DOI: 10.24850/j-tyca-2025-03-07

Articles

A simple operational process for the flood frequency analysis with partial duration series

Proceso operativo simple para el análisis de frecuencias de crecientes con series de duración parcial

Daniel Francisco Campos-Aranda[1], ORCID: https://orcid.org/0000-0002-9876-3967

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

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

Abstract

Through frequency analysis, the *design floods* (DF) are estimated, maximum river flows corresponding to low probabilities of being exceeded. With the DF, dikes and protection walls, bridges and urban drainage are hydrologically dimensioned. Frequency analysis generally processes the observed annual maximum flows or Annual Series of Maximums (SAM), but all flows that exceed a *threshold value* (*vu*) or Partial Duration Series (SDP) can also be processed, complying with the condition of being independent. The essential disadvantage of the SAM lies in considering only the maximum annual event, ignoring secondary values, which may exceed the maximums of other years. As the SDP is made up of a greater number of events, its *predictions* or DFs are more reliable. In this study, the SDP are integrated from the record of maximum monthly flows, adopting as the minimum *vu* x_0, the lowest maximum annual flow, and as the maximum *vu* the one that defines as many excess flows as years of the SAM record. The *adopted vu* accepts, graphical and numerically, the Poisson-Pareto distribution and leads to the lowest standard error of fit for the analyzed SDP data. The suggested operational process is applied in five hydrometric stations in two hydrological regions of Mexico and concludes with the contrast of SAM and SDP predictions. Finally, conclusions are formulated, which highlight the simplicity of the process and the accuracy of its predictions.

Keywords: Annual series of maximums, partial duration series, threshold values, Poisson-Pareto distribution, graph of average residual exceedances, dispersion index, predictions, relative error.

Resumen

A través del análisis de frecuencias se estiman las *crecientes de diseño* (CD), gastos máximos del río correspondientes a bajas probabilidades de ser excedidos. Con las CD se dimensionan hidrológicamente los diques y muros de protección, los puentes y los drenajes urbanos. El análisis de frecuencias procesa, por lo general, los gastos máximos anuales observados o serie anual de máximos (SAM), pero también se pueden procesar todos los gastos que exceden un *valor umbral* (*vu*) o serie de duración parcial (SDP), cumpliendo con la condición de ser independientes. La desventaja esencial de la SAM radica en considerar solo el evento máximo anual, ignorando sus valores secundarios, los cuales pueden exceder a los máximos de otros años. Como la SDP está formada por un mayor número de eventos, sus *predicciones* o CD resultan más confiables. En este estudio, las SDP se integran a partir del registro de gastos máximos mensuales, adoptando como *vu* mínimo x_0, el menor gasto máximo anual, y como *vu* máximo el que define tantos gastos excedentes como años del registro de la SAM. El *vu adoptado* acepta, gráfica y numéricamente, la distribución Poisson-Pareto y conduce al menor error estándar de ajuste para los datos de la SDP analizada. El proceso operativo sugerido se aplica en cinco estaciones hidrométricas de dos regiones hidrológicas de México y concluye con el contraste de predicciones de la SAM y de la SDP. Por último, se formulan las conclusiones, las cuales destacan la sencillez del proceso y la exactitud de sus predicciones.

Palabras clave: serie anual de máximos, serie de duración parcial, valores umbral, distribución Poisson-Pareto, gráfico de excedencias residuales medias, índice de dispersión, predicciones, error relativo.

Received: 18/10/2023

Accepted: 25/03/2024

Publicado *online*: 02/05/2024

Introduction

Generalities

A common task or challenge of the hydrologist is to estimate the risk or probability of *floods* or maximum flows occurrence in a basin, or a specific location of a river. Such floods cause inundations, causing considerable economic and environmental damage; in addition, they can damage hydraulic works, such as reservoirs, bridges, dikes and retaining walls, as well as urban drainage. Preventing and mitigating inundations damage also requires significant investments (Stedinger, 2017).

Therefore, the estimation of the *Design Floods* (DF), which are used to dimension the hydraulic works and provide hydrological security, is an activity that must be carried out with the greatest possible accuracy. DFs are maximum flows associated with low probabilities of being exceeded, their estimation is carried out through the *frequency analysis*, a statistical technique that processes the available record of maximum flows, through a probability distribution function (PDF), to obtain inferences or *predictions* that are feasible to occur in the future (Bobée & Ashkar, 1991; Lang, Ouarda, & Bobée, 1999; Singh & Zhang, 2017).

For *Floods Frequency Analysis* (FFA) to lead to reliable results, the following four conditions must be met: (1) the record of maximum available flows must have been originated by a stationary random process; (2) the PDF used must be suitable for the processed record, therefore several that are considered appropriate are tested; (3) the fit of each PDF to the available data must be carried out with efficient methods, and (4) the selection of the results or predictions must be exhaustive and objective (Meylan, Favre, & Musy, 2012).

In the FFA and other extreme hydrological data, two types of records can be used: (I) maximum annual flows or annual series of maximums (SAM, by its acronym in Spanish) and (II) values greater than a *threshold value* (x_0) or partial duration series (SDP, by its acronym in Spanish), universally known as POT, peaks-over-threshold analysis. The basic objection to the SAM is that it considers only the annual maximum event, ignoring secondary values, which may exceed the maximums of other years. The latter is valid for the sequence of dry years present in the record, whose small values can influence the estimation of the DFs. The above is exacerbated in arid and semi-arid regions (Stedinger, Vogel, & Foufoula-Georgiou, 1993; Madsen, Rasmussen, & Rosbjerg, 1997; Lang *et al*., 1999; Bhunya, Mishra, Ojha, & Berndtsson, 2008; Bhunya, Singh, Berndtsson, & Panda, 2012; Bezak, Brilly, & Sraj, 2014).

The SDP dodges the disadvantages of the SAM, by using a *threshold level* and processing all flows above that basic level (x_0). However, its use has not become widespread, due to the difficulty in integrating it, since its values must meet the condition of independence and then, a criterion for the selection of events must be adopted (Madsen *et al*., 1997; Lang *et al*., 1999). Furthermore, the SDP requires two PDFs, one for counting exceedances per year and another for modeling their magnitudes (Bhunya *et al*., 2008). As the average and variance of exceedances per year changes with the record and with x_0, different PDFs must be applied to model their occurrence (Önöz & Bayazit, 2001; Bezak *et al*., 2014; Pan, Rahman, Haddad, & Ouarda, 2022).

According to Stedinger (2017), if a peak flow is properly defined, the SDP will be composed of independent events with equal distribution, meeting the basic requirement (*iid*) of frequency analysis. In general, the SDP is more convenient than the SAM, in short records of less than 14 years and in records from arid zones (Bhunya *et al*., 2008; Bezak *et al*., 2014).

Objectives

Campos-Aranda (2000) presented the basic theory of the statistical analysis of the SDP, now he formulates a simple procedure for integrating such series, finding by trial and error the *value of the threshold* (*vu*) that achieves the best fit; he verifies, graphically and numerically, the acceptance of the Poisson-Pareto probabilistic model and contrasts its predictions with those adopted when previously processing the SAM.

The *objectives* of this study were the following 4: (1) present a summary of theoretical concepts related to the integration and probabilistic processing of SDPs; (2) process five hydrometric records with SDP, integrated from the available records of monthly maximum flows, showing nine years per page in the BANDAS system (IMTA, 2003); (3) verification of the Poisson-Pareto distribution, using the graph of average residual exceedances and the dispersion index test and (4) contrast the fit errors and the predictions obtained with the processed samples of the SAM and the adopted SDP.

Although in a section prior to the presentation of the five SDP records that will be processed, the six stages that make up the suggested operating procedure are described in detail, Figure 1 shows a descriptive scheme of such sequence, to establish a prior conception and basic of its approach.

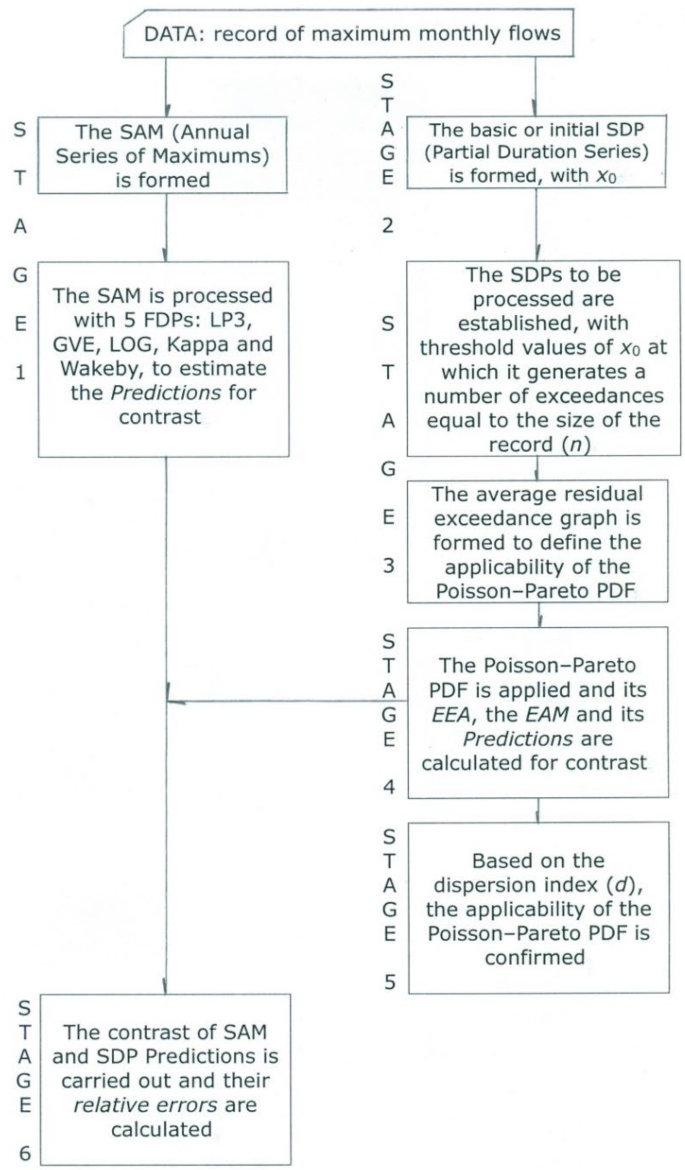

Figure 1. Basic descriptive diagram of the stages of the suggested operating procedure.

Tecnología y Ciencias del Agua
OPEN ACCESS
Check for updates

Theoretical concepts related to SDP

Independence criteria

When observed hydrographs are analyzed and two maximum flows occur (Q_1, Q_2) that define two peak hydrographs, their independence is established by the minimum period (θ) in days between them, as well as by the decrease that the flow reaches (x_{min}) between they. Bezak *et al*. (2014) present and apply the criterion followed in the USA, which is the following:

$$\theta > 5 \text{ days} + \ln(A) \quad (1)$$

$$x_{\min} < \left(\frac{3}{4}\right) \min(Q_1, Q_2) \quad (2)$$

In Equation (1), *A* is the basin area in square miles. The second peak flow in magnitude must be eliminated if equations (1) and (2) are not met. Lang *et al*. (1999) and Pan *et al*. (2022) present and analyze other similar criteria.

Threshold value selection (x_0)

The other difficulty associated with the integration of the SDP, is the ideal estimation of x_0 or minimum or base *threshold value*. According to Lang *et al*. (1999) there are two different approaches for the selection of the base value x_0: the first one is based on a physical criterion and therefore consists of identifying the maximum flow or flood that overflows the channel towards the flood plains. The second approach is based on mathematical and/or statistical considerations, which seek to select exceedances that are independent, and their occurrence can be described by a Poisson process.

For practical purposes, magnitudes that lead to a certain number of exceedances per year (λ) with a minimum of one and a maximum of five have been suggested; passing through the value of 1.63 to achieve a lower sample variance compared to that obtained with the SAM (Pan *et al*., 2022). Madsen *et al*. (1997) recommend estimating the value of x_0 with the expression:

$$x_0 = \mu_x + FC \cdot \sigma_x \quad (3)$$

Where:

μ_x = mean of the basic series of daily flows

σ_x = standard deviation of the basic series of daily flows

$FC = 3$ = frequency factor cited by Lang *et al*. (1999) and adopted by Bezak *et al*. (2014).

Madsen, Rosbjerg and Harremoës (1993) apply Equation (3) in SDP of extreme rainfall.

On the other hand, Bezak *et al*. (2014) indicate that Langbein (1949) suggested that the *minimum threshold value* (x_0) be equal to the lowest value of the SAM. Fischer and Schumann (2016) adopt the previous criterion and work with maximum monthly flow records, for which they accept that they may have a certain weak degree of dependence, which they ignore due to the gain in accuracy due to the increase in events or exceedances in the SDP formed.

In the author's opinion, this selection as the *minimum* threshold value (x_0) in an SDP leads to the acceptance of all values higher than the annual minimum, so that they integrate the exceedances and therefore, implies a relationship or correspondence between the SAM and the SDP. On the other hand, the *threshold value vu* that will be *adopted* is related to the best fit achieved and the verification of applicability of the Poisson-Pareto probabilistic model.

Relationship between return periods T_a and T_p

For an SDP with λ average exceedances per year and PDF of events exceeding x_0 equal to G(x), the independent events that do not exceed x and therefore occur in the interval from x_0 to x, will be defined by the following expression (Stedinger *et al*., 1993; Stedinger, 2017):

$$\lambda^* = \lambda[1 - G(x)] \quad (4)$$

If $F_a(x)$ is the PDF of the SAM, in other words, the probability that an annual maximum value does not exceed x. Then, for independent events, the probability of not having exceedances of x in a period of one year is given by the Poisson distribution (Stedinger *et al*., 1993; Stedinger, 2017), with the following expression:

$$F_a(x) = \exp(-\lambda^*) = \exp(-\lambda[1 - G(x)] = \exp(-1/T_p) \quad (5)$$

From the previous expression, the equation that relates the return period (T_p) of the SDP to the corresponding one (T_a) of the SAM is obtained, which is (Stedinger, 2017):

$$1/T_a = 1 - \exp(-1/T_p) \quad (6)$$

For infrequent return periods, for example, greater than 10 years, with the previous expression we obtain that $T_a \cong T_p + 0.50$, thereby indicating that the return periods of the SAM and the SDP are basically the same.

PDF of the number of exceedances per year (*λ*)

The Poisson probability distribution is commonly applied to estimate the probability of the number of exceedances (λ) per year greater than x_0. An important property of this distribution is that it also defines a Poisson process for any other upper threshold (Metcalfe, 1997; Önöz & Bayazit, 2001). The Poisson distribution is applicable when the mean value of the exceedances per year (E) is equal to its variance (V).

The binomial random variable is a count of the number of successes in a certain number of trials, while the negative binomial distribution variable counts the failures. In both distributions the number of successes is predetermined, and the number of trials is random (Bhunya *et al*., 2012). In the Binomial distribution the mean value is greater than the variance and in the Negative Binomial the variance is greater than its mean (Önöz & Bayazit, 2001; Bezak *et al*., 2014; Pan *et al*., 2022).

Test for selecting the distribution of *λ*

Önöz and Bayazit (2001) present the test developed by Cunnane (1979) to confirm the selection between the Poisson, Binomial and Negative Binomial distributions, since its initial adoption depends on the value of the average number of exceedances per year (*E*) and its variance (*V*). As already indicated, when $E \cong V$ the Poisson distribution is adopted, with $E > V$ the Binomial and with $E < V$ the Negative Binomial. The test statistic is the *dispersion index* (*d*) expressed as:

$$d = \sum_{i=1}^{n} \frac{(\lambda_i - \bar{\lambda})^2}{\bar{\lambda}} = \frac{(n-1)V}{E} \quad (7)$$

For a significance level of 5 %, the Poisson distribution is accepted if *d* is between $\lambda^2_{0.025}$ and $\lambda^2_{0.975}$, defined in Table 1, for *n* - 1 degrees of freedom (*d.l.*), being *n* the number of years from the record. If *d* is less than $\lambda^2_{0.025}$, the Binomial distribution is adopted and if *d* is greater than $\lambda^2_{0.975}$, the negative Binomial distribution is accepted (Önöz & Bayazit, 2001).

Table 1. Values of the Chi-square distribution for a 5 % significance level (Ostle & Mensing, 1975).

g.l.	$\chi^2_{0.025}$	$\chi^2_{0.975}$	*g.l.*	$\chi^2_{0.025}$	$\chi^2_{0.975}$	*g.l.*	$\chi^2_{0.025}$	$\chi^2_{0.975}$
30	16.8	47.0	54	35.6	76.2	78	55.5	104.3
32	18.3	49.5	56	37.2	78.6	80	57.2	106.6
34	19.8	52.0	58	38.8	80.9	82	58.8	108.9
36	21.3	54.4	60	40.5	83.3	84	60.5	111.2
38	22.9	56.9	62	42.1	85.7	86	62.2	113.5
40	24.4	59.3	64	43.8	88.0	88	63.8	115.8
42	26.0	61.8	66	45.4	90.3	90	65.6	118.1
44	27.6	64.2	68	47.1	92.7	92	67.4	120.4
46	29.2	66.6	70	48.8	95.0	94	69.1	122.7
48	30.8	69.0	72	50.4	97.4	96	70.8	125.0
50	32.4	71.4	74	52.1	99.7	98	72.5	127.3
52	34.0	73.8	76	53.8	102.0	100	74.2	129.6

Graph of average residual exceedances

Coles (2001) exposes two procedures to select the threshold value greater than x_0. The first one is based on the *average residual exceedances* graph, where the abscissa and ordinate are defined by the following expression:

$$\left(vu, \frac{1}{n_{vu}} \sum_{i=1}^{n_{vu}} (x_i - vu)\right) \tag{8}$$

Where *vu* is a threshold value greater than x_0 and x_i are the maximum flows that exceed *vu*, whose number is n_{vu}.

Therefore, when above an initial threshold value x_0, which accepts the Generalized Pareto PDF (PAG, by its acronym in Spanish), the graph of average residual exceedances must be approximately linear, for the aforementioned distribution to remain valid. The other procedure by Coles (2001) is based on the similarity that the location and scale parameters of the PAG distribution must have.

Magnitude of exceedances distribution

The exponential distribution was used by Shane and Lynn (1964) to model the magnitude of exceedances of x_0, in the pioneering work of probabilistic analyzes of SDPs. Subsequently, the Generalized Pareto distribution has been used to describe the PDF of $G(x)$, for the magnitude of events greater than the threshold x_0, this is (Metcalfe, 1997; Stedinger *et al*., 1993; Singh & Zhang, 2017):

$$G(x) = F = 1 - \left[1 - k\left(\frac{x - x_0}{\alpha}\right)\right]^{\frac{1}{k}} \text{ for } k \neq 0 \quad (9)$$

Where:

x_0 = location

α = scale

For positive values of the shape parameter k this PDF has an upper boundary at $x_{max} = x_0 + \alpha/k$, for $k < 0$ it has no boundary, and its right tail is thicker or denser. When $k = 0$, the exponential PDF of two parameters is obtained.

Poisson-Pareto distribution

Accepting that the generalized Pareto PDF describes $G(x)$, then Equation (9) is substituted in Equation (4) and a GVE distribution can be obtained, for values of the SAM greater than x_0 (Stedinger *et al.*, 1993; Campos-Aranda, 2000; Meylan *et al.*, 2012), whose equations are:

$$F_a(x) = \exp\left[-\left(1 - k\frac{x-u^*}{\alpha^*}\right)^{1/k}\right] \text{ for } k \neq 0 \quad (10)$$

Being:

$$k = \frac{\mu - x_0}{\lambda_2} - 2 \quad (11)$$

$$\alpha = (\mu - x_0)(1 + k) \quad (12)$$

$$u^* = x_0 + \frac{\alpha(1-\lambda^{-k})}{k} \quad (13)$$

$$\alpha^* = \alpha \cdot \lambda^{-k} \quad (14)$$

In Equation (9), $\mu = \beta_0$ is the average of the exceedances or values greater than x_0 and λ_2 is the second-order moment L, a function of the heavy probability moments β_1 and β_0 with the following expressions (Hosking & Wallis, 1997):

$$\lambda_2 = 2 \cdot \beta_1 - \beta_0 \quad (15)$$

$$\beta_r = \frac{1}{n} \sum_{j=r+1}^{n} \frac{(j-1)(j-2)\cdots(j-r)}{(n-1)(n-2)\cdots(n-r)} x_{j:n} \quad \text{for } r = 0 \text{ y } 1 \quad (16)$$

A sample of n size is considered, ordered in a progressive form, that is: $x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}$. u^* and α^* are the corrected location and scale parameters of the Generalized Pareto PDF, since they consider the Poisson process, through the average number of exceedances per year (λ).

Finally, the inverse solution of Equation (9) leads to the quantile function, which allows estimating the *predictions* associated with a certain probability of non-exceedance (F), obtained with the SDP, since they are a function of u^* and α^*, such an expression is:

$$x_{Tr} = u^* + \left(\frac{\alpha^*}{k}\right)\left[1 - (1 - F)^k\right] \quad (17)$$

Theoretical concepts related to SAM

Probability distributions used

For this contrast, between the quality of the fit and the predictions obtained with the SAM and SDP, the SAM records were processed based on the three PDFs that have been suggested as a standard or reference: the Log-Pearson type III (LP3), the General Extreme Values (GVE, by its acronym in Spanish) and Generalized Logistics (LOG, by its acronym in Spanish). In addition, two highly versatile FDPs were applied, the Kappa and Wakeby distributions.

The LP3 distribution was applied with the method of moments in the logarithmic (WRC, 1977) and real domains (Bobée, 1975), adopting the one with the best fit, and the rest of the distributions were applied with the method of L moments (Hosking & Wallis, 1997).

Standard error of fit

In the mid-1970s (Kite, 1977), the *standard error of fit* (*EEA*, by its acronym in Spanish) was established as a quantitative statistical indicator that estimates the quality of the fit and that also allows objective comparison between the various PDFs that are tested in a sample, since it has the data units. At present, it is the most common indicator for contrasting PDFs with real data (Chai & Draxler, 2014).

It has been applied in Mexico using the Weibull empirical formula. It is currently recommended to apply it using the Cunnane formula

(Equation (18)), which leads to approximately unbiased non-exceedance probabilities (*F*) for many PDFs, according to Stedinger (2017), this is:

$$F = \frac{i-0.40}{n+0.20} \tag{18}$$

The expression for the standard error of fit (*EEA*) is:

$$EEA = \left[\frac{\sum_{i=1}^{n}(x_i-\hat{x}_i)^2}{(n-np)}\right]^{1/2} \tag{19}$$

Where:

x_i = values of the variable studied sorted from lowest to highest, whose number is *n*

$\hat{x}_i$ = estimated magnitudes, for the probability calculated with Equation (18) and the inverse solution of the PDF that is contrasted

np = number of fitting parameters, with 3 for the LP3, GVE and LOG distributions and 4 and 5 for the Kappa and Wakeby models, that will be applied

When Equation (18) is applied in the SDP, *n* is equal to the number of exceedances that defines the tested *vu*; logically, sorted from lowest to highest, whose order number is *i* and *np* is 3 for the Poisson-Pareto distribution (Equation (10)).

Mean absolute error

Abbreviated *EAM*, by its acronym in Spanish, its advantages lie in having the units of the variable (Willmott & Matsuura, 2005), like the *EEA*, and preventing the impact of the dispersed values from being squared and therefore, *EEA* ≥ *EAM*. Its expression is:

$$EAM = \frac{\sum_{i=1}^{n}|x_i - \hat{x}_i|}{n-np} \tag{20}$$

Stages of the suggested operating procedure

Stage 1: As indicated in the objectives section, the proposed operating procedure includes the contrast of predictions obtained with the SAM and SDP. Therefore, the first stage was adopting the maximum annual values in the record of maximum monthly flows, which displays nine years per page (IMTA, 2003), thus integrating the SAM. The five adopted PDFs (LP3, GVE, LOG, Kappa and Wakeby) are applied to this record to obtain their fit errors (equations (18) to (20)) and respective predictions, with the following seven return periods for contrast: 25, 50, 100, 500, 1 000, 5 000 and 10 000 years.

Stage 2: Again, from the file of maximum monthly flows, the basic or initial SDP is integrated, adopting the lowest maximum annual flow as the lower threshold value (x_0). Next, in each year, the maximum monthly flows that are greater than x_0 were selected and with them the SDP to be processed are integrated, as shown in the following section.

Stage 3: Starting from x_0, constant increments are assigned to the threshold value (*vu*) and the graph of average residual exceedances is obtained with Equation (8). In that graph, the linear trend section that allows to define, approximately, the threshold values are determined, where the Poisson-Pareto distribution (Equation (10)) is applicable.

Stage 4: For the selected threshold values (*vu*), equations (10) to (17) are applied to obtain the contrast predictions, in the aforementioned return periods. In addition, quality indicators of the achieved fit are obtained, based on equations (18) to (20). This stage concludes, finding the threshold value (*vu*) that leads to the lowest standard error of fit.

Stage 5: For each processed *vu*, the selection test of the distribution of the average number of exceedances per year is applied, based on the dispersion index (*d*) and the applicability of the Poisson-Pareto distribution is ratified.

Stage 6: Finally, the predictions of the SAM and the adopted SDP are contrasted, based on the relative error (*ER*, by its acronym in Spanish) in percentage, whose expression is:

$$ER = \frac{Q_{\text{SDP}}^{Tr} - Q_{\text{SAM}}^{Tr}}{Q_{\text{SDP}}^{Tr}} 100 \quad (21)$$

Where the *ER* is negative when the SAM prediction (Q_{SAM}^{Tr}) is greater than the prediction obtained with the SDP (Q_{SDP}^{Tr}); On the other hand, when the *ER* is positive, the flood estimated with the SDP was greater than that calculated with the SAM. Finally, the conclusions are formulated, relative to each processed record of floods.

Records to be processed

The five records that were integrated as SDP, are displayed and processed in progressive order of watershed area (*A*). Table 2 shows the SDP of the *Guamúchil* station on the Mocorito River in Hydrological Region No. 10 (Sinaloa), Mexico, with A = 1 645 km^2, n = 33 years, x_0 = 65.3 m^3/s and number of exceedances (NE) = 101. This record covers the year of the operation start, until the construction of the Eustaquio Buelna Dam.

Table 2. Maximum monthly flows (m^3/s) higher than the annual minimum at the *Guamúchil* hydrometric station, Mexico.

Year	Exceedances					
	1	**2**	**3**	**4**	**5**	**6**
1939	299.0	299.0	164.4	78.0		
1940	254.5					
1941	65.3					
1942	445.0	298.0	161.0	83.0		
1943	1 550.0	1 236.4	298.0	284.0	92.3	71.5
1944	391.8	125.0				
1945	916.0	336.0	276.0	228.7		
1946	241.0	197.3				
1947	530.0	133.0				
1948	648.0	548.0	195.9			
1949	375.0	145.4	89.2	72.8		
1950	272.3	74.2	69.3			
1951	422.3	409.7	82.8			
1952	376.8					

Year	Exceedances					
	1	**2**	**3**	**4**	**5**	**6**
1953	1 173.0	261.3	213.0			
1954	219.0	115.4	101.2			
1955	3 507.0	189.0	117.0			
1956	165.0	148.2	76.4			
1957	526.0	342.0				
1958	1 014.0	534.0	221.0	168.0		
1959	1 610.0	374.0	372.8	203.8		
1960	137.0	130.0	94.5	69.3		
1961	524.5	302.0	211.0	134.1		
1962	985.0	524.0	168.0	112.5		
1963	459.5	311.2	211.5	90.7		
1964	390.0	202.1	123.2			
1965	449.0	382.5				
1966	793.9	687.8				
1967	719.5	325.0	105.7			
1968	200.0	146.2	132.5	129.6		
1969	312.0	126.6	84.7			
1970	520.0	295.0	256.0			
1971	1 045.0	790.0	175.0			
SAM statistics:		μ_x	σ_x	*Cv*	*Cs*	*Ck*
		652.6	640.2	0.981	3.061	14.916

Table 3 presents the SDP, of the *Santa Rosa* station in the Valles River of Hydrological Region No. 26 (Pánuco), Mexico, with A = 3 521 km^2, n = 45 years, x_0 = 65.0 m^3/s and NE = 185. Table 4 presents the SDP, of the *Tempoal* station on the river of the same name in Hydrological

Region No. 26 (Pánuco), Mexico, with A = 5 275 km^2, n = 48 years, x_0 = 449.0 m^3/s and NE = 142.

Table 3. Maximum monthly flows (m^3/s) higher than the annual minimum at the *Santa Rosa* hydrometric station, Mexico.

Year	Exceedances					
	1	2	3	4	5	6
1958	1 176.0	732.0	326.0	187.0	163.0	103.0
1959	505.0	291.0	185.0			
1960	341.5	88.7	84.2			
1961	728.0	370.0	234.0	210.7		
1962	324.0	108.5	96.2	72.0		
1963	973.0	250.4	106.0			
1964	145.2	116.5	101.0	83.9		
1965	360.0	165.5	162.7	122.5		
1966	1 691.5	1 187.0	274.2	158.1	81.7	75.0
1967	912.0	550.0	440.0	391.0		
1968	769.5	208.6	130.0	127.8	125.7	120.3
1969	1 244.0	1 193.0	191.0	160.0		
1970	803.0	532.0	515.3	441.0	252.4	
1971	636.0	581.7	390.0	330.4	223.8	
1972	393.6	392.0	301.0	139.1	67.7	
1973	951.3	608.0	326.8	250.5	183.0	66.3
1974	1 480.0	302.7	217.6	106.8	66.3	
1975	662.0	575.0	341.0	90.0		
1976	2 588.0	288.0	251.5	211.6	196.0	87.9
1977	1 884.0	250.0	127.7	126.8		
1978	372.2	138.5	135.4	113.8		
1979	283.4	249.4	85.8			

Year	Exceedances					
	1	2	3	4	5	6
1980	161.8	158.3	73.3			
1981	289.0	150.0	128.9	121.0	72.3	
1982	65.0					
1983	525.0	261.2	170.6			
1984	461.7	348.4	124.9	121.7	77.1	
1985	361.4	244.4	240.9	127.3	120.7	
1986	353.3	247.0	195.0	179.2	82.1	
1987	187.0	184.0	162.4	74.4		
1988	898.0	329.0	148.9	103.5		
1989	267.0	96.4				
1990	800.0	154.9	152.0	126.9	107.9	
1991	1 163.0	255.0	100.0	96.3		
1992	1 272.8	158.7	130.1	122.9	97.4	83.9
1993	1 933.7	1249.0	1 057.0	281.5	102.2	77.0
1994	233.6	108.3				
1995	479.9	109.1	85.2			
1996	771.5	541.0	207.0	119.3		
1997	221.9	91.1	73.9			
1998	234.2	187.2				
1999	286.7	147.6	72.8			
2000	663.2	227.7	146.2	96.8		
2001	307.0	115.7	96.6			
2002	222.7	213.1	112.4	97.7	83.5	
SAM statistics:		μ_x	σ_x	*Cv*	*Cs*	*Ck*
		697.4	552.1	0.792	1.501	5.378

Table 4. Maximum monthly flows (m^3/s) higher than the annual minimum at the *Tempoal* hydrometric station, Mexico.

Year	Exceedances					
	1	2	3	4	5	6
1955	6 000.0	4 905.0	2 186.0			
1956	4 424.0	1 094.0	533.0			
1957	449.0	449.0				
1958	4 100.0	2 640.0	2 075.0	828.0	524.1	
1959	1 507.6	1 442.4	520.2			
1960	1 277.0	1 210.7	514.4	507.0		
1961	852.9	710.0	654.0	652.9	585.2	506.4
1962	739.2	702.0	694.0			
1963	1 800.0	533.0				
1964	748.0	737.0	540.0			
1965	792.7	738.0				
1966	1 778.0	813.0	508.2	494.9		
1967	2 245.0	1 184.0	1082.0			
1968	1 145.0	656.0	520.0	498.0		
1969	1 948.0	748.8	469.0			
1970	1 418.0	1 227.0	560.0	501.2		
1971	1 630.0	627.2				
1972	989.0	702.0	523.0			
1973	1 668.0	1 142.0	1 140.0	821.4	656.0	
1974	4 950.0	2 410.0	880.0	757.0		
1975	4 040.0	613.0	555.7			
1976	1 275.0	1 237.4	1 125.0	921.0	820.0	
1977	514.0					
1978	3 725.0	1 052.0	820.0	784.6		
1979	1 655.9	722.3	529.9			

Year	Exceedances					
	1	**2**	**3**	**4**	**5**	**6**
1980	1 162.0					
1981	2 020.0	1 805.0	1 520.0	1 492.0	503.0	
1982	539.6					
1983	868.0	559.1				
1984	4 030.0	1 680.0	467.7	467.1		
1985	1 882.0	1 064.0	711.9			
1986	476.0					
1987	1 765.0	1 080.0	702.0	463.0		
1988	3 265.0					
1989	649.0	614.0				
1990	1 611.0	569.0	483.0			
1991	3 532.0	2 100.0	795.0	580.0		
1992	2 291.0	887.2	645.0	510.4		
1993	6 120.0	1 773.6	1 050.0	764.8	678.4	
1994	1 133.0	709.0				
1995	742.0	513.0				
1996	683.0					
1997	905.0					
1998	1 266.9	1 096.3	489.0			
1999	2 693.7	978.5	615.7			
2000	641.2	561.2	537.4			
2001	1 847.9					
2002	926.4					
SAM statistics:		μ_x	σ_x	*Cv*	*Cs*	*Ck*
		1 931.7	1 453.3	0.752	1.394	4.380

Table 5 shows the SDP, of the *Huites* station on the Fuerte River in Hydrological Region No. 10 (Sinaloa), Mexico, with A = 26 057 km^2, n =

51 years, x_0 = 593.0 m^3/s and NE = 185. This record covers the year it began operation until the construction of the Luis Donaldo Colosio Dam.

Table 5. Maximum monthly flows (m^3/s) higher than the annual minimum at the *Huites* hydrometric station, Mexico.

Year	Exceedances					
	1	**2**	**3**	**4**	**5**	**6**
1942	2 531.0	2 037.6	1 868.8	780.0		
1943	14 376.0	3 283.0	1 396.9	1 085.0		
1944	2 580.0	1 262.5	1 024.8	976.0	768.0	
1945	1 499.2	1 250.0	1 191.3			
1946	1 164.8	808.4				
1947	1 127.3	754.8	718.8	634.0		
1948	3 215.0	799.0	623.2			
1949	10 000.0	2297.5	942.4	895.3	826.4	
1950	3 225.3	1 384.0	961.0			
1951	677.0					
1952	1 266.0	895.0				
1953	1 025.0	885.0				
1954	954.8	715.0				
1955	4 780.3	1 069.9	662.0			
1956	695.7					
1957	593.0					
1958	3 010.0	1 045.0	894.0	849.0	608.5	
1959	1 908.0	1 831.0	1 345.5	652.0		
1960	15 000.0	1 046.0	985.2	790.0	721.4	
1961	1 396.3	905.9	831.6	771.2	682.0	
1962	1 620.0	912.0	892.8			
1963	2 702.0	1 054.0	980.1	969.2		

Year	Exceedances					
	1	2	3	4	5	6
1964	1 319.1	938.5	912.4			
1965	1 944.0	1 787.6	663.1			
1966	2 420.0	892.4	738.0	688.9		
1967	2 505.8	1 310.5	1 192.8			
1968	1 534.3	1 478.3	1 118.0	1 019.6	703.8	
1969	1 508.0	736.0				
1970	1 558.0	1 330.0	970.0			
1971	2 200.0	1 176.0				
1972	2 225.0	2 040.0	1 142.0	1 109.0	732.1	732.0
1973	7 960.0	2 256.5	1 546.1	1 380.0	800.0	
1974	3 790.0	3 315.0	1 120.0	886.7		
1975	1 095.0	965.5	944.6			
1976	2 677.0	1 360.2	1 211.0			
1977	1 135.0	1 113.7	622.0			
1978	4 790.0	1 750.0	1 119.2	756.0		
1979	6 860.0	3 347.5	1 001.0	820.0		
1980	1 496.0	1 475.0	1 197.8	660.2		
1981	4 828.1	2 448.0	2 280.0	2 052.0	1 000.3	
1982	2 450.0	2 422.0	1 571.6	1 085.2	624.0	
1983	8 275.0	1 439.0	1 400.0	1 006.7	893.0	877.6
1984	5 580.0	1 623.0	1 132.0	960.0	624.5	
1985	3 585.0	1 250.0	1 121.4	925.0	820.4	
1986	1 348.8	1 329.3	944.7			
1987	1 429.2	1 218.9	679.2			
1988	1 866.3	1 494.3	666.7			
1989	1 868.5	1 413.9	1 378.9	1 230.3	1 164.7	
1990	11 558.6	3 544.2	970.1	815.8	693.3	
1991	2 563.2	2 370.0	1 721.5	1 517.6	1 266.7	

Year	Exceedances					
	1	2	3	4	5	6
1992	2 025.3	1 564.9	1 348.9	787.3		
SAM statistics:		μ_x	σ_x	*Cv*	*Cs*	*Ck*
		3 328.3	3 312.7	0.995	2.210	7.668

Finally, Table 6 presents the SDP, of the *Pánuco* station on the river of the same name in Hydrological Region No. 26 (Pánuco), Mexico, with A = 65 577 km^2, n = 31 years, x_0 = 829.9 m^3/s and NE = 109. In this record, in 1992, a seventh exceedance of 883.0 m^3/s was not included, to seek to standardize the process of capturing and reading all records to six exceedances, including zeros. The above facilitates the numerical processing of each record, with any value of the threshold *vu* greater than x_0.

Table 6. Maximum monthly flows (m^3/s) higher than the annual minimum at the *Pánuco* hydrometric station, Mexico.

Year	Exceedances					
	1	2	3	4	5	6
1972	2 880.0	2 469.5	2 375.4	1 129.0		
1973	3 234.0	3 150.0	3 071.5	2 777.8	2 150.7	
1974	7 300.0	6 928.0	3 708.4	1 653.0	900.0	845.6
1975	4 138.0	2 402.0	2 181.0	1 676.0		
1976	3 886.0	2 819.0	2 492.0	2 313.0	2 064.0	
1977	1 968.0	1 067.0	955.8			
1978	3 471.0	2 794.0	1 278.0	1 106.0	859.5	
1979	3 525.0	1 544.0				

Tecnología y Ciencias del Agua
OPEN ACCESS
Check for updates

Year	Exceedances					
	1	**2**	**3**	**4**	**5**	**6**
1980	2 753.0	2 187.0				
1981	3 240.0	3 154.0	2 987.0	2 717.0	2 710.0	
1982	829.9					
1983	2631.0	2 017.0	1 683.0	1 013.0		
1984	4 236.0	3 105.0	2 834.0	1 265.0	893.2	887.0
1985	2 906.0	2 430.0	1 422.0			
1986	1 641.0	1 538.9	1 414.0			
1987	2 775.0	2 745.0	1 064.0	1 052.0		
1988	3 170.0	1 935.0	1 145.0	1 088.0		
1989	1 530.0	1 467.0				
1990	3 445.0	1 635.0	1 560.0	1 235.0		
1991	3 619.0	3 360.0	2 100.0	1 100.0	847.2	
1992	2 890.0	2 700.0	2 146.0	1 180.0	1 165.0	1 045.4
1993	5 400.0	4 033.0	3 691.0	3 257.2	1 420.5	
1994	2 325.0	1 179.0				
1995	1 976.0	1 289.2				
1996	2 546.0	2 193.0				
1997	2 334.0					
1998	2 486.6	2 282.6	1 252.7			
1999	2 701.3	2 006.0	1 351.4			
2000	1 817.9	1 396.4	1 063.7			
2001	2 306.4	1 468.4	1 160.1			
2002	1 899.4	872.2				
SAM statistics:		μ_x	σ_x	*Cv*	*Cs*	*Ck*
		2 963.2	1 220.6	0.412	1.579	7.609

Description of results

Randomness verification in SAMs

Based on the Wald-Wolfowitz Test, presented and applied by Bobée and Ashkar (1991), Rao and Hamed (2000), and Meylan *et al*. (2012), the independence and stationarity of the SAM was tested, presented in the second columns of Table 2, Table 3, Table 4, Table 5 and Table 6. The *U* statistic of the test, for each hydrometric station were: -1.414, 2.152, -0.111, 0.446 and 1.190, respectively.

The previous statistics detect the SAM of the *Santa Rosa* station as non-random, therefore, the following tests were applied: a general test, the Von Neumann test, and six specific tests, the Anderson and Sneyers tests for persistence, and the Kendall and Spearman tests for trend, Bartlett's variability and Cramer's change in the mean. Such SAM was found to show persistence with a serial correlation coefficient of order one of 0.305.

Khaliq, Ouarda, Ondo, Gachon and Bobée (2006) present three methods to eliminate persistence from hydrological records. It has been verified —with the simplest criterion (Campos-Aranda, 2018; Campos-Aranda, 2023)— that by eliminating such a deterministic component, the results of frequency analysis or *predictions* are lower. Therefore, such SAM will not be corrected, but its acceptance or functionality will be verified with the average residual exceedance graph.

AFC at *Guamúchil* station

In Table 7, the GVE distribution has been selected because it leads to the smallest fit errors and predictions that are considered representative of those obtained with the other four PDFs.

Table 7. Fitting and prediction errors (m^3/s) of the three reference PDFs and two widely applied PDFs in the annual flood record of the *Guamúchil* hydrometric station, Mexico.

PDF	*EEA*	*EAM*	Return periods, in years						
			25	50	100	500	1 000	5 000	10 000
LP3	176.0	69.7	1 948	2 507	3 150	5 024	6 020	8 861	10 345
GVE	153.4	61.9	1 927	2 609	3 492	6 701	8 813	16 506	21 569
LOG	161.7	63.9	1 873	2 562	3 487	7 089	9 612	19 479	26 397
KAP	159.3	64.0	1 894	2 588	3 515	7 071	9 530	18 999	25 557
WAK	153.6	66.5	1 957	2 645	3 513	6 507	8 384	14 859	18 920

In Figure 2, corresponding to the graph of average residual exceedances, of the *Guamúchil* station, a barely linear or very approximate trend can be accepted, from x_0 = 65.3 m^3/s to *vu* = 350 m^3/s.

Tecnología y Ciencias del Agua
OPEN ACCESS
Check for updates

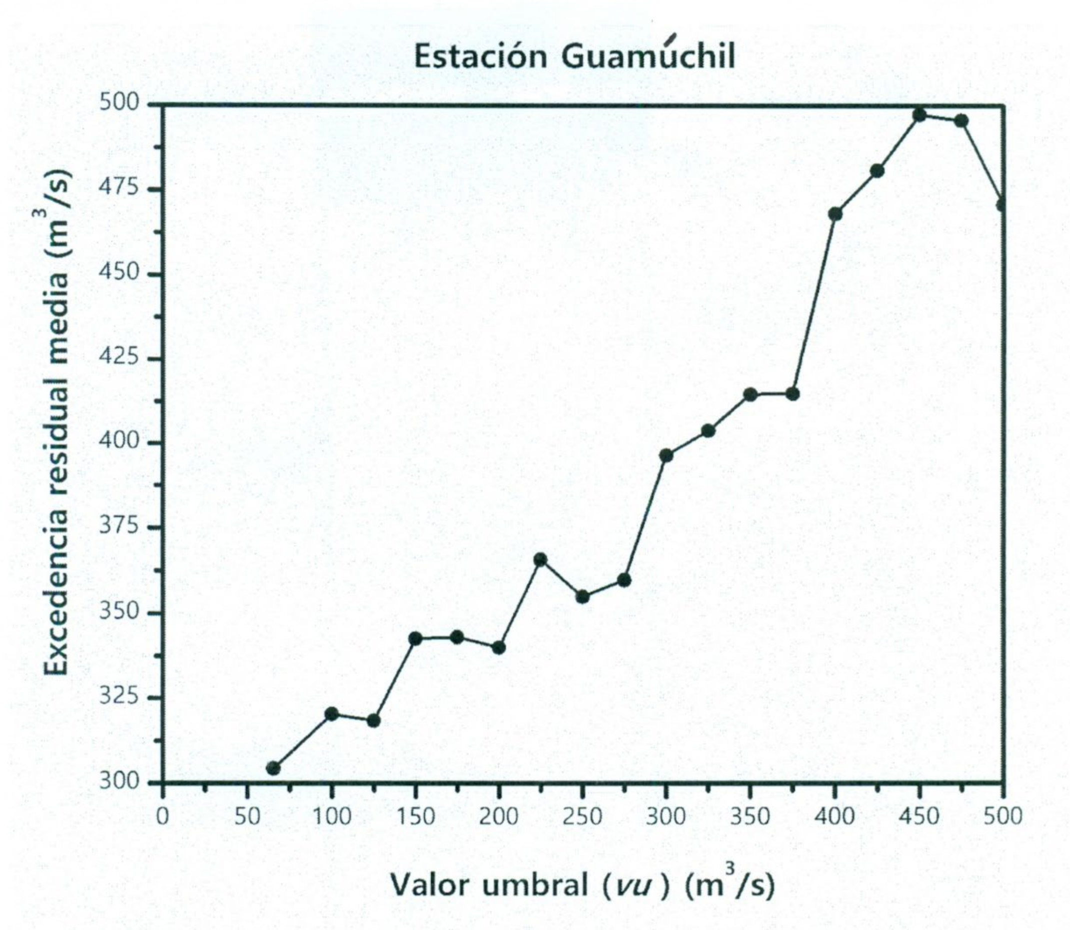

Figure 2. Graph of average residual exceedances in the SDP of the *Guamúchil* station, Mexico. Abscissa axis: Threshold value (m^3/s); ordinate axis: Average residual exceedance (m^3/s).

In Table 8, for the 12 threshold values adopted from x_0 to *vu* = 350 m^3/s, which leads to a number of exceedances equal to the record size (*n* = 33), the minimum *EEA* was obtained with a *vu* = 300 m^3/s, but its predictions are lower than those obtained with a *vu* = 275 m^3/s, which reports the most severe predictions and also defines a low *EEA*. For the shaded row, its *ER* (Equation (21)) ranges from 2.6 to 16.4 %, for the return periods of 25 and 10 000 years.

Table 8. Fitting errors and predictions (m^3/s) of the partial duration series obtained with the Poisson-Pareto distribution applied to the flood record of the *Guamúchil* hydrometric station, Mexico.

x_0 *vu*	NE	*EEA*	*EAM*	Return periods, in years						
				25	50	100	500	1 000	5 000	10 000
65.3	101	419	403	1 950	2 541	3 266	5 642	7 067	11 741	14 542
100	86	364	351	1 929	2 497	3 186	5 411	6 724	10 963	13 464
125	80	338	322	1 976	2 613	3 411	6 134	7 826	13 597	17 181
150	69	292	283	1 924	2 484	3 159	5 322	6 589	10 644	13 018
175	64	261	250	1 952	2 558	3 309	5 814	7 340	12 433	15 533
200	60	230	214	1 982	2 657	3 522	6 592	8 567	15 579	20 089
225	52	193	182	1 956	2 570	3 331	5 886	7 449	12 694	15 901
250	50	162	148	1 978	2 677	3 588	6 916	9 115	17 146	22 448
275	46	131	112	1 979	2 711	3 684	7 365	9 872	19 356	25 813
300	39	121	75	1 981	2 659	3 530	6 632	8 634	15 765	20 366
325	36	125	59	1 977	2 667	3 563	6 807	8 933	16 634	21 678
350	33	145	66	1 973	2 667	3 572	6 873	9 051	16 998	22 238

x_0 = minimum threshold value, in m^3/s.

vu = tested threshold value, in m^3/s.

NE = number of exceedances.

EEA = standard error of fit, in m^3/s.

EAM = mean absolute error, in m^3/s.

Since the number of years of the record (n) is 33, then $\chi^2_{0.025}$ = 18.3 and $\chi^2_{0.975}$ = 49.5, according to values in Table 1. Table 9 verifies that the fit of the Poisson distribution-Pareto is acceptable for *vu* greater than 150 m^3/s. Furthermore, a similarity of orders of magnitude is observed in the shape parameters (k) starting at *vu 0* = 200 m^3/s.

Table 9. Results of the λ distribution selection test and fit parameters of the Poisson-Pareto distribution applied to the SDP of the floods of the *Guamúchil* hydrometric station, Mexico.

x_0 *vu*	*E*	*V*	*d*	DE	ERM	Fit parameters		
						*u**	*a**	*k*
65.3	3.061	1.148	12.001	BI	304.4	350.0387	298.6488	-0.293721
100	2.606	1.027	12.606	BI	320.3	353.4125	301.6087	-0.279857
125	2.424	1.032	13.624	BI	318.4	345.2760	286.4088	-0.326209
150	2.091	1.295	19.816	PO	342.6	353.5177	304.6174	-0.272875
175	1.939	1.208	19.939	PO	343.1	349.4684	291.1793	-0.308004
200	1.818	1.179	20.752	PO	340.0	345.4118	270.2383	-0.358669
225	1.576	1.032	20.960	PO	365.8	347.9105	289.9379	-0.312327
250	1.515	0.977	20.635	PO	355.0	348.7276	256.9866	-0.382545
275	1.394	1.027	23.568	PO	359.9	350.4437	243.0065	-0.411040
300	1.182	0.755	20.438	PO	396.4	343.5819	268.8417	-0.361499
325	1.091	0.749	21.980	PO	403.7	347.3032	260.5336	-0.375388
350	1.000	0.667	21.333	PO	414.6	350.0000	256.4193	-0.381468

x_0 = minimum threshold value, in m^3/s.

vu = tested threshold value, in m^3/s.

E = average of the number of exceedances per year.

V = variance of the number of exceedances per year.

d = dispersion index (Equation (7)).

DE = PDF of the number of exceedances per year (λ).

BI = Binomial DE.

PO = Poisson DE.

BN = Negative binomial DE.

ERM = average residual exceedance (by its acronym in Spanish), m^3/s.

AFC at *Santa Rosa* station

In Table 10, it is observed that the PDFs that lead to the lowest fitting errors are the Wakeby and the Kappa, but their predictions are very reduced compared to those of the LOG and GVE models. Therefore, the LP3 distribution was adopted, which shows low fitting errors and superior predictions.

Table 10. Fitting and prediction errors (m^3/s) of the three reference PDFs and two widely applied PDFs in the annual flood record of the *Santa Rosa* hydrometric station, Mexico.

PDF	*EEA*	*EAM*	Return periods, in years						
			25	50	100	500	1 000	5 000	10 000
LP3	64.2	49.8	1 997	2 480	3 005	4 393	5 066	6 818	7 657
GVE	76.7	61.6	1 941	2 453	3 051	4 860	5 869	8 935	10 646
LOG	92.6	74.5	1 897	2 444	3 122	5 422	6 846	11 701	14 713
KAP	50.0	43.4	1 964	2 296	2 603	3 223	3 456	3 928	4 104
WAK	48.7	40.2	1 967	2 370	2 780	3 761	4 197	5 241	5 705

Figure 3 of the average residual exceedance graph of the *Santa Rosa* station is one of the well-formed ones, whose section with an approximate linear trend varies from x_0 = 65 m^3/s to a vu = 400 m^3/s.

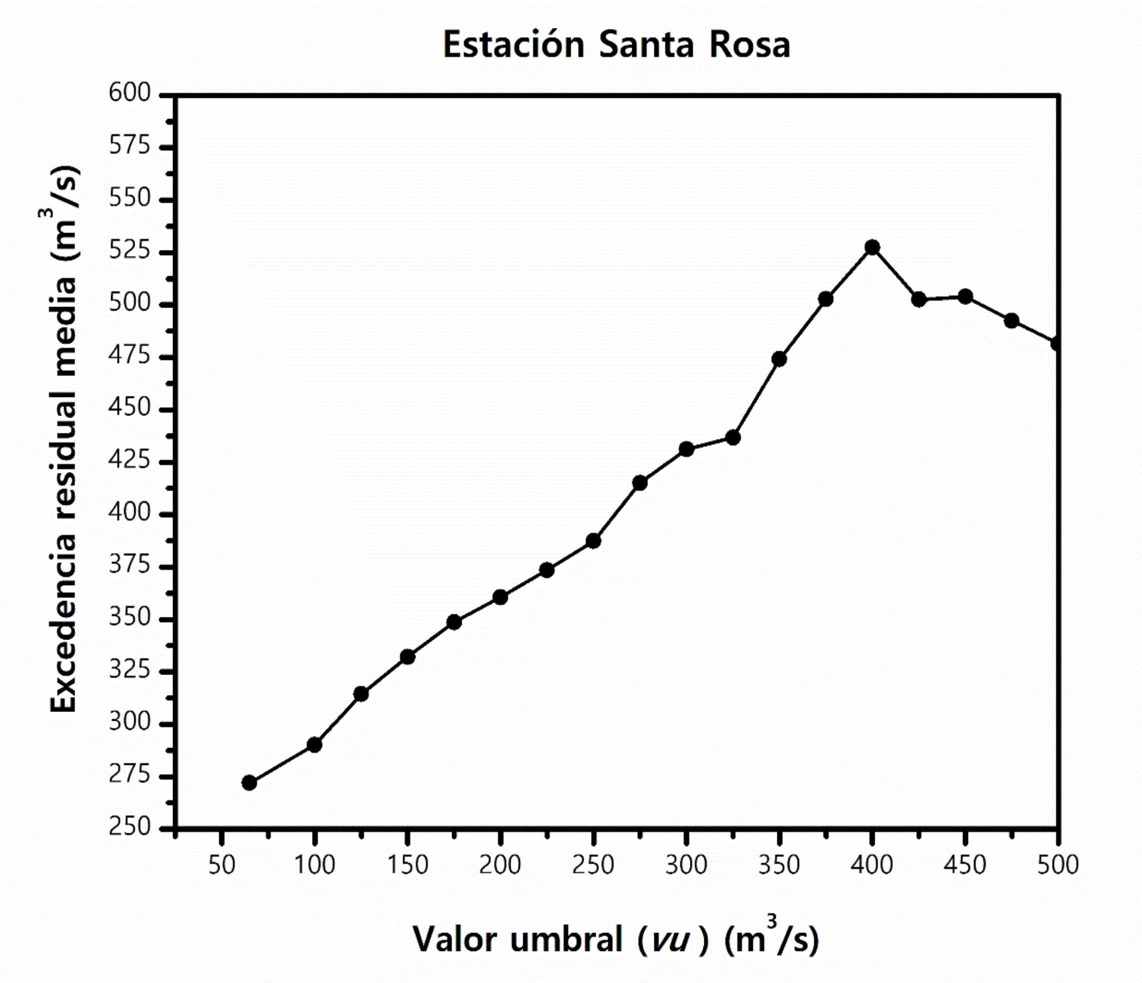

Figure 3. Graph of average residual exceedances in the SDP of the *Santa Rosa* station, Mexico. Abscissa axis: Threshold value (m^3/s); ordinate axis: Average residual exceedance (m^3/s).

In Table 11, it is observed that the *vu* that leads to the lowest fit error is 375 m^3/s, but its predictions in high return periods are reduced, therefore, the *vu* value of 325 m^3/s was adopted, with stable predictions, before a significant reduction. The predictions adopted define *ER*

(Equation (21)) ranging from 3.5 to 32.9 %, for the return periods of 25 and 10 000 years.

Table 11. Fitting errors and predictions (m^3/s) of the SDP obtained with the Poisson-Pareto distribution applied to the flood record of the *Santa Rosa* hydrometric station, Mexico.

x*0** ***vu	**NE**	***EEA***	***EAM***	**Return periods, in years**						
				25	**50**	**100**	**500**	**1 000**	**5 000**	**10 000**
65	185	595	502	2 195	2 967	3 970	7 614	10 007	15 704	24 415
100	153	546	447	2 216	3 015	4 063	7 930	10 508	20 022	26 359
125	130	472	400	2 175	2 915	3 865	7 242	9 419	17 159	22 145
150	114	417	357	2 150	2 855	3 749	6 853	8 814	15 628	19 929
175	101	367	316	2 131	2 807	3 654	6 537	8 324	14 411	18 182
200	91	329	278	2 129	2 804	3 648	6 521	8 300	14 353	18 100
225	82	292	239	2 129	2 802	3 643	6 501	8 268	14 268	17 978
250	74	253	199	2 125	2 789	3 615	6 400	8 110	13 872	17 410
275	65	189	153	2 087	2 670	3 365	5 539	6 786	10 685	12 920
300	59	148	116	2 072	2 622	3 265	5 210	6 291	9 561	11 377
325	55	125	90	2 070	2 620	3 264	5 214	6 298	9 584	11 411
350	48	64	47	2 022	2 462	2 938	4 198	4 815	6 451	7 252
375	43	57	46	1 980	2 337	2 697	3 543	3 912	4 778	5 155
400	39	103	96	1 938	2 225	2 494	3 050	3 264	3 709	3 880

Identical to that of Table 8.

Since the number of years of the record (n) is 45, then $\chi^2_{0.025}$ = 27.6 and $\chi^2_{0.975}$ = 64.2, according to values in Table 1. In Table 12, the acceptance of the Poisson-Pareto model is verified from a *vu* of 125 m³/s and it is observed that as the *vu* increases the shape parameters (k) decrease, showing a significant reduction from the *vu* of 325 to that of 350 m³/s.

Table 12. Results of the λ distribution selection test and fit parameters of the Poisson-Pareto distribution applied to the SDP of the floods of the *Santa Rosa* hydrometric station, Mexico.

x_0 *vu*	*E*	*V*	*d*	DE	ERM	Fit parameters		
						*u**	*a**	*k*
65	4.111	1.521	16.279	BI	372.0	381.2178	288.7109	-0.377817
100	3.400	1.840	23.812	BI	290.2	377.4422	285.2420	-0.390893
125	2.889	1.832	27.904	PO	314.4	384.8855	294.7532	-0.360205
150	2.533	1.849	32.112	PO	332.2	389.3301	300.4796	-0.341370
175	2.244	1.918	37.601	PO	348.7	392.7313	306.1809	-0.324451
200	2.022	1.666	36.253	PO	360.6	392.8934	306.6225	-0.323680
225	1.822	1.524	36.798	PO	373.5	392.5919	307.1636	-0.322220
250	1.644	1.607	42.996	PO	387.4	392.7665	310.1531	-0.315750
275	1.444	1.358	41.368	PO	415.1	394.3980	340.0730	-0.253666
300	1.311	1.370	45.972	PO	431.1	393.2664	354.9367	-0.225466
325	1.222	1.373	49.422	PO	436.8	394.3612	353.5649	-0.226593
350	1.067	1.084	44.733	PO	474.1	377.2306	423.4738	-0.113324
375	0.956	1.154	53.118	PO	502.7	352.3812	497.3375	-0.010191
400	0.867	0.960	48.738	PO	527.5	316.5242	587.4352	-0.098024

Identical to that of Table 9.

AFC at *Tempoal* station

In Table 13, the distributions that define the smallest fit errors are the Kappa and the Wakeby, the second was adopted because it leads to somewhat higher predictions than those obtained with the LP3; but both are very small compared to the predictions of the LOG and GVE distributions, which also report the largest fit errors.

Table 13. Fitting and prediction errors (m^3/s) of the three reference PDFs and two of general application in the record of annual floods from the *Tempoal* hydrometric station, Mexico.

PDF	*EEA* (m^3/s)	*EAM* (m^3/s)	Return periods, in years						
			25	50	100	500	1 000	5 000	10 000
LP3	273.0	206.2	4 993	5 918	6 845	9 005	9 935	12 089	13 011
GVE	359.7	210.9	5 220	6 630	8 300	13 494	16 464	25 730	31 031
LOG	397.8	240.0	5 100	6 590	8 460	14 923	18 999	33 165	42 110
KAP	221.0	170.6	5 301	6 178	6 979	8 578	9 168	10 343	10 776
WAK	267.3	159.7	5 324	6 438	7 577	10 321	11 547	14 501	15 821

Figure 4 of the graph of the average residual exceedance of the *Tempoal* station does not show a wide section with a linear trend and even at the *vu* of 1050 m^3/s a marked decrease is observed, because the number of exceedances at such threshold is not reduces (Table 14).

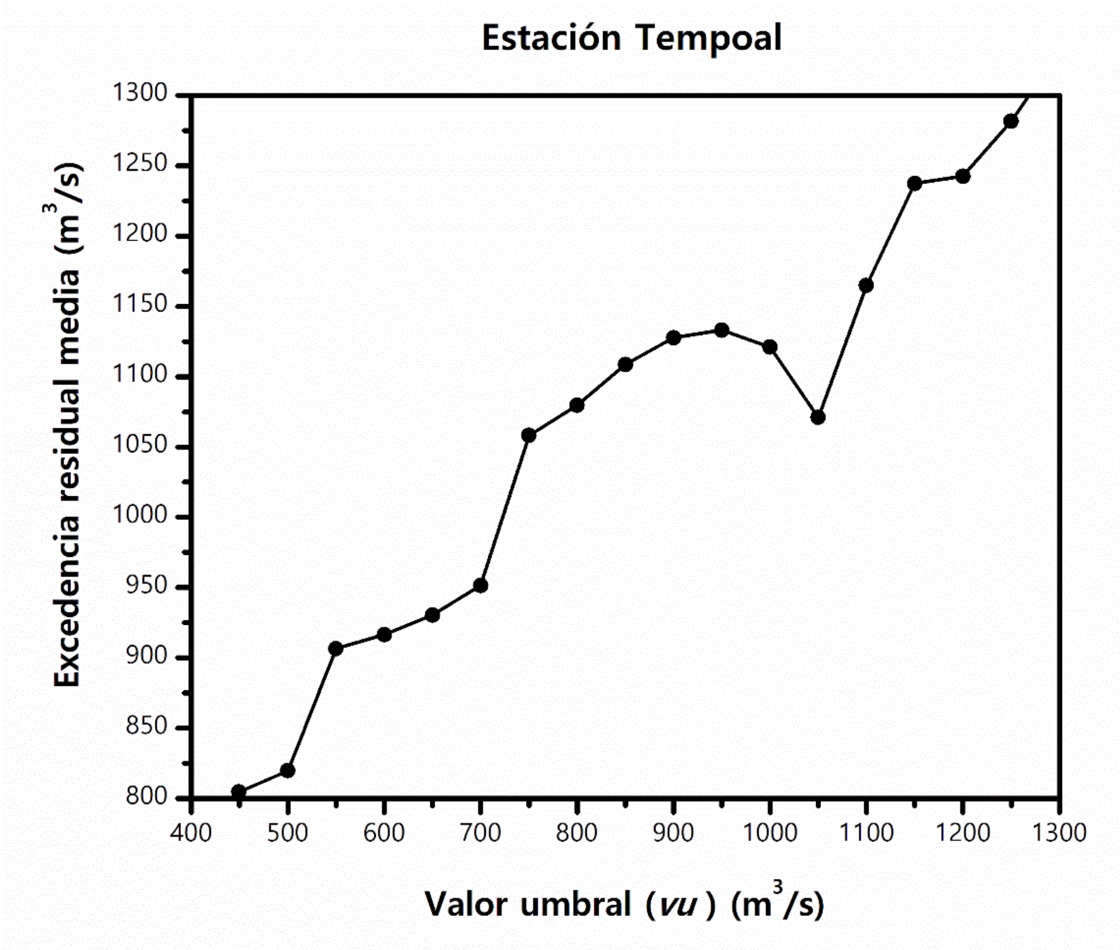

Figure 4. Graph of average residual exceedances in the SDP of the *Tempoal* station, Mexico. Abscissa axis: Threshold value (m^3/s); ordinate axis: Average residual exceedance (m^3/s).

Table 14. Fitting errors and predictions (m^3/s) of the SDP obtained with the Poisson-Pareto distribution applied to the flood record of the *Tempoal* hydrometric station, Mexico.

x_0 *vu*	NE	*EEA*	*EAM*	Return periods in years						
				25	50	100	500	1 000	5 000	10 000
449	142	1 290	1 051	5 820	7 819	10 417	19 847	26 040	48 524	63 283
600	104	952	789	5 673	7 444	9 666	17 256	21 976	38 099	48 119
700	90	842	642	5 702	7 537	9 867	17 987	23 130	41 061	52 411
750	77	653	519	5 520	6 986	8 715	14 025	17 020	26 220	31 407
800	72	559	453	5 506	6 946	8 636	13 775	16 648	25 389	30 272
850	67	486	388	5 471	6 846	8 436	13 152	15 727	23 364	27 531
900	63	432	335	5 454	6 801	8 347	12 882	15 330	22 510	26 385
950	60	402	297	5 459	6 823	8 396	13 047	15 579	23 061	27 131
1000	58	407	272	5 491	6 943	8 657	13 923	16 895	26 026	31 176
1050	58	477	272	5 541	7 214	9 298	16 336	20 666	35 287	44 279
1100	51	371	191	5 493	6 930	8 615	13 747	16 616	25 352	30 234
1150	46	337	195	5 450	6 714	8 126	12 070	14 106	19 797	22 732
1200	44	365	230	5 450	6 720	8 143	12 137	14 207	20 019	23 029
1250	41	399	308	5 425	6 606	7 890	11 324	13 023	17 565	19 811

Identical to that of Table 8.

In Table 14, it is observed that the *vu* that leads to the smallest fitting error is 1 150 m^3/s, but its predictions are considered reduced compared to those of the previous three *vu*. The predictions adopted correspond to *vu* = 1 100 m^3/s which shows the lowest *EAM*. Such predictions define *ER*

(Equation (21)) ranging from 3.1 to 47.7 %, for the return periods of 25 and 10 000 years. Such predictions are quite similar to those obtained with the GVE distribution in Table 13.

Since the number of years of the record (*n*) is 48, then $\chi^2_{0.025}$ = 30.0 and $\chi^2_{0.975}$ = 67.8, according to values in Table 1. In Table 15, the acceptance of the Poisson-Pareto model is verified from the *vu* of 600 m^3/s and it is observed that as the *vu* increases the shape parameters (*k*) decrease, showing a significant reduction from the *vu* of 1 100 to 1 150 m^3/s.

Table 15. Results of the λ distribution selection test and fit parameters of the Poisson-Pareto distribution applied to the SDP of the floods of the *Tempoal* hydrometric station, Mexico.

x_0 *vu*	*E*	*V*	*d*	DE	ERM	Fit parameters		
						*u**	*a**	*k*
449	2.958	1.748	27.775	BI	804.7	1 115.899	749.2466	-0.377337
600	2.167	1.681	36.455	PO	916.5	1 142.518	794.1446	-0.327161
700	1.875	1.609	40.342	PO	951.5	1 138.045	774.8971	-0.343966
750	1.604	1.781	52.176	PO	1058.1	1 153.030	901.7773	-0.238638
800	1.500	1.542	48.306	PO	1079.6	1 153.048	912.0601	-0.230576
850	1.396	1.197	40.321	PO	1108.6	1 152.777	939.9544	-0.209301
900	1.313	1.215	43.503	PO	1127.6	1 152.220	952.9009	-0.199577
950	1.250	1.188	44.650	PO	1133.1	1 155.341	941.5792	-0.206447
1000	1.208	1.207	46.932	PO	1121.0	1 165.150	892.5770	-0.238993
1050	1.208	1.207	46.932	PO	1071.0	1 192.564	776.2032	-0.317538
1100	1.063	1.017	44.984	PO	1164.7	1 154.680	908.2670	-0.231036
1150	0.958	0.915	44.871	PO	1237.3	1 105.870	1033.393	-0.159129
1200	0.917	0.910	46.644	PO	1242.5	1 110.129	1025.571	-0.162816
1250	0.854	0.833	45.830	PO	1281.6	1 074.198	1104.654	-0.121403

Identical to that of Table 9.

AFC at *Huites* station

In Table 16, the Kappa distribution leads to the smallest fit errors and its predictions, although reduced, compared to those obtained with the other PDFs, are confirmed by the Wakeby model. In this registry, the predictions of the SDP will have relevant importance.

Table 16. Fitting and prediction errors (m^3/s) of the three reference PDFs and two widely applied PDFs in the annual flood record of the *Huites* hydrometric station, Mexico.

PDF	***EEA***	***EAM***	**Return periods, in years**						
			25	**50**	**100**	**500**	**1 000**	**5 000**	**10 000**
LP3	957.1	418.3	11 008	15 613	21 776	45 128	60 939	119 793	159 113
GVE	930.1	465.4	10 113	14 166	19 665	41 313	56 611	116 988	159 645
LOG	984.4	496.9	9 812	13 837	19 452	42 767	60 041	132 094	185 524
KAP	766.5	398.9	10 709	14 470	19 089	34 271	43 383	73 314	91 245
WAK	798.5	423.5	10 580	14 355	19 057	34 929	44 692	77 639	97 904

Figure 5 of the average residual exceedance graph of the *Huites* station shows a reduced section with a linear trend between *vu* of 1 300 to 1 600 m^3/s. This graph at its beginning and end shows slightly curved sections.

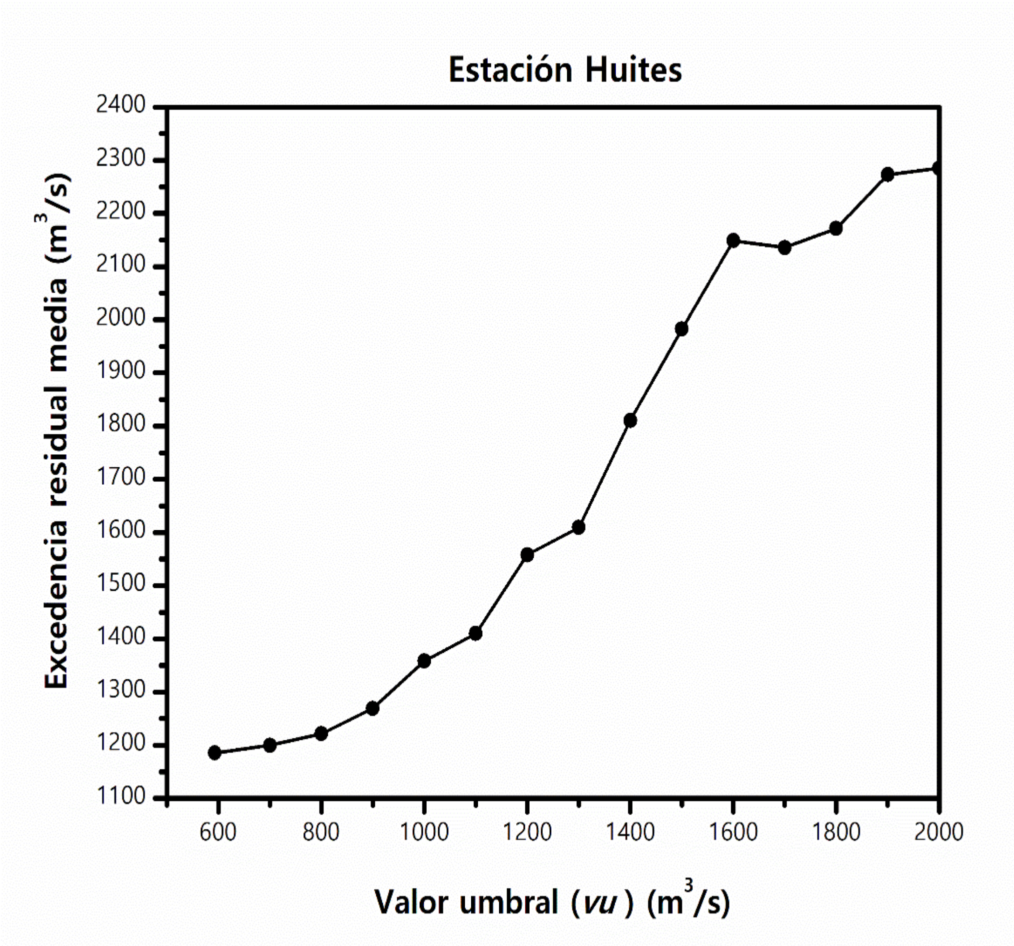

Figure 5. Graph of average residual exceedances in the SDP of the *Huites* station, Mexico. Abscissa axis: Threshold value (m^3/s); ordinate axis: Average residual exceedance (m^3/s).

In Table 17, the *vu* that leads to the lowest standard error of fit is 1 500 m^3/s, whose predictions were only lower than the maximum ones obtained with the LOG distribution. The *ER* (Equation (21)) of the return periods of 25 and 10 000 years were -5.1 and 44.9 %.

Table 17. Fitting errors and predictions (m^3/s) of the SDP obtained with the Poisson-Pareto distribution applied to the flood record of the *Huites* hydrometric station, Mexico.

x_0 *vu*	NE	*EEA*	*EAM*	Return periods, in years						
				25	**50**	**100**	**500**	**1 000**	**5 000**	**10 000**
593	185	2 157	1 939	9 347	12 544	16 695	31 752	41 635	77 485	101 002
700	167	2 119	1 800	9 622	13 226	18 063	36 650	49 489	98 835	132 902
800	151	2 074	1 648	9 798	13 756	19 221	41 315	57 268	121 770	168 317
900	134	1 967	1 467	9 886	14 099	20 038	44 964	63 552	141 577	199 732
1000	116	1 742	1 261	9 894	14 153	20 183	45 682	64 821	145 765	206 488
1100	104	1 614	1 066	9 886	14 328	20 732	48 722	70 331	164 780	237 659
1200	88	1 303	838	9 899	14 195	20 299	46 262	65 850	149 185	212 022
1300	80	1 201	695	9 850	14 281	20 679	48 711	70 399	165 429	238 913
1400	67	951	568	9 945	14 161	20 081	44 751	63 035	139 230	195 671
1500	58	853	463	10 188	14 059	19 600	41 728	57 536	120 678	165 761
1600	51	873	426	10 089	13 923	19 030	38 423	51 676	102 000	136 373
1650	49	972	473	10 001	13 988	19 428	40 984	56 275	116 860	159 815

Identical to that of Table 8.

Since the number of years of the record (*n*) is 51, then $\chi^2_{0.025}$ = 32.4 and $\chi^2_{0.975}$ = 71.4, according to values in Table 1. In Table 18, the acceptance of the Poisson-Pareto model is verified from a *vu* of 1 000 m^3/s and it is observed that in the *vu* interval of 900 to 1 300 m^3/s, the shape parameter (*k*) fluctuates from -0.49 to -0.53.

Table 18. Results of the *λ* distribution selection test and fit parameters of the Poisson-Pareto distribution applied to the SDP of the floods of the *Huites* hydrometric station, Mexico.

x_0 vu	*E*	*V*	*d*	DE	VRM	Fit parameters		
						*u**	*a**	*k*
593	3.627	3.627	21.600	BI	1185.5	1 818.554	1 200.563	-0.376704
700	3.275	3.275	29.388	BI	1199.8	1 765.255	1 142.750	-0.424001
800	2.961	1.881	31.762	BI	1221.6	1 722.484	1 082.439	-0.465250
900	2.627	1.606	30.567	BI	1269.2	1 693.315	1 033.405	-0.495535
1000	2.275	1.689	37.137	PO	1358.3	1 688.471	1 022.298	-0.501614
1100	2.039	1.606	39.385	PO	1409.9	1 675.564	969.239	-0.528153
1200	1.725	1.415	40.998	PO	1558.0	1 683.167	1 013.682	-0.506420
1300	1.569	1.265	40.319	PO	1609.4	1 684.469	959.961	-0.530167
1400	1.314	1.117	42.523	PO	1810.7	1 669.746	1 056.082	-0.489724
1500	1.137	1.099	48.310	PO	1982.7	1 643.044	1 145.051	-0.455322
1600	1.000	0.941	47.059	PO	2148.9	1 600.000	1 259.063	-0.414093
1700	0.961	0.940	48.900	PO	2135.7	1653.323	1156.329	-0.448775

Identical to that of Table 9.

AFC at *Pánuco* station

In Table 19, the Wakeby distribution reports the lowest fit errors, but it is observed that the predictions are quite high in return periods greater than 500 years, compared to those of the rest of the FDPs applied. The LOG distribution is adopted to define the following smallest fit errors, whose predictions are confirmed by the Kappa model.

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

Table 19. Fitting and prediction errors (m^3/s) of the three reference PDFs and two widely applied PDFs in the annual flood record of the *Pánuco* hydrometric station, Mexico.

PDF	*EEA*	*EAM*	Return periods, in years						
			25	50	100	500	1 000	5 000	10 000
LP3	265.4	182.8	5 519	6 257	7 010	8 846	9 682	11 744	12 690
GVE	269.3	168.1	5 442	6 158	6 583	6 888	8 658	9 455	12 266
LOG	232.8	140.7	5 389	6 239	7 202	9 992	11 488	15 853	18 199
KAP	235.6	144.6	5 396	6 800	7 216	10 021	11 526	15 925	18 292
WAK	196.0	128.6	5 387	6 475	7 804	12 130	14 706	23 098	28 095

In Figure 6 relative to the average residual exceedance graph of the *Pánuco* station, a decreasing trend of the *vu* drawn from 1 450 to 2 700 m^3/s is observed. The behavior of the *vu* from 2 500 to 2 800 m^3/s is influenced by the number of exceedances, as seen in Table 20.

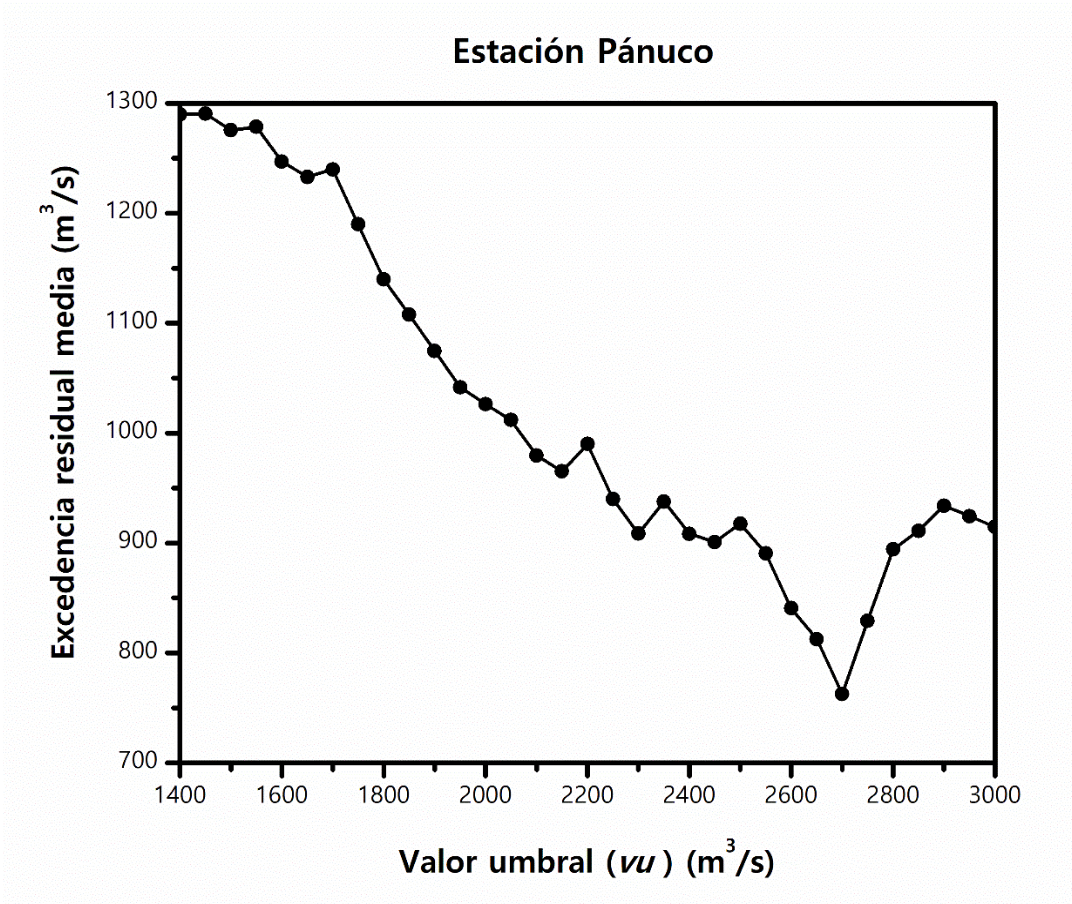

Figure 6. Graph of average residual exceedances in the SDP of the *Pánuco* station, Mexico. Abscissa axis: Threshold value (m^3/s); ordinate axis: Average residual exceedance (m^3/s).

Table 20. Fitting errors and predictions (m^3/s) of the SDP obtained with the Poisson-Pareto distribution applied to the flood record of the *Pánuco* hydrometric station, Mexico.

x_0 *vu*	NE	*EEA*	*EAM*	Return periods, in years						
				25	50	100	500	1 000	5 000	10 000
829.9	109	1 561	1 557	5 337	5 639	5 886	6 295	6 420	6 627	6 690
1 000	100	1 435	1 435	5 343	5 652	5 905	6 329	6 459	6 677	6 744
1 250	84	1 229	1 227	5 009	5 204	5 349	5 560	5 615	5 695	5 716
1 500	73	1 066	1 064	4 926	5 095	5 219	5 390	5 432	5 492	5 507
1 750	64	920	914	4 923	5 094	5 219	5 394	5 438	5 499	5 514
2 000	59	736	739	5 434	5 857	6 236	6 970	7 233	7 743	7 925
2 250	50	541	532	5 668	6 282	6 888	8 263	8 843	10 158	10 711
2 400	44	425	398	5 762	6 482	7 226	9 048	9 875	11 902	12 823
2 500	39	373	317	5 743	6 442	7 158	8 887	9 661	11 530	12 367
2 550	38	333	272	5 813	6 615	7 472	9 698	10 768	13 547	14 882
2 600	38	287	224	5 911	6 910	8 061	11 448	13 284	18 692	21 623
2 650	37	259	188	5 956	7 097	8 477	12 900	15 493	23 803	28 674
2 700	37	250	162	5 977	7 340	9 114	15 575	19 832	35 339	45 549
2 750	32	266	114	5 968	7 195	8 722	13 866	17 025	27 666	34 196
2 800	28	336	188	5 959	7 055	8 355	12 380	14 668	21 754	25 780
3 000	22	552	422	5 923	7 091	8 528	13 277	16 141	25 609	31 312

Identical to that of Table 8.

In Table 20, it is observed that the lowest standard error of fit is obtained with a *vu* of 2 700 m^3/s, whose predictions are the highest of all the *vu* tested. In this case, such results are not adopted because the acceptance of the Poisson-Pareto distribution is not met, as seen in Table

21. A *vu* of 2 600 m^3/s is adopted, for which the model Poisson-Pareto is still acceptable, according to the results of Table 21. The *ER* (Equation (21)) in the return periods of 25 and 10 000 years of the predictions adopted are 8.8 and 15.8 %.

Table 21. Results of the λ distribution selection test and fit parameters of the Poisson-Pareto distribution applied to the SDP of the floods of the *Pánuco* hydrometric station, Mexico.

x_0 *vu*	*E*	*V*	*d*	DE	VTM	Fit parameters		
						*u**	*a**	*k*
829.9	3.516	2.056	17.544	PO	1 402.3	2 738.089	1 252.490	0.296228
1 000	3.226	1.852	17.226	PO	1 354.2	2 734.251	1 244.537	0.288664
1 250	2.710	1.625	17.995	PO	1 340.8	2 801.020	1 252.427	0.420709
1 500	2.355	1.713	21.821	PO	1 275.4	2 818.323	1 255.689	0.460355
1 750	2.065	1.867	27.127	PO	1 189.8	2 818.849	1 244.772	0.454779
2 000	1.903	2.281	35.954	PO	1 026.3	2 727.458	1 073.663	0.158639
2 250	1.613	1.850	34.413	PO	940.1	2 706.037	949.5336	0.019501
2 400	1.419	1.856	39.238	PO	908.4	2 705.831	880.4129	-0.046399
2 500	1.258	1.804	43.027	PO	917.6	2 704.300	893.3889	-0.034004
2 550	1.226	1.852	45.331	PO	890.6	2 715.482	820.7762	-0.096273
2 600	1.226	1.852	45.331	PO	840.6	2 739.234	698.1020	-0.203154
2 650	1.194	1.898	47.707	BN	812.5	2 756.943	619.1960	-0.273937
2 700	1.194	1.898	47.707	BN	762.5	2 786.495	505.4875	-0.380161
2750	1.032	1.515	44.032	PO	829.3	2 768.109	573.2537	-0.315650
2800	0.903	1.378	45.760	PO	894.2	2 732.191	657.9195	-0.245625
3000	0.710	1.174	49.619	BN	914.6	2 791.176	578.1785	-0.299486

Identical to that of Table 9.

Since the number of years of the record (n) is 31, then $\chi^2_{0.025}$ = 16.8 and $\chi^2_{0.975}$ = 47.0, according to values in Table 1. In Table 21, the acceptance of the Poisson-Pareto model is verified from x_0 to a *vu* of 2 600 m^3/s, for which the shape parameter (k) shows a negative value of the order of magnitude of those obtained with vu greater than the one adopted.

It is observed in Table 21, in its first seven *vu* analyzed that its shape parameter (k) is positive, which coincides with the results of a pioneering study (Campos-Aranda, 1996), with the GVE distribution of shape parameter positive or Weibull model, which has an upper limit at its extreme of maximum observed levels.

Discussion of results

In Table 22, a summary of predictions for the five flood records processed has been integrated. For each SDP adopted, its number of exceedances (NE), the maximum flow (Q_{max}) and the minimum (Q_{min}) of such surpluses are cited. The PDF applied to the SAM is presented first, which led to the most severe predictions. In the first four records it was the FDP Generalized Logistic (LOG) and at the *Pánuco* station, it was the Wakeby distribution (WAK).

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

Table 22. Contrast of fit errors and predictions (m^3/s) of the SAM with the indicated distributions and of the SDP with the adopted threshold value, in the five hydrometric stations processed.

PDF *vu*	*EEA*	*EAM*	Return periods, in years						
			25	50	100	500	1 000	5 000	10 000
Guamúchil (n = 33, NE = 46; Q_{max} = 3 507 m^3/s; Q_{min} = 276 m^3/s)									
LOG	162	64	1 873	2 562	3 487	7 089	9 612	19 479	26 397
GVE	153	62	1 927	2 609	3 492	6 701	8 813	16 506	21 569
275	131	112	1 979	2 711	3 684	7 365	9 872	19 356	25 813
ER	-	-	2.6	3.8	5.2	9.0	10.7	14.7	16.4
Santa Rosa (n = 45, NE = 55; Q_{max} = 2588 m^3/s; Q_{min} = 326 m^3/s)									
LOG	93	75	1 897	2 444	3 122	5 422	6 846	11 701	14 713
LP3	64	50	1 997	2 480	3 005	4 393	5 066	6 818	7 657
325	125	90	2 070	2 620	3 264	5 214	6 298	9 584	11 411
ER	-	-	3.5	5.3	7.9	15.7	19.6	28.9	32.9
Tempoal (n = 48, NE = 51; Q_{max} = 6 120 m^3/s; Q_{min} = 1 125 m^3/s)									
LOG	398	240	5 100	6 590	8 460	14 923	18 999	33 165	42 110
WAK	267	160	5 324	6 438	7 577	10 321	11 547	14 501	15 821
1100	284	179	5 493	6 930	8 615	13 747	16 616	25 352	30 234
ER	-	-	3.1	7.1	12.0	24.9	30.5	42.8	47.7
Huites (n = 51, NE = 58; Q_{max} = 15 000 m^3/s; Q_{min} = 1 508 m^3/s)									
LOG	984	497	9 812	13 837	19 452	42 767	60 041	132 094	185 524
KAP	766	399	10 709	14 470	19 089	34 271	43 383	73 314	91 245
1500	853	463	10 188	14 059	19 600	41 728	57 536	120 678	165 761
ER	-	-	-5.1	-2.9	2.6	17.9	24.6	39.2	44.9
Pánuco (n = 31, NE = 38; Q_{max} = 7 300 m^3/s; Q_{min} = 2 631 m^3/s)									
WAK	196	129	5 387	6 475	7 804	12 130	14 706	23 098	28 095
LOG	233	141	5 389	6 239	7 202	9 992	11 488	15 853	18 199
2600	287	224	5 911	6 910	8 061	11 448	13 284	18 692	21 623
ER	-	-	8.8	9.7	10.7	12.6l	13.5	15.2	15.8

Then the PDFs with the best fit in the SAM are cited, these were: GVE, LP3, WAK, KAP and LOG. Such distributions define the lowest fitting errors in four of the records in Table 22; the exception occurs at the *Guamúchil* station, for the *EEA* of the SDP.

Thirdly, the fit adopted with the Poisson-Pareto distribution in the processed SDPs is cited. Finally, the relative errors are presented (Equation (21)), obtained between the prediction of the SAM and the SDP.

Recent operational processes for SDPs

Recent approaches to processing SDP range from the use of other probabilistic models for exceedances, case of references Bhunya *et al*. (2012), and Ashkar and Ba (2017); to new procedures such as those described below, which have been cited by Pan *et al*. (2022).

The Eastoe and Tawn (2010) approach is applicable in records that show excess dispersion, having a variance greater than the average of the exceedances per year. In such cases, the Poisson process is extended to consider over-dispersion using mixed and regression models.

Another approach, diametrically opposed, is that of Solari and Losada (2012), who propose a mixed model, with a central truncated Log-normal distribution and two Generalized Pareto for the maximum and minimum values of its upper and lower tails. The threshold values define the limits of the three regimes and are obtained with the maximum likelihood method.

Finally, Solari, Egüen, Polo and Losada (2017) develop a process for automatic estimation of the threshold value, based on the Anderson-

Darling test, which is combined with simulation techniques to quantify the uncertainty in the estimates of the threshold value and predictions with high return periods.

On the other hand, recent operational processes can also be considered, approaches that address the presence of climate change in extreme hydrological data records, which are therefore non-stationary. Among the new approaches, we can mention those of Roth, Buishand, Jongbloed, Klein-Tank and van Zanten (2012), as well as Durocher, Burn and Ashkar (2019).

Conclusions

It can be broadly observed from Table 22, that the predictions obtained with the SDP are close to the maximums generated by the Generalized Logistic distribution (LOG) with the SAM. The above is quite important, since such predictions come from very different samples, from their number of events (NE > n), but above all from their magnitudes, since the SAMs include n values greater than x_0 or minimum annual flow and the SDP only contain events greater than the adopted *vu*, that is, NE flows greater than that indicated as Q_{min}, in Table 22.

The relevance of the prediction estimates with the SDP is demonstrated in the relative errors in Table 22, for return periods greater than 100 years, which fluctuate from 2.6 to 10.7 % and from 15.8 to 47.7 % in the return period of 10 000 years.

Since it is not difficult to integrate flood records as SDP, based on the hydrometric information of maximum monthly flows available in the

BANDAS system (IMTA, 2003) and also, the application and verification of the Poisson-Pareto distribution is quite simple, it is recommended to incorporate the use of the SDP in the analysis of flood frequencies with the suggested operational process.

On the other hand, the approaches of Roth *et al*. (2012) and Durocher *et al*. (2019), seem convenient to start addressing the impacts of non-stationarity in frequency analyzes using partial duration series, of extreme hydrological data records.

References

Ashkar, F., & Ba, I. (2017). Selection between the generalized Pareto and kappa distributions in peaks-over-threshold hydrological frequency modelling. *Hydrological Sciences Journal*, 62(7), 1167-1180. DOI: 10.1080/02626667.2017.1302089

Bezak, N., Brilly, M., & Sraj, M. (2014). Comparison between the peak-over-threshold and the annual maximum method for flood frequency analysis. *Hydrological Sciences Journal*, 59(5), 959-977. DOI: 10.1080/02626667.2013.831174

Bhunya, P. K., Singh, R. D., Berndtsson, R., & Panda, S. N. (2012). Flood analysis using generalized logistic models in partial duration series. *Journal of Hydrology*, 420-421, 59-71. DOI: 10.1016/j.jhydrol.2011.11.037

Bhunya, P. K., Mishra, S. K., Ojha, C. S. P., & Berndtsson, R. (2008). Chapter 17. Flood estimation in ungauged basins. In: Singh, V. P. (ed.). *Hydrology and hydraulics* (pp. 605-646). Highlands Ranch, USA: Water Resources Publications, LLC.

Bobée, B. (1975). The Log-Pearson type 3 distribution and its application to Hydrology. *Water Resources Research*, 11(5), 681-689. DOI: 10.1029/WR011i005p00681

Bobée, B., & Ashkar, F. (1991). Theme 1.1. Statistical criteria and tests. In: *The Gamma Family and derived distributions applied in hydrology* (pp. 1-9). Littleton, USA: Water Resources Publications.

Campos-Aranda, D. F. (2023). Capítulo 1. Datos locales: aumento y aceptación En: *Temas recientes del análisis de frecuencias hidrológico* (pp. 1-51). Jiutepec, México: Instituto Mexicano de Tecnología del Agua.

Campos-Aranda, D. F. (2018). *Ponencia núm. 78. Análisis de frecuencia de datos hidrológicos extremos con persistencia o dependientes*. XXV Congreso Nacional de Hidráulica 5 al 9 de noviembre, Ciudad de México, México.

Campos-Aranda, D. F. (2000). Descripción y aplicación del modelo Poisson-Pareto en el análisis de frecuencia de crecientes con series de duración parcial. *Ingeniería. Investigación y Tecnología*, 1(5), 199-207.

Campos-Aranda, D. F. (1996). Caracterización hidrológica de crecientes en la cuenca baja del río Pánuco con base en niveles máximos anuales. *Ingeniería Hidráulica en México*, 11(2), 15-31.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7(3), 1247-1250. DOI: 10.5194/gmd-7-1247-2014

Coles, S. (2001). Chapter 4. Threshold models. In: *An introduction to statistical modeling of extreme values* (pp. 74-91). London, England: Springer-Verlag.

Cunnane, C. (1979). A note on the Poisson assumption in partial duration series models. *Water Resources Research*, 15(2), 489-494. DOI: 10.1029/WR015i002p00489

Durocher, M., Burn, D. H., & Ashkar, F. (2019). Comparison of estimation methods for a nonstationary index-flood model in flood frequency analysis using peaks over threshold. *Water Resources Research*, 55(11), 9398-9416. DOI: 10.1029/2019WR025305

Eastoe, E., & Tawn, J. A. (2010). Statistical models for over dispersion in the frequency of peaks over threshold data for a flow series. *Water Resources Research*, 46(2), 1-12. DOI: 10.1029/2009WR007757

Fischer, S., & Schumann, A. (2016). Robust flood statistics: Comparison of peak over threshold approaches based on monthly maxima and TL-moments. *Hydrological Sciences Journal*, 61(3), 457-470. DOI: 10.1080/02626667.2015.1054391

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.

IMTA, Instituto Mexicano de Tecnología del Agua. (2003). *Banco Nacional de Datos de Aguas Superficiales (BANDAS).* Hidrometría y Sedimentos hasta 2002 (8 CD). Jiutepec, México: Comisión Nacional del Agua, Secretaría de Medio Ambiente y Recursos Naturales, Instituto Mexicano de Tecnología del Agua.

Khaliq, M. N., Ouarda, T. B. M. J., Ondo, J. C., Gachon, P., & Bobée, B. (2006). Frequency analysis of a sequence of dependent and/or non-stationary hydro-meteorological observations: A review. *Journal of Hydrology*, 329(3-4), 534-552. DOI:10.1016/j.jhydrol.2006.03.004

Kite, G. W. (1977). Chapter 12. Comparison of frequency distributions. In: *Frequency and risk analyses in hydrology* (pp. 156-168). Fort Collins, USA: Water Resources Publications.

Lang, M., Ouarda, T. B. M. J., & Bobée, B. (1999). Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225(3-4), 103-117. DOI: 10.1016/S0022-1694(99)00167-5

Langbein, W. B. (1949). Annual floods and the partial-duration flood series. *Transactions of the American Geophysical Union*, 30(6), 879-881. DOI: 10.1029/TR030i006p00879

Madsen, H., Rasmussen, P. F., & Rosbjerg, D. (1997). Comparison of annual maximum series and partial duration series for modeling extreme hydrologic events. 1. At-site modeling. *Water Resources Research*, 33(4), 747-757. DOI: 10.1029/96WR03848

Madsen, H., Rosbjerg, D., & Harremoës, P. (1993). Application of the partial duration series approach in the analysis of extreme rainfalls. In: *Extreme hydrological events: Precipitation, floods and droughts* (pp. 257-266). Wallingford, England: IAHS Publication No. 213.

Metcalfe, A. V. (1997). Theme 4.5. Generalized Pareto distribution and peaks over threshold (POT) analysis. In: *Statistics in Civil Engineering* (pp. 94-98). London, England: Arnold Publishers.

Meylan, P., Favre, A. C., & Musy, A. (2012). Chapter 3. Selecting and checking data series and Theme 8.2.2. Peak over a threshold series. In: *Predictive hydrology. A frequency analysis approach* (pp. 29-70, 139-144). Boca Raton, USA: CRC Press.

Önöz, B., & Bayazit, M. (2001). Effect of the occurrence process of the peaks over threshold on the flood estimates. *Journal of hydrology*, 244(1-2), 86-96. DOI: 10.1016/S0022-1694(01)00330-4

Ostle, B., & Mensing, R. W. (1975). Appendix 4. Cumulative Chi-square distribution. In: *Statistic in research* (pp. 540-543) (3rd ed.). Ames, USA: The Iowa State University Press.

Pan, X., Rahman, A., Haddad, K., & Ouarda, T. B. M. J. (2022). Peaks-over-threshold model in flood frequency analysis: A scoping. *Stochastic Environment Research and Risk Assessment*, 36(9), 2419-2435. DOI: 10.1007/s0047/022-02174-6

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.

Roth, M., Buishand, T. A., Jongbloed, G., Klein-Tank, A. M. G., & van Zanten, J. H. (2012). A regional peaks-over-threshold model in a nonstationary climate. *Water Resources Research*, 48(11), 1-12. DOI: 10.1029/20126WR012214

Shane, R. M., & Lynn, W. R. (1964). Mathematical model for flood risk evaluation. *Journal of the Hydraulics Division*, 90(HY6), 1-20.

Singh, V. P., & Zhang, L. (2017). Frequency distributions. In: Singh, V. P. (ed.). *Handbook of applied hydrology* (pp. 21.1-21.11) (2nd ed.). New York, USA: McGraw-Hill Education.

Solari, S., Egüen, M., Polo, M. J., & Losada, M. A. (2017). Peaks over threshold (POT): A methodology for automatic threshold estimation using goodness of fit *p*-value. *Water Resources Research*, 53(4), 2833-2849. DOI: 10.1002/2016WR019416

Solari, S., & Losada, M. A. (2012). A unified statistical model for hydrological variables including the selection of threshold for the peak over threshold method, *Water Resources Research*, 48(10), 1-15. DOI: 10.1029/2011WR011475

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

Stedinger, J. R., Vogel, R. M., & Foufoula-Georgiou, E. (1993). Theme 18.6.1. Partial duration series. In: Maidment, D. R. (ed.). *Handbook of hydrology* (pp. 18.37-18.39). New York, USA: McGraw-Hill, Inc.

Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30(1), 79-82. DOI: 10.3354/cr030079

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