valores dispersos o crecientes que se apartan de la tendencia general. Se contrastan las predicciones del modelo TCEV, ajustado con cada método numérico, contra las obtenidas con las distribuciones de aplicación establecida bajo precepto (LP3, GVE y LOG) y la Wakeby. Aceptando el error estándar de ajuste como criterio de selección, se deduce que la distribución TCEV es la mejor opción en dos de los seis registros procesados. Por último, se formulan las conclusiones, las cuales sugieren la aplicación sistemática de la distribución TCEV con ambos métodos numéricos en los registros que presentan dos poblaciones mezcladas.

Palabras clave: distribución TCEV, proceso de Poisson, máxima verosimilitud, algoritmo de Rosenbrock, error estándar de ajuste, homogeneidad y estacionariedad, contraste de predicciones.

Received: 26/09/2019

Accepted: 28/06/2020

2021, Instituto Mexicano de Tecnología del Agua

Introduction

Flood Frequency Analysis (FFA) Overview

Flood Frequency Analysis (FFA) is one of the basic estimates of Surface Hydrology, since it allows the planning, design and review of all the waterworks required by society, being these for use like reservoirs, or for protection like the embankments, channelings, control dams, bridges, urban drainage works, etc. FFA tries to define the relationship between the maximum flows of annual floods and their probability of being exceeded. Selected such probability or risk for *Design Flood*, its value should be estimated as accurately as possible, since a default error leads to an increase in the adopted risk, and an excess error leads to an increase in the cost of the project (Botto, Ganora, Laio, & Claps, 2014).

The most accurate procedure for performing an FFA is to represent the available record of annual maximum flows by a probability distribution function (PDF) and based on such a probabilistic model, make the inferences sought or *predictions*. This technique can lead to inaccurate estimates mainly due to the following four factors: (1) measurement errors in data; (2) finite amplitude of the available record; (3) presence of different mechanisms for generating the floods and (4) choice of PDF (Merz & Blöschl, 2008).

Francés-García (1995) states that historically the evolution of the FFAs has gone through three stages: (I) search for and fit of a better PDF to the annual maximum flows until the 1960s; (II) improved methods for estimating PDF fitting parameters during the 1970s and (III) increase in estimate accuracy, since the 1980s, by increasing the maximum flow records, through regional analysis and/or incorporation of flood historical information.

A fourth evolutionary stage of FFAs, beginning in this 21st century, refers to the processing of non-stationary records, because they show trends and/or increase or decrease in variability, as a result of physical changes in basins or climate change (Khaliq, Ouarda, Ondo, Gachon, & Bobée, 2006; Katz, 2013; López-de-la-Cruz & Francés, 2014; Prosdocimi, Kjeldsen & Svensson, 2014; Campos-Aranda, 2018).

FFA Based on Mixed PDFs

Since the middle of the last century, the interest for the presence of two statistical properties observed in the records of maximum annual flows raised; the first of them was called "*dog leg effect*", named so by Potter (1958), and it is observed at the moment of drawing data on the probability paper and finding a sudden slope change, due to the presence of outliers. The second property was named "phenomenon of separation"; it was studied and documented by Matalas, Slack and Wallis (1975), who drew on the x-axis the average values of the asymmetry or bias coefficients, and on the y-axis their respective standard deviations from the synthetic records generated with the most common PDFs. By plotting the points in the available maximum flow historical records in each of the 14 hydrologically homogeneous regions of the USA; PDFs were observed to be below actual points, thereby indicating that such PDFs could not represent flood sample variability.

The above-mentioned PDF deficiencies identified by hydrologists used in late 1970s, their solution was proposed through the *mixed probabilistic models*, which consider the different origins in the floods of a record. Mixed models are applied under two different approaches: (a) independent and separable floods and (b) independent floods which are mixed. The classic example of the first approach is the model suggested by Waylen and Woo (1982), applicable when snow melt and rainfall floods are fully identifiable, by their date of occurrence, in the available record of annual maximum flows. Snow melting floods lead to low flows and occur in spring-summer, and rainfall floods lead to higher flows and occur in winter from October to February. For the second approach, PDF mixtures have been proposed, for example: (1) the so-called Double Gumbel suggested by González-Villarreal (1970); (2) the Wakeby distribution proposed by Houghton (1978), which attempts to model the left and right portions separately; (3) the TCEV model developed by Rossi, Fiorentino, and Versace (1984), and (4) the mixed Gumbel function recently applied by Molina-Aguilar, Gutiérrez-López and Aparicio-Mijares

(2018). Rulfova, Buishand, Roth and Kysely (2016) proposed the TCGEV (two-component generalized extreme value) distribution of two components with generalized extreme values, such as a more flexible model for the analysis of maximum annual precipitation of six hours duration.

TCEV distribution is described in detail in this article, because it has not been applied systematically in Mexico; neither under precept, nor as an option of the probabilistic models suggested in records of floods with mixed populations.

Objective

In this paper, the theoretical origin of TCEV (two-component extreme value) distribution is detailed, with four fitting parameters, which leads to a probabilistic model appropriate to annual flood records that come from two physically different mechanisms, but that are mixed. The TCEV distribution fitting method through maximum likelihood is described in its two versions: (1) Successive Substitution Method and (2) Optimization Method. With both methods, six records of floods taken from the specialized literature are processed, which vary from 31 to 72 data and have between three and six outliers. Then, six return period predictions for 10, 25, 50, 100, 500 and 1 000 years are contrasted, each with six values coming from the fit of TCEV, Log-Pearson type III, General Extreme Values, Generalized Logistics and Wakeby distributions. Finally, the study conclusions are formulated.

Other Fitting Methods and Applications of TCEV Distribution

Francés (1998) expanded the maximum likelihood method for the joint use of systematic and historical information of floods. Beran, Hosking and Arnell (1986) put forward three equations to quantify the probability weighted moments (PWM) and indicate that such expressions must be resolved in an iterative manner, which makes them more complicated than the maximum likelihood method. The procedure of the PWM method, has been exposed by Singh (1998), as well as the one based on the maximum entropy principle, originally suggested by Fiorentino, Arora and Singh (1987).

Moreover, Fiorentino, Versace and Rossi (1985) were the first ones to state that TCEV distribution showed a good fit in maximum annual daily precipitation records in southern Italy, which were processed as series of partial duration or magnitudes greater than a threshold value. Later, Cannarozzo, D'Asaro, and Ferro (1995) applied the TCEV distribution regionally in Sicily, Italy, using rainfall and flood records. Ferro and Porto

(2006) use the TCEV model with the regional hierarchical approach in an FFA for Sicily, Italy. Such a regional approach was originally proposed by Fiorentino, Gabriele, Rossi and Versace (1987). Boni, Parodi and Rudari (2006), as well as Aronica and Candela (2007) use the TCEV distribution in the regional rainfall analyses they have described. Escalante-Sandoval and Reyes-Chávez (2004) perform a flood bivariate analysis using TCEV as marginal distributions.

Theory and data descriptions

Mixed Models PDF

Waylen and Woo (1982) call *R* the random variable representing the annual floods caused by precipitation, which are the largest ones, and *S* to those generated by the melting of snow, which are the basic or minor ones; X_i is a new random variable that is the maximum of both, *i* is the year and *n* the amplitude of the available record. As the hypothesis is accepted that *R* and *S* are independent, its PDF is equal to the product of both, and each one is estimated separately with the corresponding *n* data taken from the registry. So, the non-exceedance probability function of X_i is (Francés-García, 1995):

$$X_i = \max(R_i, S_i) \tag{1}$$

$$F_X(x) = P(X \leq x) = P(R \leq x, S \leq x) = P(R \leq x) \cdot P(S \leq x) = F_R(x) \cdot F_S(x) \tag{2}$$

In the simplest case, $F_R(x)$ and $F_S(x)$ are Gumbel PDFs with two fitting parameters, i.e., location and scale, each. The application of five fitting parameter Double Gumbel distribution has been solved in a practical way through not restricted optimization (Gómez, Aparicio, & Patiño, 2010); its PDF is (González-Villarreal, 1970):

$$F_X(x) = P(X \leq x) = F_R(x) \cdot [p + (1 - p) \cdot F_S(x)] \tag{3}$$

In the above equation, p is the probability of having ordinary events. In the Wakeby distribution (Houghton, 1978), the new random variable is defined by the addition, since the annual maximum flood is produced by the combination in time of the two generating mechanisms and then its PDF is explicitly expressed as:

$$x = -a \cdot [1 - F_X(x)]^b + c \cdot [1 - F_X(x)]^d + e \tag{4}$$

The estimation of its five fitting parameters has been solved by means of the L-moment method (Hosking & Wallis, 1997) and by restricted optimization (Campos-Aranda, 2001). Finally, the mixed Gumbel distribution has been fitted based on modern search algorithms (Molina-Aguilar *et al*., 2018), its PDF is similar to Equation (3), this is:

$$F_X(x) = P(X \leq x) = p \cdot F_R(x) + (1 - p) \cdot F_S(x) \quad (5)$$

TCEV Distribution Genesis

Poisson Compound Process

Rossi *et al.* (1984) established that a theoretical principle is necessary to support a certain PDF, so that it can be considered and accepted as a model of origin or source (*parent distribution*) of floods of certain geographical region. Due to the limitations of available records of floods, in addition to their statistical particularities, the PDF to be selected must have a probabilistic structure that simulates the real origin of the floods of the area under study.

In this regard, Rossi *et al.* (1984) state that the basic hypothesis of the extreme value theory is very restrictive, considering that annual floods come from the maximum value of a wide series of independent and identically distributed random variables (*iid*). A more flexible approach, which shows a greater similarity with the physical reality, considers that annual floods (X) are the maximum of a sequence of k random numbers with Poisson distribution, which are non-negative *iid* random variables Z_i, with i = 1, 2,…, k, which are also independent of k. The above is equivalent to modeling X as the maximum of a compound Poisson process, whose PDF is (Todorovic & Zelenhasic, 1970):

$$F_X(x) = P[X \leq x] = \exp\{-\lambda[1 - F_Z(x)\} \text{ for } x \geq 0 \tag{6}$$

where $\lambda = E[k]$ is the parameter of the Poisson process. If Z is adopted as an exponential random variable, we have that:

$$F_Z(z) = 1 - \exp(-x/\theta) \text{ for } z \geq 0 \tag{7}$$

being $\theta = E[Z]$. Substituting Equation (7) in Equation (6) it gives:

$$F_X(x) = \exp[-\lambda \cdot \exp(-x/\theta)] \text{ for } x \geq 0 \tag{8}$$

The above expression can be transformed into the Gumbel PDF or Type I of extreme values, making its location parameter ε equal to θ by the natural logarithm of λ, obtaining:

$$F_X(x) = \exp\left\{-\exp\left[-\frac{x-\varepsilon}{\theta}\right]\right\} \text{ for } -\infty < x < \infty \tag{9}$$

being:

$$\varepsilon = \theta \cdot \ln \lambda \tag{10}$$

And:

$$\lambda = \exp(\varepsilon/\theta) \tag{11}$$

Note that Equation (8) has a discontinuity in $x = 0$ which is exp(-λ), i.e., it exhibits a delta component, while Equation (9) extends to the negative values of x. The exp(-λ) value is close to zero, except perhaps in arid climates for the occurrence of null values (Rossi *et al*., 1984).

Mixed Model Development

A possible approach to take into account in FFAs, the presence of outliers and series of high asymmetry, is to accept that flood record comes from two different mechanisms of generation, one of them originates low magnitude but frequent floods and the other generates extraordinary sporadic floods. Then, assuming that there are two sequences of independent increments of variables *iid*, Z_{1i}, i = 1, 2,…, K_1 y Z_{2j}, j = 1, 2, . . . , K_2, each one defined by a compound Poisson process with parameters $\lambda_1 = E[K_1]$ and $\lambda_2 = E[K_2]$, respectively, with $\lambda_1 > \lambda_2$. The total number $K = K_1 + K_2$ of independent floods in a year will also be a compound Poisson process with parameter $\lambda = \lambda_1 + \lambda_2$, while the magnitude Z of the annual floods shall be defined by the mixture of PDF, that means (Rossi *et al*., 1984):

$$F_Z(z) = p \cdot F_{Z_1}(z) + (1-p) \cdot F_{Z_2}(z) \text{ for } z \geq 0 \qquad (12)$$

where, $p = \lambda_1/\lambda$ is the ratio of Z_1 in the mixture (basic or ordinary floods) and (1 - p) that of extraordinary floods with much greater variability. Substituting Equation (6) in Equation (12) it gives:

$$F_X(x) = \exp\{-\lambda_1[1 - F_{Z_1}(x)] - \lambda_2[1 - F_{Z_2}(x)]\} \text{ for } x \geq 0 \qquad (13)$$

Designating X_1 and X_2 to the annual maximum of Z_1 and Z_2, with:

$$F_X(x) = F_{X_1}(x) \cdot F_{X_2}(x) \qquad (14)$$

Finally, if Z_1 and Z_2 are exponential random variables, as in Equation (8), Equation (14) becomes the two-component extreme value (TCEV) distribution:

$$F_X(x) = \exp[-\lambda_1 \cdot \exp(-x/\theta_1) - \lambda_2 \cdot \exp(-x/\theta_2)] \text{ for } x \geq 0 \quad (15)$$

Its four fitting parameters characterize the basic floods and the extraordinary with the average number of independent floods per year ($\lambda_1 > 0$, $\lambda_2 \geq 0$) and the average amplitude of the maximum annual flow ($\theta_2 \geq \theta_1 > 0$). Fiorentino *et al*. (1985) emphasize that Equation (15) has a finite probability close to zero, whose value is exp(-λ) and that in the absence of the extraordinary value component, the TCEV distribution is reduced to the Gumbel distribution. The probability density function of X, according to the equation above is:

$$f_X(x) = F_X(x) \cdot \psi(x) \quad (16)$$

in which:

$$\psi(x) = (\lambda_1/\theta_1) \cdot \exp(-x/\theta_1) + (\lambda_2/\theta_2) \cdot \exp(-x/\theta_2) \quad (17)$$

Generalizing Equation (11) and applying it to Equation (15), it is obtained that TCEV model is equal to the product of two Gumbel distributions (Rossi *et al*., 1984; Metcalfe, 1997), this is:

$$F_X(x) = \exp\{-\exp[-(x-\varepsilon_1)/\theta_1]\} \cdot \exp\{-\exp[-(x-\varepsilon_2)/\theta_2]\} \qquad (18)$$

In the equation above, ε_1 and ε_2 are the location parameters and θ_1 and θ_2 those of the scale of each Gumbel distribution.

Probability of outliers

Beran *et al*. (1986) expose and analyze various complementary theoretical aspects relating to the TCEV distribution. They find that, like the Wakeby distribution, the TCEV model has a denser right tail than most of the PDFs used in hydrology. They also show that the TCEV distribution is able to show the phenomenon of separation, using extensive records from England.

Beran *et al*. (1986) suggest that the proportion of outliers in a record is of interest and can be used to validate the acceptance of the TCEV distribution. This ratio called q is a function of θ and λ, according to the following expressions:

$$\theta = \theta_2/\theta_1 \tag{19}$$

$$\lambda = \lambda_2/\lambda_1^{1/\theta} \tag{20}$$

$$q = -\frac{\lambda}{\theta}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!}\lambda^j \cdot \Gamma\left(\frac{j+1}{\theta}\right) \tag{21}$$

Beran *et al*. (1986) expose a graphic to estimate q whose lines vary from 0.01 to 0.90, with $\theta \cdot \ln \lambda$ in the x-axis and θ in the y-axis. They also state that in Equation (21) the series converges rapidly for $q < 0.90$ and that, when the TCEV distribution was used with 2 334 stations-year, from 57 records of England, one outlier was obtained for every 33 annual maximum data, i.e. q = 3.03 %; in contrast, Rossi *et al*. (1984) find in the Italian rivers one outlier for every 7 annual data, i.e., q = 14.3 %. The Stirling approximation (Davis, 1972) can be used to estimate the value of the Gamma function:

$$\Gamma(\omega) \cong e^{-\omega} \cdot \omega^{\omega-1/2} \cdot \sqrt{2\pi}\left(1 + \frac{1}{12\omega} + \frac{1}{288\omega^2} - \frac{139}{51840\omega^3} - \frac{571}{2488320\omega^4} + \cdots\right) \tag{22}$$

Fitting by Maximum Likelihood

1. Successive Substitution Method

Rossi *et al*. (1984) expose the natural logarithm of the maximum likelihood function and designate it by L_{fmv}, its expression is:

$$L_{fmv} = \sum_{i=1}^{n} \ln f_X(x_i) = \sum_{i=1}^{n} \ln F_X(x_i) + \sum_{i=1}^{n} \ln \psi(x_i) \tag{23}$$

Equalizing to zero the partial derivatives of L_{fmv} with respect to the four fitting parameters and performing several algebraic operations, Rossi *et al*. (1984) obtain the following four equations, which are solved by a successive substitution technique:

$$\lambda_j = \lambda_j \frac{\sum_{i=1}^{n} \frac{\exp\left(-x_i/\theta_j\right)}{\psi(x_i)}}{\theta_j \sum_{i=1}^{n} \exp\left(-x_i/\theta_j\right)} \text{ for } j = 1, 2 \tag{24}$$

$$\theta_j = \frac{\sum_{i=1}^{n} \frac{x_i \cdot \exp\left(-x_i/\theta_j\right)}{\psi(x_i)}}{\sum_{i=1}^{n} x_i \cdot \exp\left(-\frac{x_i}{\theta_j}\right) + \sum_{i=1}^{n} \frac{exp\left(-x_i/\theta_j\right)}{\psi(x_i)}} \text{ for } j = 1, 2 \tag{25}$$

In order to ensure rapid convergence, it is appropriate to start with values of λ and θ as close as possible. For the search of the initial values, flood record data are first drawn on the Gumbel-Powell paper (Chow, 1964), assigning them a graphical position or non-exceedance probability by means of the Weibull formula (Benson, 1962):

$$P(X < x) = \frac{m}{n+1} \tag{26}$$

where *m* is the order number of data when they have been sorted from lowest to highest, *n* is the number of observations, floods or data. The basic series is then identified and represented by a straight line or Gumbel distribution (Equation (9)), defining a point ($F1$, $X1$) in the beginning, and another one towards the end of designated data ($F2$, $X2$). Then, another straight line or Gumbel model is drawn, with a higher slope to represent extreme values or extraordinary floods, using ($F2$, $X2$) and defining an endpoint in ($F3$, $X3$), towards the last of the data. These three pairs of values are taken to the following formulas that come from Equation (9) (Campos-Aranda, 2002), to define the initial values λ_j and θ_j:

$$\theta_1 = \frac{(X2 - X1)}{\{-\ln[-\ln(F2)]\}} - \{-\ln[-\ln(F1)]\} \tag{27}$$

$$\varepsilon_1 = X1 - \theta_1\{-\ln[-\ln(F1)]\} \tag{28}$$

$$\lambda_1 = \exp(\varepsilon_1/\theta_1) \tag{29}$$

$$\theta_2 = \frac{(X3 - X2)}{\{-\ln[-\ln(F3)]\}} - \{-\ln[-\ln(F2)]\} \tag{30}$$

$$\varepsilon_2 = X2 - \theta_2\{-\ln[-\ln(F2)]\} \tag{31}$$

$$\lambda_2 = \exp(\varepsilon_2/\theta_2) \tag{32}$$

Based on the initial values, Equation (24) and Equation (25) are applied for the first time and the logarithmic function of maximum likelihood is evaluated (L_{fmv}), with Equation (23) and auxiliary numbers Equation (15), Equation (16) and Equation (17). The standard error of fit is also calculated with Equation (33).

Next, the new values of λ_j and θ_j become the initials and apply Equation (24) and Equation (25) again. If the value of L_{fmv} decreased, the process is repeated; if increased, the successive substitution is suspended. This process was programmed in *Basic* language, and such code was called TCEVMV.

At the end of the process, the optimal values of the fitting parameters are obtained and Equation (15) is applied repeatedly with them to obtain pairs of values of x and $F_X(x)$ to construct the TCEV model on Gumbel-Powell paper and obtain the desired predictions associated with return periods (*Tr*) or average recurrence intervals of 10, 25, 50, 100, 500 and 1 000 years. The *Tr* is the reciprocal of the exceedance probability and then the mentioned *Tr* correspond to non-exceedance probabilities [$F_X(x)$] of 0.90, 0.96, 0.98, 0.99, 0.998 y 0.999, respectively.

2021, Instituto Mexicano de Tecnología del Agua

2. Objective Function Maximization Method

Metcalfe (1997) proposed to maximize the logarithm of the maximum likelihood function (Equation (23)), for TCEV distribution fitting, and therefore, Campos-Aranda (2002) used the negative value of such expression (L_{fmv}) as *objective function* to minimize by means of a numerical algorithm of multiple non restricted variables (Rosenbrock, 1960), through the code available in Fortran (Kuester & Mize, 1973), which was translated into the Basic language for ease to handle data and printings (Campos-Aranda, 2003).

Numerical tests showed that the use of Equation (18) in Equation (23) is more convenient than the application of Equation (15), due to greater numerical stability and faster convergence towards the desired minimum. This numerical algorithm was programmed in *Basic* language, and such code was called TCEVROS.

For the start of the Rosenbrock algorithm, initial values of the variables to be optimized or fitting parameters of the TCEV distribution are defined, based on Equation (27), Equation (28), Equation (29), Equation (30) and Equation (31). At such initial values: ε_1, θ_1, ε_2 and θ_2 the first magnitude of the objective function is calculated (Equation (23), aided from Equation (17)) and the search for the minimum is started by means of the Rosenbrock algorithm.

At the end of the process, the stages and evaluations of the objective function (OF) are defined, and the optimal values of the fitting parameters are obtained and the Equation (18) is applied repeatedly with them in order to obtain pairs of values of x and $F_X(x)$, as well as predictions, as it was done in the successive substitution method.

Standard Error of Fit

For purposes of quantitative comparison of the fitting achieved with the TCEV distribution and the method of maximum likelihood, as opposed to other probabilistic models, it was proposed to estimate the *standard error of fit* (*SEF*), defined as (Kite, 1977):

$$SEF = \left[\frac{\sum_{i=1}^{n}(Qo_i - Qc_i)^2}{n - np}\right]^{1/2} \qquad (33)$$

being n total number of data, Qo_i the maximum annual flow observed measured from lowest to highest , and Qc_i the maximum flow calculated with the TCEV distribution (Equation (15) or Equation (18)), for the same non-exceedance probability assigned to the observed flow by the Weibull formula (Equation (26)). Finally, np is the number of fitting parameters, in this case four (λ_1 or ε_1, θ_1, λ_2 or ε_2, θ_2).

Since Equation (15), Equation (16), Equation (17) and Equation (18) have no inverse solution, the approach was to use the bisection method with error tolerance equal to 0.0001, between the calculated probability (Equation (26)) and that obtained with Equation (15) or Equation (18), using lower limit values of 0.001 of Qo_i and upper of 3 times Qo_i. The evaluation of the *SEF* was attempted at the end of the calculation of the initial values of the fitting parameters (Equation (27), Equation (28), Equation (29), Equation (30), Equation (32) and Equation (32)) and sometimes the method did not work; for example, when the TCEV distribution does not match the data. For such cases, the use of the *SEF* subroutine is suppressed. The other *SEF* calculation was performed upon completion of the numerical substitution method or Rosenbrock algorithm and generally concluded satisfactorily.

Flood Records to be processed

Six series of maximum annual flows (m^3/s) were processed, taken from the specialized literature, which are presented in progressive order of amplitude (*n*), in Table 1. Its origin is as follows: (1) taken from Haan, Barfield, and Hayes (1994) for the place called Beargrass Creek with 31 values; (2) the record of 37 values of Santa Cruz station has been exposed by Molina-Aguilar *et al.* (2018); (3) Francés-García (1995) presented the Turia river one at station E-25 with 41 data; (4) the record of 53 data at Huites station comes from Campos-Aranda (1999); (5) Gómez *et al*. (2010) showed the record of 58 data in La Cuña station; and

2021, Instituto Mexicano de Tecnología del Agua

finally, (6) Kite (1991) presented the data of St. Mary's River, with 72 data.

Table 1. Maximum annual flows (m^3/s) in the six processed records of hydrometric stations: Beargrass Creek, Santa Cruz, E-25 Turia River, Huites, La Cuña, and St. Mary's River.

No.	Record number								
	1	2	3	4		5		6	
1	51.3	2 142.0	139	2 085	1 530	784.0	595.2	565	824
2	22.4	1 023.4	90	2 531	8 000	736.8	110.2	294	292
3	23.8	837.6	63	14 376	5 496	510.0	523.9	303	345
4	49.6	1 161.2	2300	2580	3 385	461.0	1 636.3	569	442
5	25.4	1 062.0	26	1 499	1 374	411.0	1 168.0	232	360
6	60.0	784.2	260	1 165	1 245	326.0	295.0	405	371
7	34.5	1 086.3	117	1 127	2 299	349.8	212.8	228	544
8	36.5	487.8	76	3 215	1 345	130.4	367.4	232	552
9	21.7	677.0	514	10 000	11 350	690.0	144.6	394	651
10	44.5	807.0	84	3 229	2 509	266.0	78.4	238	190
11	35.1	553.0	90	677	2 006	199.0	261.9	524	202
12	30.0	1252.0	3700	1 266	1 180	690.0	196.3	368	405
13	42.2	369.5	88	1 025	-	340.6	46.8	464	583
14	25.0	293.0	155	955	-	249.6	313.8	411	725
15	37.4	1 157.2	199	4 780	-	350.0	319.6	368	232
16	93.4	762.2	60	696	-	317.0	621.1	487	974

17	68.0	1 074.0	58	593	-	732.6	824.5	394	456
18	27.6	1 280.0	79	3010	-	265.1	-	337	289
19	26.0	1 002.0	150	1 908	-	743.6	-	385	348
20	111.0	3 680.0	918	15 000	-	463.9	-	351	564
21	32.6	861.0	90	1 396	-	1474.9	-	518	479
22	24.7	888.8	133	1 620	-	323.0	-	365	303
23	20.2	1 166.4	25	2 702	-	160.4	-	515	603
24	41.1	950.0	150	1 319	-	763.8	-	280	514
25	20.0	7000.0	136	1 944	-	578.0	-	289	377
26	147.2	484.0	35	2 420	-	191.8	-	255	318
27	60.9	920.6	43	2 506	-	2440.0	-	334	342
28	33.1	812.0	34	1 534	-	238.4	-	456	593
29	58.9	3 332.4	40	1 508	-	622.1	-	479	378
30	35.4	898.0	238	1 558	-	1374.0	-	334	255
31	64.3	2790.0	37	2 200	-	439.7	-	394	292
32	-	620.0	49	2 225	-	280.2	-	348	-
33	-	1 495.0	32	7 960	-	267.2	-	428	-
34	-	836.0	42	4 001	-	287.3	-	337	-
35	-	940.0	34	1 067	-	280.7	-	311	-
36	-	3 080.0	117	3 233	-	156.5	-	453	-
37	-	1 550.0	64	1 119	-	455.5	-	328	-
38	-	-	48	6178	-	501.2	-	564	-
39	-	-	48	4443	-	385.0	-	527	-
40	-	-	42	1474	-	698.2	-	510	-

41	-	-	144	2508	-	184.7	-	371	-

For the three gauging stations in Mexico, their keys, locations and registration periods are placed immediately. (1) Santa Cruz: 10040, Rio San Lorenzo, of the Hydrological Region No. 10 (Sinaloa), 1944-1980; (2) Huites: 10037, Río Fuerte, of the Hydrological Region No. 10 (Sinaloa), 1941-1993 and (3) La Cuña: 12054, Río Verde, of the Hydrological Region No. 12-3 (Río Santiago), 1947-2004.

Record Homogeneity Verification

For the results of the FFA to be reliable, the data to be used must come from a *stationary random process*, which implies that it has not changed over time. Then, flood records must be composed of independent data, which are free of deterministic components, for such a record to be *homogeneous*.

To verify the aforementioned, seven statistical tests were applied, i.e., one general, the Von Neumann test, and six specific: two persistence tests (Anderson and Sneyers), two trend tests (Kendall and Spearman), one for the change in the mean (Cramer) and the last one to look for inconsistencies in dispersion (Bartlett). These tests are available at WMO (1971), and Machiwal and Jha (2008). All the tests mentioned were applied with a level of significance (α) of 5 % and six of them show that the selected records are homogeneous. Bartlett test detects loss of homogeneity for excess variability due to the presence of outliers.

In addition, the Wald-Wolfowitz test was applied, which is a non-parametric test that has been used by Bobée and Ashkar (1991), and by Rao and Hamed (2000) to test *independence* and *stationarity* in annual maximum flow records (x_i). Based on this test, the statistical quality of the records to be processed was also approved.

Results and analysis

Fittings According to Numerical Substitution Method

Table 2 and Table 3 show the results of the TCEV distribution fitting to the six processed records, based on Equation 24 and Equation (25). It is noted that the number of iterations varied from 1 to 15 and that only in the Turia river record 3 could the initial *SEF* not be assessed. The evaluation of the quality of the achieved fitting with this method will be estimated during the prediction contrast.

Table 2. Results of TCEV distribution fitting, according to numerical substitution method, in the three hydrometric stations indicated.

Station	Beargrass Creek	Santa Cruz	Turia river
Data in	Haan *et al.* (1994)	Molina-Aguilar *et al.*(2018)	Francés-García (1995)
Number of data	31	37	41
minimum and maximum:	20.0, 147.2	293.0, 7000.0	25.0, 3700.0
$F1$, $X1$	0.20, 24.0	0.022, 300.0	0.10, 35.0
$F2$, $X2$	0.83, 62.0	0.830, 1 600.0	0.90, 500.0
$F3$, $X3$	0.97, 148.0	0.974, 7 000.0	0.98, 1 750.0
$(\lambda_1)_{initial}$	6.282	7.662	2.904
$(\theta_1)_{initial}$	17.6	430.5	150.8
$(\lambda_2)_{initial}$	0.688	0.333	0.204
$(\theta_2)_{initial}$	47.5	2760.3	756.9
$OF(\ln L)_{initial}$	-139.490	-293.149	-278.073
$SEF_{initial}(m^3/s)$	11.6	375.3	-
Number of iterations:	4	1	15
$(\lambda_1)_{final}$	12.120	6.558	5.030
$(\theta_1)_{final}$	15.8	337.0	38.0

$(\lambda_2)_{final}$	0.784	0.238	0.138
$(\theta_2)_{final}$	46.4	2 180.2	1 401.5
$OF(\ln L)_{final}$	-137.440	-290.976	-243.446
$SEF_{final}(m^3/s)$	6.9	328.3	246.8

Table 3. Results of TCEV distribution fitting, according to numerical substitution method, in the three hydrometric stations indicated.

Station:	Huites	La Cuña	St. Mary's River
Data in:	Campos-Aranda (1999)	Gómez *et al.* (2010)	Kite (1991)
Number of data:	53	58	72
minimum and maximum:	593.0, 15 000.0	46.8, 2 440.0	190.0, 974.0
***F*1, *X*1**	0.10, 1 000.0	0.018, 50.0	0.05, 220.0
***F*2, *X*2**	0.80, 3 500.0	0.905, 860.0	0.95, 660.0
***F*3, *X*3**	0.98, 16 700.0	0.984, 2 475.0	0.99, 1 040.0
$(\lambda_1)_{initial}$	5.857	5.047	22.894
$(\theta_1)_{initial}$	1071.1	219.2	108.2
$(\lambda_2)_{initial}$	0.422	0.263	0.870
$(\theta_2)_{initial}$	5495.4	886.0	233.1
$OF(\ln L)_{initial}$	-473.167	-412.894	-452.755
$SEF_{initial}(m^3/s)$	790.1	116.4	33.9

Number of iterations:	5	5	1
$(\lambda_1)_{final}$	6.450	4.945	21.770
$(\theta_1)_{final}$	655.4	160.9	100.1
$(\lambda_2)_{final}$	0.358	0.253	0.706
$(\theta_2)_{final}$	5088.6	734.5	181.1
$OF(\ln L)_{final}$	-466.344	-408.736	-452.673
$SEF_{final}(m^3/s)$	478.7	83.1	30.3

In Figure 1, the Gumbel-Powell probability paper shows data recorded at the Huites station in five-value intervals from 1 to 40, and then one at a time. Dotted lines show the straight lines representing the two populations, and continuous line shows the TCEV distribution calculated with the successive substitution method. It is observed that the TCEV model follows approximately the basic floods up to number 42, but it deviates from the extraordinary ones from number 43 on, and coincides with the last one, which is data 53.

2021, Instituto Mexicano de Tecnología del Agua

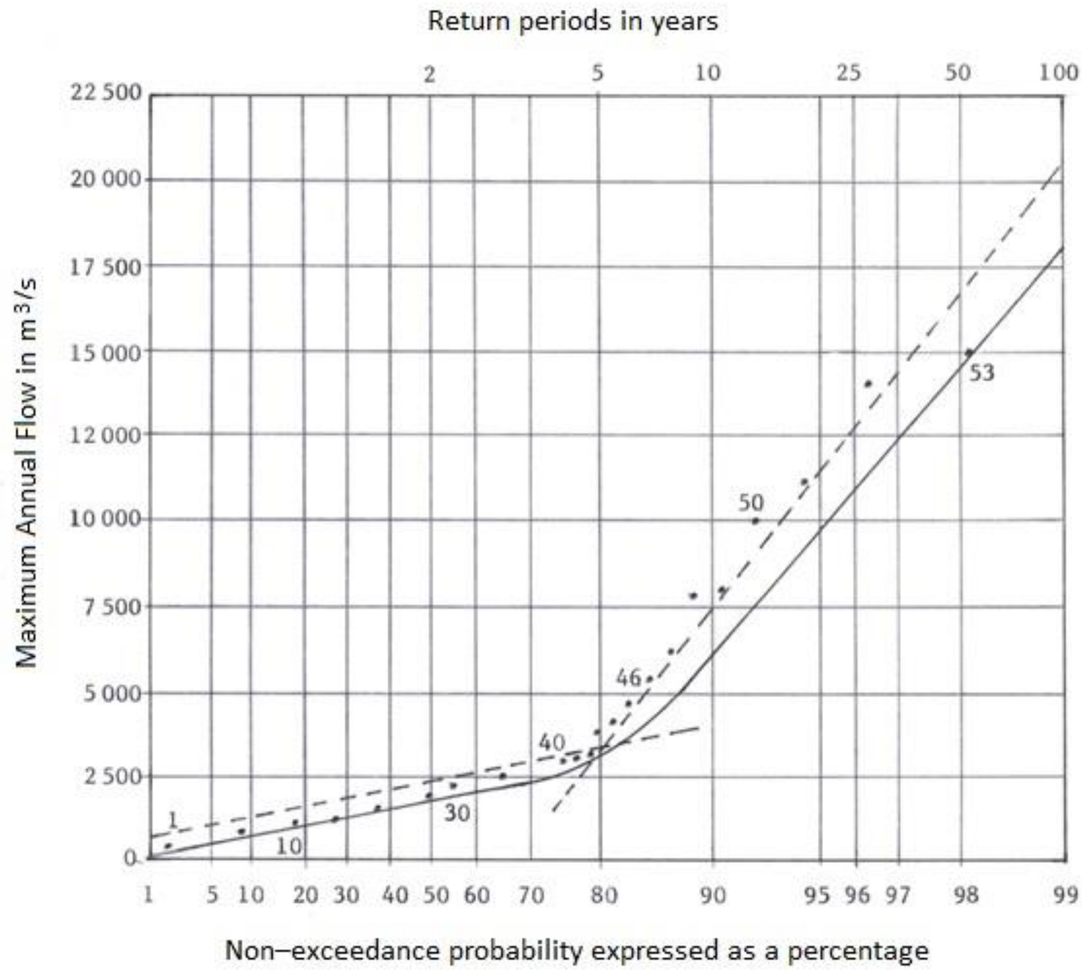

Figure 1. Fitting by the successive substitution method of the TCEV distribution to the Huites station floods, in the Gumbel-Powell probability paper.

Fittings According to the Optimization Method

In the records of Santa Cruz and La Cuña, the initial values defined by the points *F*1, *X*1 to *F*3, *X*3 did not allow the start of the Rosenbrock algorithm and were therefore slightly modified, as shown in Table 4 and Table 5, from the results of the TCEV function fitting to the six processed records. It is observed that sometimes this method improved fit (lower L_{fmv} and *SEF*), as in the first and last records processed, and on other occasions, it did not exceed the results of the numerical substitution method, as was the case with the records of the Turia and Huites rivers. Therefore, it is recommended that both maximum likelihood fitting methods be applied normally.

Table 4. Results of the fitting of TCEV distribution according to optimization method, in the three hydrometric stations indicated.

Station:	Beargrass Creek	Santa Cruz	Stat. 25 Turia River
***F*1, *X*1**	0.20, 24.0	0.10, 550.0	0.10, 35.0
***F*2, *X*2**	0.83, 62.0	0.78, 1 300.0	0.90, 500.0
***F*3, *X*3**	0.97, 148.0	0.95, 3830	0.98, 1750.0
$(\varepsilon_1)_{initial}$	32.387	830.945	160.738
$(\theta_1)_{initial}$	17.624	336.852	150.759
$(\varepsilon_2)_{initial}$	-17.785	-932.923	-1 203.201
$(\theta_2)_{initial}$	47.848	1 603.572	756.855
$OF(\ln L)_{initial}$	-139.490	-291.499	-278.073

***SEF*initial(m³/s)**	11.6	432.9	-
Number of stages:	6	8	17
Number of evaluations:	72	118	232
$(\varepsilon_1)_{final}$	27.918	729.148	60.849
$(\theta_1)_{final}$	8.133	277.394	36.710
$(\varepsilon_2)_{final}$	-19.230	-1274.057	-1 913.153
$(\theta_2)_{final}$	43.013	1 719.063	1 133.334
$OF(\ln L)_{final}$	-134.291	-289.773	-243.634
$SEF_{final}(m^3/s)$	5.9	397.9	267.0

Table 5. Results of the fitting of TCEV distribution according to optimization method, in the three hydrometric stations indicated.

Station:	Huites	La Cuña	St. Mary's River
***F*1, *X*1**	0.10, 1 000.0	0.05, 80.0	0.01, 185.0
***F*2, *X*2**	0.80, 3 500.0	0.88, 750.0	0.90, 580.0
***F*3, *X*3**	0.98, 16 700.0	0.97, 1 850.0	0.99, 1 010.0
$(\varepsilon_1)_{initial}$	1 893.361	313.058	344.690
$(\theta_1)_{initial}$	1 071.135	212.414	104.565
$(\varepsilon_2)_{initial}$	-4742.809	-827.543	168.192
$(\theta_2)_{initial}$	5495.425	766.904	182.996
$OF(\ln L)_{initial}$	-473.167	-411.328	-455.004
$SEF_{initial}$ (m³/s)	790.1	93.9	50.4

Number of stages:	12	6	8
Number of evaluations:	184	127	98
$(\boldsymbol{\varepsilon_1})_{\text{final}}$	1 445.455	280.490	315.582
$(\boldsymbol{\theta_1})_{\text{final}}$	639.558	162.207	96.423
$(\boldsymbol{\varepsilon_2})_{\text{final}}$	-6 417.104	-1 157.494	165.375
$(\boldsymbol{\theta_2})_{\text{final}}$	5 969.293	852.125	145.344
OF(ln L)$_{\text{final}}$	-466.337	-408.863	-451.528
$\boldsymbol{SEF_{\text{final}}(m^3/s)}$	540.6	54.4	22.5

In Figure 2, the Gumbel-Powell probability paper shows data recorded at La Cuña station in five-value intervals from 1 to 50 and then one at a time. Dotted lines show the straight lines representing the two populations and continuous line shows the TCEV distribution calculated with the maximization of the objective function method. It is observed that TCEV model represents in an excellent way all the floods.

2021, Instituto Mexicano de Tecnología del Agua

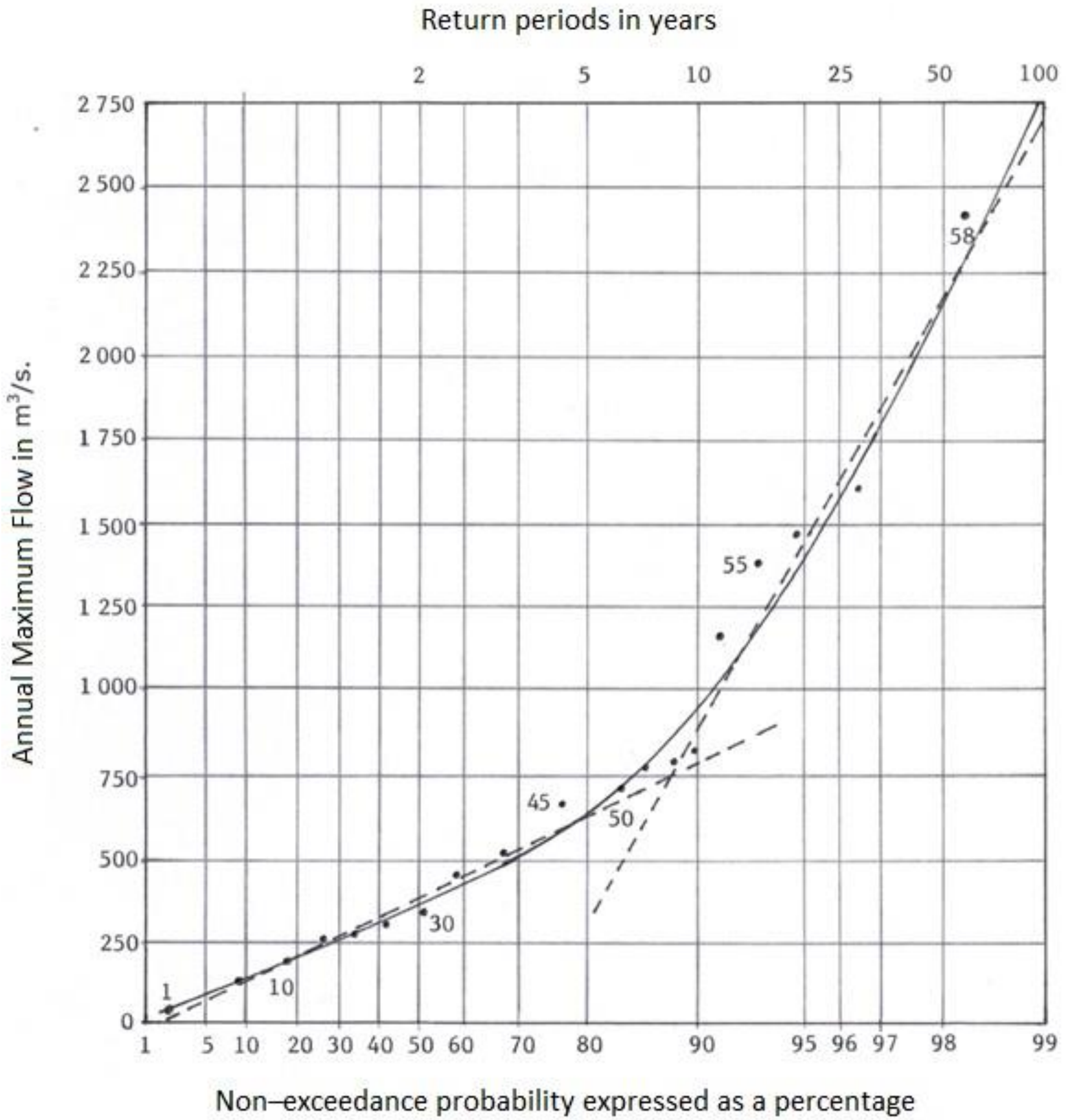

Figure 2. Fitting with the maximization of the objetive function method of the TCEV distribution to La Cuña station floods, in the Gumbel-Powell probability paper.

2021, Instituto Mexicano de Tecnología del Agua

Contrasts of TCEV Distribution Predictions

A first contrast was made between the predictions obtained with each of the two TCEV distribution fitting methods applied. It follows that both methods lead to similar predictions in all the records; the latter outstanding due to its similarity, and Huites station one due to its differences. This confirms the suggestion to fit the TCEV model with both exposed methods.

The second prediction contrast was carried out using the three distributions whose application has been suggested under precept. The Log-Pearson Type III (LP3) in USA; the Generalized Extreme Value (GEV) in England, from 1975 to 2000, and the Generalized Logistics (LOG) currently applied in England (Shaw, Beven, Chappell, & Lamb, 2011). In addition, Wakeby distribution (Houghton, 1978) was applied, which has shown great versatility and descriptive ability, being a mixed function.

LP3 distribution was fitted with the moment method, in the logarithmic domain (WRC, 1977) and in the real domain (Bobée, 1975), adopting the lower *SEF*; in contrast, GEV, LOG, and Wakeby functions were applied with the L-moment method (Stedinger, Vogel & Foufoula-Georgiou, 1993; Hosking & Wallis, 1997; Campos-Aranda, 2018). Table 6 shows the estimated predictions for each of the four cited probabilistic models.

2021, Instituto Mexicano de Tecnología del Agua

Table 6. Predictions of the indicated return periods obtained with five probability distributions, in the six processed records.

Station(*)	***SEF***	**Return periods, in years**					
PDF (**)	**(m^3/s)**	**10**	**25**	**50**	**100**	**500**	**1 000**
Beargrass Creek (*nvd* = 3; *q* = 9.7 %)							
TCEV (ns)	6.9	102	139	171	203	277	309
TCEV (no)	5.9	78	118	149	179	248	278
LP3 (dl)	3.4	77	111	143	184	320	404
GEV (mL)	7.7	76	107	138	177	313	398
LOG (mL)	4.9	74	105	136	178	334	440
WAK (mL)	7.2	79	110	138	170	266	319
Santa Cruz (*nvd* = 6; *q* = 16.2 %)							
TCEV (ns)	328.3	2 080	3 845	5 375	6 898	10 417	11 929
TCEV (no)	397.9	2 620	4 225	5 435	6 635	9 408	10 600
LP3 (mdl)	255.6	2 510	3 943	5 445	7 440	14 947	20 026
GEV (mL)	499.8	2 348	3 677	5 127	7 133	15 296	21 234
LOG (mL)	340.5	2 290	3 571	5 003	7 033	15 694	22 265
WAK (mL)	502.2	2 348	3 725	5 205	7 213	15 084	20 616
Stat. 25 Turia river (*nvd* = 4; *q* = 9.8 %)							
TCEV (ns)	246.8	381	1 708	2 693	3 671	5 932	6 905
TCEV (no)	267.0	638	1 712	2 510	3 300	5 129	5 915
LP3 (mdl)	207.0	408	1 093	2 323	4 975	30 201	66 688
GEV (mL)	441.0	453	958	1 666	2 883	10 223	17 612

2021, Instituto Mexicano de Tecnología del Agua

LOG (mL)	416.9	355	750	1 307	2 269	8 144	14 119
WAK (mL)	501.0	369	793	1 385	2 395	8 404	14 381
Huites (*nvd* = 6; *q* = 11.3 %)							
TCEV (ns)	478.7	6 240	11 050	14 630	18 200	26 387	29 916
TCEV (no)	540.6	7 025	12 680	16 870	21 045	30 675	34 814
LP3 (mdl)	949.9	6 290	10 492	15 061	21 303	45 851	63 053
GEV (mL)	1 007.1	5 948	9 614	13 600	19 091	41 299	57 363
LOG (mL)	984.2	5 786	9 322	13 263	18 830	42 460	60 302
WAK (mL)	893.1	6 303	10 070	13 834	18 626	35 412	46 091
La Cuña (*nvd* = 5; *q* = 8.6 %)							
TCEV (ns)	83.1	847	1 359	1 858	2 369	3 554	4 064
TCEV (no)	54.4	926	1 575	2 168	2 762	4 137	4 729
LP3 (mdr)	74.8	991	1 389	1 719	2 078	3 025	3 486
GEV (mL)	97.1	941	1 373	1 784	2 290	3 964	4 979
LOG (mL)	61.1	913	1 338	1 764	2 313	4 304	5 614
WAK (mL)	95.2	958	1 400	1 799	2 265	3 668	4 446
St. Mary's River (*nvd* = 4; *q* = 5.6 %)							
TCEV (ns)	30.3	568	680	769	862	1 099	1 211
TCEV (no)	22.3	595	707	795	888	1 122	1 233
LP3 (mdl)	16.8	600	712	797	883	1 093	1 188
GEV (mL)	22.1	603	713	796	881	1 079	1 167
LOG (mL)	15.5	589	706	807	919	1 240	1 410
WAK (mL)	23.3	605	715	794	868	1 025	1 087

**nvd* = number of outliers.

n = number of registration data.

q = (*nvd*/*n*)·100.

***Fitting method:*

ns = numerical substitution.

no = numerical optimization.

mdl = moments in the logarithmic domain.

mdr = moments in the real domain.

mL = L-moments.

In general terms, LP3, GEV, LOG and WAK distributions lead to higher predictions in all processed records, in return periods (*Tr*) of 500 and 1 000 years. Adopting the criterion of selection of the smallest *SEF*, in the first three records, function LP3 would be the option to take, but this leads to high predictions in the 500- and 1 000-year *Tr*. The above is remarkable in the Turia River and in Santa Cruz, and it could be acceptable in record 1. The opposite occurs in the Huites and La Cuña records, in which the best option is the TCEV distribution, in the first one, through the numerical substitution method and in the second one, with the optimization method (Figure 1 and Figure 2). Finally, in the St. Mary's River records, the distribution with the lowest *SEF* is Generalized Logistics, whose predictions are the largest, but of the same order of magnitude as those obtained with the other probabilistic models.

It should be noted that the predictions of the TCEV distribution in the Huites station (Figure 1), practically coincide with those obtained with the global approach by Campos-Aranda (1999).

Another appreciation detected in Table 5 is the following: when the percentage of outliers (*q*) is close to or exceeds 10 %, distributions that

2021, Instituto Mexicano de Tecnología del Agua

are applied under precept and Wakeby can lead to very high predictions in the *Tr* over 100 years, especially if such outliers are very large compared to the with the ordinary floods, case of the records of the river Turia, Huites and Santa Cruz. In such records, TCEV model is an option that must always be used, since its theoretical origin justifies its application in samples of floods that come from two different physical mechanisms of formation, but that are mixed and therefore show outliers.

Conclusions

First: TCEV distribution is a very important option for flood modeling, when the available record consists of events associated with two different hydrometeorological processes, that is, when there are flows that deviate considerably from the general trend of data, when drawing them on Gumbel-Powell paper.

Second: according to the results of Tables 2 to 5, the fitting of TCEV distribution using the maximum likelihood method, according to the successive substitution process and optimization method using the Rosenbrock algorithm, are simple procedures that generally converge and complement each other, as sometimes the first one provides better results (lowers L_{fmv} and *SEF*) and other times, the second one does.

Third: the contrast of predictions with six TCEV distribution return periods shown in Table 6 evidences that its two maximum likelihood fitting methods lead to similar predictions in all records. This confirms the suggestion to fit the TCEV model with both exposed processes.

Fourth: the predictions of the distributions that are applied under precept (LP3, GEV, and LOG) and of the probabilistic Wakeby model can serve to ratify or limit the estimates obtained with the TCEV function and help in the selection of design floods, taking into account flood genesis, *SEF*, and predictive capabilities of each distribution.

Acknowledgments

The comments and corrections suggested by the two anonymous referees A and B are appreciated, which allowed to improve the wording of the text, incorporate the processed data and correct descriptive omissions of the calculations made.

References

Aronica, G. T., & Candela, A. (2007). Derivation of flood frequency curves in poorly gauged Mediterranean catchments using a simple stochastic hydrological rainfall-model. *Journal of Hydrology*, 347, 132-142.

Benson, M. A. (1962). Plotting positions and economics of engineering planning. *Journal of Hydraulics Division, 88*(6), 57-71.

Beran, M., Hosking, J. R. M., & Arnell, N. (1986). Comment on "Two-component extreme value distribution for flood frequency analysis, by Rossi *et al*., (1984)". *Water Resources Research*, 22(2), 263-266.

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

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 raingauge time series. *Journal of Hydrology*, 327(3-4), 304-314.

Botto, A., Ganora, D., Laio, F., & Claps, P. (2014). Uncertainty compliant design flood estimation. *Water Resources Research*, 50(5), 4242-4253.

Campos-Aranda, D. F. (2018). Ajuste con momentos L de las distribuciones GVE, LOG y PAG no estacionarias en su parámetro de ubicación, aplicado a datos hidrológicos extremos. *Agrociencia*, 52(2), 169-189.

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

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. *Avances en Hidráulica*, 9, 527-532.

Campos-Aranda, D. F. (2001). Contraste de dos procedimientos de ajuste de la distribución Wakeby en modelación probabilística de crecientes. *Agrociencia*, 35(4), 429-439.

Campos-Aranda, D. F. (1999). Hacia el enfoque global en el análisis de frecuencia de crecientes. *Ingeniería Hidráulica en México*, 14(1), 23-42.

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.

Chow, V. T. (1964). Statistical and probability analysis of hydrologic data. Section 8-I: Frequency Analysis. In: Chow, V. T. (ed.). *Handbook of applied hydrology* (pp. 8.1-8.42). New York, USA: McGraw-Hill Book Co.

Davis, P. J. (1972). Chapter 6. Gamma Function and related functions. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (pp. 253-296). New Work, USA: Dover Publications.

Escalante-Sandoval, C. A., & Reyes-Chávez, L. (2004). *Análisis bivariado de gastos máximos anuales con distribuciones marginales TCEV*. En: XVIII Congreso Nacional de Hidráulica. San Luis Potosí, S.L.P., México. *Avances en Hidráulica*, 11, 523-529.

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

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.

Fiorentino, M., Arora, K., & Singh, V. P. (1987). The two-component extreme value distribution for flood frequency analysis: Derivation of a new estimation method. *Stochastic Hydrology and Hydraulics*, 1(3), 199-208.

Fiorentino, M., Gabriele, S., Rossi, F., & Versace, P. (1987). Hierarchical approach for regional flood frequency analysis. In: Singh, V. P. (ed.). *Regional flood frequency analysis* (pp. 35-49). Dordrecht, The Netherlands: Reidel Publishing Company.

Francés-García, F. (1995). *Utilización de la información histórica en el análisis regional de las avenidas* (nomografía No. 27). Barcelona, España: Centro Internacional de Métodos Numéricos en Ingeniería.

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.

Gómez, J. F., Aparicio, M., & Patiño, C. (2010). *Manual de análisis de frecuencias en hidrología*. Jiutepec, México: Instituto Mexicano de Tecnología del Agua.

González-Villarreal, F. J. (1970). *Contribución al análisis de frecuencias de valores extremos de los gastos máximos en un río* (No. 277)., México, DF, México: Instituto de Ingeniería de la Universidad Nacional Autónoma de México.

Haan, C. T., Barfield, B. J., & Hayes, J. C. (1994). Chapter 2. Hydrologic frequency analysis. *Design hydrology and sedimentology for small catchments*. (pp. 5-36). San Diego, USA: Academic Press.

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

Houghton, J. C. (1978). Birth of a parent: The Wakeby distribution for modeling flood flows. *Water Resources Research*, 14(6), 1105-1109.

Katz, R. W. (2013). Chapter 2. Statistical Methods for Nonstationary Extremes. In: Aghakouchak, A., Easterling, D., Hsu, K., Schubert, S., & Sorooshian, S. (eds.). *Extremes in a changing climate* (pp. 15-37). Dordrecht, The Netherlands: Springer.

Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P., & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 329(3-4), 534-552.

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.

Kite, G. W. (1991). Chapter 4. Frequency analysis. In: *Hydrologic Applications: Computer programs for water resources engineering*. (pp. 72-98). Littleton, USA: Water Resources Publications.

Kuester, J. L., & Mize, J. H. (1973). Chapter 9. Multivariable unconstrained methods. In: *Optimization techniques with Fortran* (pp. 297-365). New York, USA: McGraw-Hill Book Co.

López-de-la-Cruz, J., & Francés, F. (2014). La variabilidad climática de baja frecuencia en la modelación no estacionaria de los regímenes de las crecidas en las regiones hidrológicas Sinaloa y Presidio-San Pedro. *Tecnología y ciencias del agua*, 5(4), 79-101.

Machiwal, D., & Jha, M. K. (2008). Comparative evaluation of statistical tests for time series analysis: Applications to hydrological time series. *Hydrological Sciences Journal*, 53(2), 353-366.

Matalas, N. C., Slack, J. R., & Wallis, J. R. (1975). Regional skew in search of a parent. *Water Resources Research*, 11(6), 815-826.

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 4. Extreme value and related distributions. In: *Statistics in Civil Engineering* (pp. 81-115). London, UK: Arnold Publishers.

Molina-Aguilar, J. P., Gutiérrez-López, M. A., & Aparicio-Mijares, F. J. (2018). Búsqueda armónica para optimizar la función Gumbel mixta univariada. *Tecnología y ciencias del agua*, 9(5), 280-322.

Potter, W. D. (1958). Upper and lower frequency curves for peak rates of runoff. *EOS. Transactions of AGU*, 39(1), 100-105.

Prosdocimi, I., Kjeldsen, T. R., & Svensson, C. (2014). Non-stationarity in annual and seasonal series of peak flow and precipitation in the UK. *Natural Hazards and Earth System Sciences*, 14(5), 1125-1144.

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.

Rosenbrock, H. H. (1960). An automatic method of finding the greatest or least value of a function. *Computer Journal*, 3(3), 175-184.

Rossi, F., Fiorentino, M., & Versace, P. (1984). Two-component extreme value distribution for flood frequency analysis. *Water Resources Research*, 20(7), 847-856.

Rulfova, Z., Buishand, A., Roht, M., & Kysely, J. (2016). A two-component generalized extreme value distribution for precipitation frequency analysis. *Journal of Hydrology*, 534, 659-668.

Shaw, E. M., Beven, K. J., Chappell, N. A., & Lamb, R. (2011). Chapter 13: Estimating floods and low flows in the UK. In: *Hydrology in Practice* (4^{th} ed.) (pp. 322-350). London, UK: Spon Press.

Singh, V. P. (1998). Chapter 22. Two-component extreme value distribution. In: *Entropy-based parameter estimation in hydrology* (Vol. 10). (pp. 347-362). Dordrecht, The Netherlands: Kluwer Academic Publishers, Water Science and Technology Library.

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Chapter 18: Frequency Analysis of Extreme Events. In: Maidment, D. R. (ed.). *Handbook of Hydrology* (pp. 18.1-18.66). New York, USA: McGraw-Hill, Inc.

Todorovic, P., & Zelenhasic, E. (1970). A stochastic model for flood analysis. *Water Resources Research*, 6(6), 1641-1648.

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

WMO, World Meteorological Organization. (1971). Annexed III. In: *Climatic Change* (pp. 58-71) (Technical Note No. 79). Geneva, Switzerland: Secretariat of the World Meteorological Organization.

Waylen, P., & Woo, M. K. (1982). Prediction of annual floods generated by mixed processes. *Water Resources Research*, 18(4), 1283-1286.