DOI: 10.24850/j-tyca-15-02-09

Notes

Contrast of frequency analysis between beta-kappa and beta-Pareto distributions with three of widespread application

Contraste de análisis de frecuencias entre las distribuciones beta-kappa y beta-Pareto con tres de aplicación generalizada

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

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

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

Abstract

The hydrological design of several hydraulic works or the revision of the constructed ones is based on the *design floods*, which are maximum flows of the river, associated with low probabilities of exceedance or *predictions*. Its most reliable estimate is made through *frequency analysis*, statistical

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

process that consists of representing the record of maximum annual flows, with a *probability distribution function* (PDF) or probabilistic model, used to make the desired predictions. In this contrast study, the beta-kappa and beta-Pareto FDPs are proposed, and the following three were considered to be widely used FDPs: Log-Pearson type III, general extreme values, and generalized logistics. Therefore, it is exposed, for the first two FDP, a summary of his theory and his method of fit for maximum likelihood is presented. Eleven annual extreme hydrological data records are processed and the fits are contrasted with two indices: The standard error of fit and the mean absolute error. The selection of the predictions in the seven return periods (*Tr*) studied was based on the lower values of the fit errors and on the search for *representative* predictions in the *Tr* ≥ 500 years. The conclusions suggest the inclusion of the beta-kappa and beta-Pareto distributions in the frequency analysis due to their versatility and fit facility.

Keywords: Beta-kappa distribution, beta-Pareto distribution, maximum likelihood fit, standard error of fit, mean absolute error, *Q-Q* graphics, predictions.

Resumen

El diseño hidrológico de varias obras hidráulicas o la revisión de las construidas se basa en las *crecientes de diseño*, que son gastos máximos del río, asociados con bajas probabilidades de excedencia o *predicciones*. Su estimación más confiable se realiza a través del *análisis de frecuencias*, proceso estadístico que consiste en representar el registro de gastos máximos anuales con una *función de distribución de probabilidades* (FDP)

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

o modelo probabilístico, utilizado para realizar las predicciones buscadas. En este estudio de contraste se proponen las FDP beta-kappa y beta-Pareto, y se consideraron FDP de uso generalizado las tres siguientes: la Log-Pearson tipo III, la general de valores extremos y la logística generalizada. Por lo anterior, se expone para las dos primeras FDP un resumen de su teoría y su método de ajuste por máxima verosimilitud. Se procesan 11 registros de datos hidrológicos extremos anuales y se contrastan los ajustes con dos índices: el error estándar de ajuste y el error absoluto medio. La selección de las predicciones en los siete periodos de retorno (*Tr*) estudiados se basó en los valores menores de los errores de ajuste y en la búsqueda de predicciones *representativas*, en los *Tr* ≥ 500 años. Las conclusiones sugieren la inclusión de las distribuciones beta-kappa y beta-Pareto en los análisis de frecuencias debido a su versatilidad y facilidad de ajuste.

Palabras clave: distribución beta-kappa, distribución beta-Pareto, ajuste por máxima verosimilitud, error estándar de ajuste, error absoluto medio, gráficos *Q*-*Q*, predicciones.

Received: 10/06/2021

Accepted: 26/07/2022

Published Online: 15/08/2022

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Introduction

Stages of frequency analysis

The hydrological dimensioning of hydraulic works, such as: protection dykes, canalizations and bridges, as well as the various urban drainage structures, is based on the *Design Floods* (CD, by its acronym in Spanish). The most accurate hydrological estimation of CDs is done through *Frequency Analysis* (AF, by its acronym in Spanish); a statistical procedure which consists on interpreting or characterizing the available record of hydrological events, for example, floods or maximum rainfall, in terms of their future probabilities of occurrence (Bobée & Ashkar, 1991).

AF involves the following five stages: (1) integration and verification of the randomness of the record or available sample; (2) selection of the *probability distribution function* (PDF) or probabilistic model that will represent the data and allow estimates or *predictions* associated with low probabilities of exceedance; (3) adjustment of the various FDPs tested, that is, obtaining their fit parameters with the various available methods; (4) evaluation of the statistical quality of the fit achieved between the data and the PDF, by means of graphs and diagnostic indices and (5) selection of the results (Kite, 1977; Bobée & Ashkar, 1991; Rao & Hamed, 2000 ; Meylan, Favre, & Musy, 2012; Stedinger, 2017; Teegavarapu, Salas, & Stedinger, 2019).

In this contrast study, in stage one, four records of peak flow and *joint* volume of annual floods and three records of annual maximum daily rainfall were selected. In total, eleven series of extreme hydrological data

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

were processed and their randomness was verified with the Wald-Wolfowitz test. In stage two, the objective of the study is addressed by selecting the beta-kappa (BEK) and beta-Pareto (BEP) FDPs to contrast them against three of general application, which were: the Log-Pearson type III (LP3), the General of Extreme Values (GVE) and the Generalized Logistics (LOG). All the cited PDFs have three fit parameters.

In stage three, the BEK and BEP PDFs were fitted using the maximum likelihood method proposed by Mielke and Johnson (1974). The LP3 distribution was fitted with its classical method of moments in the logarithmic domain (WRC, 1977) and the GVE and LOG models with the method of L moments (Hosking & Wallis, 1997).

For stage four, two indices (*EEA* and *EAM*) were calculated, the standard error of fit (Kite, 1977; Chai & Draxler, 2014) and the mean absolute error (Willmott & Matsuura, 2005). Finally, for the fhfth stage of result selection, the *EEA* and *EAM* values were taken into account, as well as the values obtained for the *predictions*.

Background on BEK and BEP distributions

Strupczewski, Markiewicz, Kochanek and Singh (2008) indicate that there are very few references on the use of PDF, with two shape parameters, to model extreme hydrological events and cite the following two. The French statistician Halphen's system of distributions proposed in 1941 has a lower bound of zero and two shape parameters, but due to its mathematical complexity it was abandoned. On the other hand, the

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

distributions designated by Mielke and Johnson (1974) such as BEK and BEP are models with two shape parameters and one scale parameter.

Wilks (1993), in a pioneering study of the PDF contrast of three fit parameters, with maximum precipitation data, processed as annual series and with magnitudes greater than a threshold value, found that the BEK distribution describes the annual series fairly well and the BEP model is a better fit for partial duration series.

Campos-Aranda (1998) exposed various applications of the BEP distribution. Mason, Waylen, Mimmack, Rajaratnam and Harrison (1999) use the BEK and BEP PDFs in a change detection study for extreme rainfall events.

Öztekin (2007) contrasts the BEK and BEP models against the Wakeby distribution, finding that the latter leads to better or similar fits in maximum rainfall records. Murshed, Kim and Park (2011) expose for the BEK distribution the estimation of its fit parameters by means of the methods of moments and moments L. Finally, Nguyen, El Outayek, Lim and Nguyen (2017) include the PDF BEK and BEP in their study of maximum annual rainfall contrast, based on the descriptive and predictive abilities of the PDF.

Objectives

From this contrasting study the objectives were three: (1) present a summary of the BEK and BEP distributions theory; (2) describe in detail its maximum likelihood fit method, through its iterative equations for the

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

estimation of its three fit parameters and (3) perform a goodness-of-fit and prediction contrast between the BEK and BEP distributions and the three general application ones (LP3, GVE and LOG).

Generally speaking, the AF statistical technique has debatable assumptions at each of its five stages. From the representativeness in the future of the available registry, to the adoption of results, based on graphics and diagnostic indexes; going through the selection of several PDFs to make the desired predictions. It is stage two of AF, which opens possibilities to test new probabilistic models, since as is known, no PDF is better and its *suitability* depends on each record processed.

Methods and materials

Equations of the BEK and BEP distributions

Mielke and Johnson (1974) expose the PDF and the probability density function (*pdf*) of the generalized beta random variable of the second kind (x):

$$F(x) = \frac{(x/\beta)^{\gamma}\,{}_{2}^{1}F\left[\alpha+1, \gamma/\theta; 1+\gamma/\theta; -(x/\beta)^{\theta}\right]}{\left(\frac{\gamma}{\theta}\right)B} \quad x \geq 0 \tag{1}$$

$$F(x) = 0 \quad x \geq 0 \tag{2}$$

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$f(x) = \frac{(x/\beta)^{\gamma-1}\left[1+(x/\beta)^{\theta}\right]^{-(\alpha+1)}}{(\beta/\theta)\,\mathrm{B}(\gamma/\theta,\alpha+1\text{-}\gamma/\theta)} \quad x > 0 \tag{3}$$

$$f(x) = 0 \quad x \geq 0 \tag{4}$$

in which, $\alpha > 0$, $\beta > 0$, $\theta > 0$ and $0 < \gamma < \theta\ (\alpha + 1)$. The numerator of Equation (1) is the following Gaussian hypergeometric series (Oberhettinger, 1972):

$${}_2^1F(a,b;c;z) = \sum_{n=0}^{\infty} \frac{\Gamma(n+a)\Gamma(n+b)\Gamma(c)}{\Gamma(a)\Gamma(b)\Gamma(n+c)} \frac{z^n}{n!} \tag{5}$$

and the denominator function is:

$$B(\delta,\epsilon) = \Gamma(\delta)\Gamma\,(\epsilon)/\Gamma\,(\delta+\epsilon) \tag{6}$$

Mielke and Johnson (1974) indicate that the calculations associated with equations (1) and (3) are not simple, but by establishing two restrictions on the parameter γ, distributions with important computational advantages are obtained. The first constraint is $\gamma = \alpha\theta$ and leads to the beta-κ (BEK) distribution, named this way due to its similarity to Mielke's (Mielke, 1973) Kappa distribution, whose PDF and *pdf* equations are:

$$F(x) = \left\{(x/\beta)^{\theta}/\left[1+(x/\beta)^{\theta}\right]\right\}^{\alpha} \quad x \geq 0 \tag{7}$$

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$f(x) = (\alpha\theta/\beta)(x/\beta)^{\alpha\theta-1}\left[1+(x/\beta)^{\theta}\right]^{-(\alpha+1)} \quad x > 0 \tag{8}$$

with $\alpha > 0$, $\beta > 0$ and $\theta > 0$. β is the scale parameter and α y θ the shape parameters. The quantile function, designating $F(x) = p$, is as follows:

$$x(p) = \beta\left[p^{1/\alpha}/\left(1-p^{1/\alpha}\right)\right]^{1/\theta} \tag{9}$$

The second restriction is $\gamma = \theta$, which defines the beta-*P* (BEP) distribution, thus designated for its resemblance to the Pareto-type probabilistic model, its PDF and *pdf* are:

$$F(x) = 1-\left[1+(x/\beta)^{\theta}\right]^{-\alpha} \quad x \geq 0 \tag{10}$$

$$f(x) = (\alpha\theta/\beta)(x/\beta)^{\theta-1}\left[1+(x/\beta)^{\theta}\right]^{-(\alpha+1)} \quad x > 0 \tag{11}$$

with $\alpha > 0$, $\beta > 0$ and $\theta > 0$. Again, β is the scale parameter and α and θ are the shape parameters. The quantile function, designating $F(x) = p$, is as follows:

$$x(p) = \beta\left[(1-p)^{-1/\alpha}-1\right]^{1/\theta} \tag{12}$$

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

BEK and BEP by maximum likelihood fit

Mielke and Johnson (1974) describe the procedures for calculating the fit parameters (α, β, θ) with the maximum likelihood method, according to three equations of iterative application (j), starting from initial values of β_0 and θ_0. For the BEK distribution such expressions are:

$$\alpha_j = n\left\{\sum_{i=1}^{n} \ln\left[1 + \left(x_i/\beta_{j-1}\right)^{-\theta_{j-1}}\right]\right\}^{-1} \tag{13}$$

$$\beta_j = \frac{1}{n}\left(1 + \frac{1}{\alpha_j}\right)\beta_{j-1}\sum_{i=1}^{n}\left[1 + (x_i/\beta_{j-1})^{-\theta_{j-1}}\right]^{-1} \tag{14}$$

$$\theta_j = n\left\{\sum_{i=1}^{n} \frac{\left[\left(x_i/\beta_j\right)^{\theta_{j-1}} - \alpha_j\right]\ln\left(x_i/\beta_j\right)}{1+\left(x_i/\beta_j\right)^{\theta_{j-1}}}\right\}^{-1} \tag{15}$$

For the BEP distribution its iterative equations are:

$$\alpha_j = n\left\{\sum_{i=1}^{n} \ln\left[1 + \left(x_i/\beta_{j-1}\right)^{\theta_{j-1}}\right]\right\}^{-1} \tag{16}$$

$$\beta_j = \frac{1}{n}\left(1 + \alpha_j\right)\beta_{j-1}\sum_{i=1}^{n}\left[1 + (x_i/\beta_{j-1})^{-\theta_{j-1}}\right]^{-1} \tag{17}$$

$$\theta_j = n\left\{\sum_{i=1}^{n} \frac{\left[\alpha_j\left(x_i/\beta_j\right)^{\theta_{j-1}} - 1\right]\ln\left(x_i/\beta_j\right)}{1+\left(x_i/\beta_j\right)^{\theta_{j-1}}}\right\}^{-1} \tag{18}$$

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Contrast distributions and their fitting

The LP3 distribution was fitted with the classic method of moments in the logarithmic domain (WRC, 1977). In contrast, the GVE and LOG distributions were fitted based on the L moments method, which has been shown to be efficient and robust, even in small samples, according to equations set forth by Hosking and Wallis (1997).

Campos-Aranda (2002) presented six methods of fitting the LP3 distribution, limiting their applicability based on the estimation ratio (*CE*), defined as:

$$CE = \frac{X_{100} - X_{50}}{X_{50} - X_{25}} \quad (19)$$

In which, X_{Tr} is the *prediction* of the return period (*Tr*) expressed in years, equal to the reciprocal of the exceedance probability (q), when series or samples of maximum annual hydrological events are processed. When the $CE \leq 1.00$, the best fitting methods are the mixture of moments and the moments in the real domain, and when $CE \geq 1.30$, the most convenient methods are the maximum likelihood and maximum entropy methods. When *CE* fluctuates between the quoted limits, the method of moments in the logarithmic domain is acceptable and generally leads to a good fit.

On the other hand, the General distribution of Extreme Values (GVE), is derived from the Fisher-Tippet-Gnedenko Theorem, which establishes

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

as the probabilistic model of the maximum values sampled annually the GVE and its three particular cases, depending on the value of its shape parameter. However, as indicated at the end of the Objectives section, other distributions continue to be proposed and used, due to the random nature of the extreme hydrological data records, to search for their *ideal* distribution, based on the statistical indicators of the fit achieved.

Q-Q diagnostic graph

Nguyen *et al*. (2017) have suggested two evaluations to select the optimal PDF to obtain a record of extreme hydrological data: (1) descriptive ability and (2) predictive ability. The first refers to the accuracy with which the PDF being tested reproduces the sample data and the second is logically associated with the variability of its predictions in relation to the dispersion of the sample predictions. There are three techniques to test descriptive ability: (1) diagnostic charts; (2) statistical tests and (3) goodness-of-fit indices (Meylan *et al.*, 2012).

Empirical versus estimated *probability* and *amount* observed versus estimated, *P-P* and *Q-Q* diagnostic plots have become popular (Coles, 2001; Wilks, 2011) and provide a simple and effective way to compare the results of a contrasted PDF. For a sample of data x_i sorted from smallest to largest, an empirical probability (p) is assigned to them, for example, with the Cunnane formula, which according to Stedinger (2017) leads to unbiased values in most of the PDFs used in hydrology, this is:

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$p = \frac{m-0.40}{n+0.20} \tag{20}$$

in which *m* is the order number of the data and *n* its total number. For each datum x_i, its probability is obtained with the equation of the tested PDF. For the case of the BEK and BEP distributions, with expressions (7) and (10). The *P-P* graph is defined with the following abscissa and ordinate points:

$$\left[\frac{m-0.40}{n+0.20}, F(x_i)\right] \quad \text{for } i = 1, 2, \ldots., n \tag{21}$$

The *Q-Q* graph uses equations (9) and (12) or inverse solutions of the PDFs BEK and BEP, to define the points of the ordinates and is made up of the following points:

$$\left[x_i, x\left(\frac{m-0.40}{n+0.20}\right)\right] \quad \text{for } i = 1, 2, \ldots., n \tag{22}$$

The disadvantage of diagnostic graphs lies in the subjective assessment that is made when comparing various PDFs, since a numerical value is not available (Nguyen *et al.*, 2017). Campos-Aranda (2019) visualizes the *Q-Q* graph as more useful, to detect overestimated predictions (because it is above the 45° line) or underestimated (because it is below).

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

Standard Error of Fit (*EEA*)

Goodness-of-fit indices have the advantage of being easy to calculate and commonly involve the difference between the observed values x_i and the estimated values $\hat{x}$ with the PDF being tested. The *EEA* is the most common (Chai & Draxler, 2014), it was established in the mid-1970s (Kite, 1977) and has been applied in Mexico using Weibull's empirical formula (Benson, 1962). It will now be applied using Cunnane Equation (20). The expression of the *EEA* is:

$$EEA = \left[\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{(n - np)}\right]^{1/2} \tag{23}$$

x_i are the n observed data ordered from smallest to largest, $\hat{x}_i$ the estimates, for the probability estimated with Equation (20) and the PDF that is contrasted; *np* is the number of fit parameters of the FDP, with three for those applied in this study.

Mean Absolute Error (*EAM*)

Its advantages lie in having the units of the variable, just like the *EEA*, and preventing the impact of the scattered values from being squared and therefore *EEA*≥*EAM* (Willmott & Matsuura, 2005). Its expression is (Nguyen *et al.*, 2017):

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$EAM = \frac{\sum_{i=1}^{n}|x_i - \hat{x}_i|}{n - np} \quad (24)$$

Processed hydrological data records

Aldama, Ramírez, Aparicio, Mejía-Zermeño and Ortega-Gil (2006) indicate the *joint* records of peak flow rate (*Qp*) in m^3/s and volume (*Vol*) in millions of m^3 (Mm^3) per year of the inflow floods at 15 important reservoirs in Mexico and one in project. Of such joint records, the three that are considered *complicated* in their probabilistic analysis were selected, since they include scattered values (*outliers*) and have large differences between their low and maximum values. The first ones correspond to the 43 entry data to the El Infiernillo dam on the Balsas River, between the states of Michoacán and Guerrero, which has a basin area of 108,000 km^2. The second records to be processed are the 52 intake data for the Huites dam on the Fuerte River, in the state of Sinaloa, with a basin area of 26,020 km^2. The third ones were the 37 entrance data to the Guamúchil dam, on the Mocorito river, also in the state of Sinaloa and with a basin area of 1,630 km^2.

In addition, Domínguez and Arganis (2012) expose the joint records of *Qp* and *Vol* with 47 input data to the Malpaso dam, on the Grijalva River, in the state of Chiapas, Mexico, with a basin area of 34,800 km^2. Such flood records were also processed.

Finally, three records of annual maximum daily precipitation (*PMD*) from a pluviometric station in each geographical area of the state of San Luis Potosi, Mexico, were analyzed. From the Altiplano, the Los Filtros

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

station (n = 66), located in the city of San Luis Potosi, was processed; from the Middle Zone, the one located in the city of Río Verde (n = 52) and from the Huasteca region, the one from the town of Tanquian de Escobedo (n = 52). The altitudes of the pluviometric stations cited are: 1904, 987 and 87 meters above sea level. These records were analyzed by Campos-Aranda (2019) to obtain their optimal PDFs and integrated based on the CONAGUA monthly Excel file, provided to the author; therefore, they are reproduced in Table 1.

Table 1. Records of annual *PMD* (millimeters) in the three pluviometric stations indicated in the state of San Luis Potosí, Mexico.

No.	Los Filtros (1949-2014)		Rio Verde (1961-2013*)		Tanquian (1961-2014**)		No.
1	15.9	66.5	46.4	28.2	107.0	120.0	34
2	20.6	26.0	52.2	61.8	55.5	62.0	35
3	50.9	31.5	33.4	37.8	90.3	64.0	36
4	40.5	46.5	27.0	55.6	171.0	132.0	37
5	63.6	44.0	39.2	39.2	81.5	80.0	38
6	41.9	41.0	79.0	44.0	176.0	72.0	39
7	60.0	55.0	43.1	74.1	176.5	82.0	40
8	35.9	21.5	40.5	51.7	71.5	88.0	41
9	48.6	29.8	57.7	34.0	78.0	185.0	42
10	63.0	41.5	63.7	41.5	109.0	105.0	43

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

No.	Los Filtros (1949-2014)		Rio Verde (1961-2013*)		Tanquian (1961-2014**)		No.
11	35.5	25.4	81.5	43.5	84.0	67.0	44
12	40.0	59.0	52.5	32.3	87.0	201.0	45
13	63.2	33.5	48.5	126.3	111.0	89.0	46
14	39.4	46.5	51.3	58.5	99.5	200.9	47
15	27.2	51.0	117.5	53.7	166.5	90.0	48
16	59.0	40.0	57.3	99.1	113.5	110.0	49
17	32.0	35.5	61.8	81.4	148.0	68.0	50
18	30.0	45.5	83.4	31.8	117.0	89.0	51
19	40.2	25.9	71.7	63.2	85.0	190.0	52
20	31.5	20.7	35.0	-	73.0	-	53
21	31.5	37.5	87.0	-	103.0	-	54
22	52.0	40.2	86.3	-	79.0	-	55
23	52.3	111.0	37.1	-	81.5	-	56
24	31.3	43.3	30.2	-	113.5	-	57
25	35.0	76.9	87.1	-	94.0	-	58
26	28.5	42.8	46.7	-	54.0	-	59
27	57.2	46.1	79.9	-	86.0	-	60
28	58.0	42.5	51.0	-	98.0	-	61
29	42.9	45.3	61.0	-	63.0	-	62

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

No.	Los Filtros (1949-2014)		Rio Verde (1961-2013*)		Tanquian (1961-2014**)		No.
30	26.4	44.5	92.3	-	200.8	-	63
31	65.5	26.0	97.4	-	106.0	-	64
32	22.0	59.1	38.7	-	370.0	-	65
33	51.2	44.1	43.9	-	84.0	-	66

* one year missing.

** two years missing.

Wald-Wolfowitz test

This nonparametric test has been used by Bobée and Ashkar (1991), Rao and Hamed (2000), and Meylan *et al.* (2012) to test *independence* and *stationarity* in records of maximum annual flows (X_i). For this reason, it was proposed to apply the test to the annual *Qp*, *Vol* and *PMD* records, which must be samples of *random* values. A. Wald and J. Wolfowitz based on the work of R. L. Anderson on the serial correlation coefficient developed such test, whose statistic is:

$$R = \sum_{i=1}^{n-1} x_i \cdot x_{i+1} + x_n \cdot x_1 \qquad (25)$$

When the size (n) of the series or sample (x_i) is not small and its data are independent, R comes from a Normal distribution with mean and variance, given by the following expressions:

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

$$E[R] = \overline{R} = \frac{S_1^2 - S_2}{n-1} \quad (26)$$

$$Var[R] = \frac{S_2^2 - S_4}{n-1} + \frac{S_1^4 - 4 \cdot S_1^2 \cdot S_2 + 4 \cdot S_1 \cdot S_3 + S_2^2 - 2 \cdot S_4}{(n-1)(n-2)} - \overline{R}^2 \quad (27)$$

in which:

$$S_w = \sum_{i=1}^{n} x_i^w \quad (28)$$

Finally, U is calculated, with the equation:

$$U = \frac{R - \overline{R}}{\sqrt{Var[R]}} \quad (29)$$

The value of U follows a Normal distribution (0.1) and can be used to test the independence of the series data with a level of significance a, commonly 5 %. In a two-tailed test, the standardized normal variable is $Z_{a/2} \cong 1.96$; then, when the absolute value of U is less than 1.96, the series will be made up of independent values (*random sample*).

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

Results and their discussion

Test of randomness and ratios of L moments

In the third column of Table 2, the values of the *U* statistic (Equation (29)) are shown, defining that the 11 records processed are random. The rest of the columns show the magnitudes of the arithmetic mean and the quotients of moments L of each record.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Table 2. General data and ratios of L moments of the 11 records processed.

No.	**Record in:**	U	$\overline{X}$	t_3	t_4
1	El Infiernillo *Qp*	-0.707	5499.512	0.49875	0.36082
2	El Infiernillo *Vol*	-0.022	2244.791	0.37866	0.16834
3	Huites *Qp*	-0.090	3305.135	0.49313	0.30438
4	Huites *Vol*	-0.602	841.769	0.30773	0.17585
5	Guamuchil *Qp*	-1.418	1610.854	0.70597	0.51098
6	Guamuchil *Vol*	1.043	38.747	0.57062	0.28975
7	Malpaso *Qp*	-0.666	2153.234	0.40782	0.25020
8	Malpaso *Vol*	0.558	1583.168	0.29264	0.13777
9	Los Filtros *PMD*	-0.616	43.005	0.13516	0.16115
10	Rio Verde *PMD*	0.179	58.442	0.20926	0.09728
11	Tanquian *PMD*	-0.746	112.086	0.36652	0.22984

U =statistical of the Wald-Wolfowitz Test.

$\overline{X}$ = arithmetic mean, in m^3/s, Mm^3 or millimeters.

t_3 = quotient of moments L of asymmetry.

t_4 = quotient of moments L of kurtosis.

The original records of the Guamúchil hydrometric station exposed by Aldama *et al*. (2006) finally include four extreme maximum *Qp* and *Vol* values, obtained at the Eustaquio Buelna dam with inverse operation of its flood transit, which originate that such records are not random (U =

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

4.156 and U = 2.597). Eliminating these four *Qp* and *Vol* data, the U values shown in Table 2 are obtained.

Fitting of the BEK distribution

Taking into account that the fit parameters of the beta distributions are obtained by successive substitutions (j) or iterations, the models considered to be in general use were fitted first; that is, the LP3, GVE and LOG distributions to each of the 11 records processed, to give them that meaning.

With the above, we have the fit errors (*EEA* and *EAM*) and the predictions related to the seven return periods analyzed (25, 50, 100, 500, 1000, 5000 and 10000 years), obtained with the three mentioned distributions. The fit error values were used as magnitudes *not to be exceeded* and, in the case of predictions, in return periods of less than 100 years, *as values to match*, with the fit by iterations of the beta distributions.

The fit parameters of the BEK distribution were estimated based on equations (13) to (15), using as initial values of β_0 and θ_0 the arithmetic mean (Table 2) and a value of 5.0, respectively. In each iteration (j), the standard errors of fit (Equation (23)) and absolute mean (Equation (24)) were evaluated, using formulas (9) and (20), to obtain the estimated values $\hat{x}_i$. The comparison between errors obtained with the BEK and its predictions allowed defining the number of iterations carried out, which

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

are shown in Table 3, as well as the fit parameters obtained. The maximum number of iterations was limited to 500.

Table 3. Number of iterations (j) and values of the fit parameters (β,a,θ) of the BEK distribution in the 11 records processed.

No.	Record in:	j	β	a	θ
1	El Infiernillo *Qp*	274	2619.301	2.576858	2.538683
2	El Infiernillo *Vol*	3	2463.009	0.517784	2.692160
3	Huites *Qp*	27	1488.418	1.888118	2.045599
4	Huites *Vol*	3	923.356	0.553249	3.152180
5	Guamuchil *Qp*	37	304.088	2.013669	1.384388
6	Guamuchil *Vol*	4	43.366	0.397784	1.791456
7	Malpaso *Qp*	500	701.759	8.699795	2.733231
8	Malpaso *Vol*	2	1936.030	0.461082	3.178971
9	Los Filtros *PMD*	1	42.399	0.862399	5.390516
10	Rio Verde *PMD*	1	57.438	0.800002	4.660408
11	Tanquian *PMD*	200	50.479	7.188239	3.506981

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Fitting of the BEP distribution

An identical procedure to the one previous one was followed for the BEP fit, but now equations (16) to (18) were used for its fit parameters and expressions (12) and (20) for the estimated values $\hat{x}_i$ and thus be able to evaluate the fitting errors (equations (23) and (24)) and their predictions (Equation (12)). The results are shown in Table 4.

Table 4. Number of iterations (j) and values of the fit parameters (β,a,θ) of the BEP distribution in the 11 records processed.

No.	Record in:	j	β	a	θ
1	El Infiernillo *Qp*	70	2919.052	0.360741	5.423902
2	El Infiernillo *Vol*	2	1572.835	1.001999	2.413874
3	Huites *Qp*	3	1920.472	0.762506	2.641226
4	Huites *Vol*	216	803.068	1.317579	2.335993
5	Guamuchil *Qp*	6	430.447	0.676817	1.973478
6	Guamuchil *Vol*	4	15.870	0.983551	1.417081
7	Malpaso *Qp*	9	1408.943	0.513508	5.187900
8	Malpaso *Vol*	150	1972.335	1.914897	1.882710
9	Los Filtros *PMD*	1	42.774	1.178656	4.616619
10	Rio Verde *PMD*	2	55.730	1.081484	4.527451
11	Tanquian *PMD*	16	76.011	0.345564	8.324678

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

Selection strategy and results

The general approach for the selection of the adopted predictions, in each processed record, was based on the minimum values of the *EEA* and *EAM* errors, as well as on the magnitudes of the extreme return period predictions (*Tr*) 1000, 5000 and 10000 years, to find *representative* estimates of the five obtained in each *Tr* (Table 5).

Table 5. Contrast of goodness-of-fit indicators and predictions between the BEK and BEP distributions and the three in general use (LP3, GVE and LOG), in the 11 hydrological data records processed.

NR	PDF	*EEA*	*EAM*	Return periods, in years						
				25	50	100	500	1 000	5 000	10 000
1	BEK	1575	622.3	13364	17659	23267	43955	57770	108934	143049
1	BEP	1072	436.1	15125	21556	30719	69927	99654	226845	323281
1	LP3	1078	458.2	15257	20645	27643	53011	69619	129550	168642
1	GVE	1727	652.3	14499	19961	27423	57155	78386	163182	223718
1	LOG	1829	707.3	14095	19509	27104	58949	82743	182964	257909
2	BEK	416.0	276.8	6235	8158	10612	19379	25084	45626	59015
2	BEP	552.8	352.7	5851	7861	10514	20521	27343	53207	70868
2	LP3	468.2	233.3	7055	9394	12240	21356	26650	43240	52705
2	GVE	407.2	257.2	6542	8552	11017	19156	24074	40360	50200
2	LOG	460.5	299.0	6373	8456	11138	20835	27199	50330	65532
3	BEK	990.8	482.7	9648	13644	19219	42339	59437	130555	183116
3	BEP	1023.0	505.6	9443	13367	18884	42024	59292	131850	186018
3	LP3	938.4	411.3	10856	15381	21438	44405	59970	117987	156804

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Tecnología y Ciencias del Agua

OPEN ACCESS

Check for updates

NR	PDF	*EEA*	*EAM*	Return periods, in years						
				25	50	100	500	1 000	5 000	10 000
3	GVE	974.8	453.0	9994	14008	19464	41019	56297	116801	159676
3	LOG	1056.0	492.8	9696	13680	19246	42424	59637	131640	185161
4	BEK	88.3	61.1	2086	2624	3284	5491	6845	11409	14211
4	BEP	94.3	59.8	2198	2799	3539	6025	7559	12772	16003
4	LP3	64.5	45.4	2109	2501	2899	3854	4278	5287	5732
4	GVE	86.0	58.1	2155	2672	3265	5005	5949	8732	10242
4	LOG	101.3	70.8	2111	2670	3354	5622	6999	11593	14387
5	BEK	1552	620.8	5044	8416	13961	44842	74022	236881	390495
5	BEP	1639	640.5	4771	8039	13522	45140	75850	253085	425248
5	LP3	1152	493.1	7779	14739	27483	112419	204067	803624	1444814
5	GVE	1693	716.1	5864	9679	15832	48841	79098	241647	390542
5	LOG	1820	729.2	5422	8962	14714	46057	75171	234269	382083
6	BEK	18.1	11.5	150	226	336	831	1224	3009	4429
6	BEP	23.3	12.3	156	259	429	1369	2253	7152	11761
6	LP3	16.6	11.2	199	346	584	1833	2933	8432	13126
6	GVE	22.3	13.4	148	224	334	825	1208	2913	4246
6	LOG	23.7	13.9	142	217	326	829	1235	3103	4611
7	BEK	213.8	106.4	4986	6453	8332	15039	19381	34957	45005
7	BEP	285.4	152.3	4715	6118	7936	14521	18837	34465	44707
7	LP3	185.3	93.6	5127	6550	8294	14040	17498	28887	35735
7	GVE	315.6	139.0	5013	6442	8243	14475	18403	32031	40620
7	LOG	346.7	160.2	4896	6357	8279	15478	20356	38726	51186
8	BEK	279.9	163.8	4093	5143	6428	10708	13324	22114	27498
8	BEP	280.2	151.4	4317	5423	6728	10826	13206	20810	25270
8	LP3	277.7	141.2	4482	5573	6767	9968	11545	15709	17733
8	GVE	226.4	140.3	4253	5265	6406	9676	11407	16383	19019
8	LOG	262.4	165.7	4168	5272	6610	10965	13567	22086	27182

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

NR	PDF	*EEA*	*EAM*	Return periods, in years						
				25	50	100	500	1 000	5 000	10 000
9	BEK	2.5	1.6	74	85	97	131	149	200	228
9	BEP	2.5	1.5	76	87	99	134	152	205	232
9	LP3	2.9	1.7	74	82	90	107	114	131	139
9	GVE	3.7	1.5	74	81	88	104	110	123	129
9	LOG	3.3	1.5	74	83	93	121	135	172	191
10	BEK	4.4	2.9	108	126	147	208	241	341	395
10	BEP	4.1	3.2	106	123	142	198	228	317	366
10	LP3	3.4	2.2	110	126	142	184	204	255	279
10	GVE	3.0	2.2	110	125	141	181	199	244	265
10	LOG	3.8	2.9	109	127	147	208	241	339	392
11	BEK	13.0	7.2	220	269	329	521	635	1005	1224
11	BEP	12.2	6.4	233	296	377	659	839	1468	1868
11	LP3	11.9	6.7	231	283	344	535	644	985	1180
11	GVE	15.6	7.6	228	280	344	551	673	1071	1307
11	LOG	16.3	7.9	223	278	348	598	759	1337	1712

NR = registration number, according to Tables 3 or 4.

PDF = probability distribution function tested.

EEA = standard error of fit, in m^3/s, Mm^3 ó mm, according to data.

EAM = mean absolute error, in m^3/s, Mm^3 ó mm, according to data.

Exclusively for *record 3* (Huites *Qp*), the LP3 distribution led to the lowest fit errors and also to the predictions adopted, given the extraordinary similarity that all the estimates showed.

It is important to highlight that the LP3 distribution led to the lowest fit errors in *records 4, 5, 6 and 7*. In the first and last, such distribution was not selected because it led to very low predictions; on the contrary,

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

in registers 5 and 6. For these four registers, their *CE*s were: 1,015, 1,831, 1,619 and 1,226; evidently, in registers 5 and 6 the method of fit by moments in the logarithmic domain is not applicable. For the rest of the processed records, the *CE* was not less than 1.00.

In record 1, the BEP distribution reported the lowest fitting errors, but all its predictions are considered high. Due to the above, the GVE distribution was adopted, as it has much lower fit error values than the LOG model. *In record 2*, the LP3 and GVE distributions led to the lowest fitting errors; the former became adopted due to its more severe predictions.

In record 4, the GVE distribution provided the following minimum fitting error values, but its predictions were reduced and, therefore, the BEK distribution was adopted with low errors and representative predictions, even similar to those of the LOG.

In records 5, 6 and 7, the adopted BEK distribution reported the second lowest fitting error values and its predictions are accepted as representative, due to the great similarity they show with those of the GVE and LOG models.

In records 8, 9 and 10, the GVE, BEP and GVE distributions were adopted, which reported the minimum fitting errors. Regarding its predictions, those of the GVE can be considered slightly scarce and those of the BEP somewhat high, when compared to those of the LOG model.

In record 11, the LP3 and BEP distributions led to the smallest fitting errors; the latter is adopted because of its more severe predictions.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Figure 1 and Figure 2 show the *Q-Q* diagnostic graphs of the two best fits achieved with the BEK distributions in register 7 and BEP in register 9.

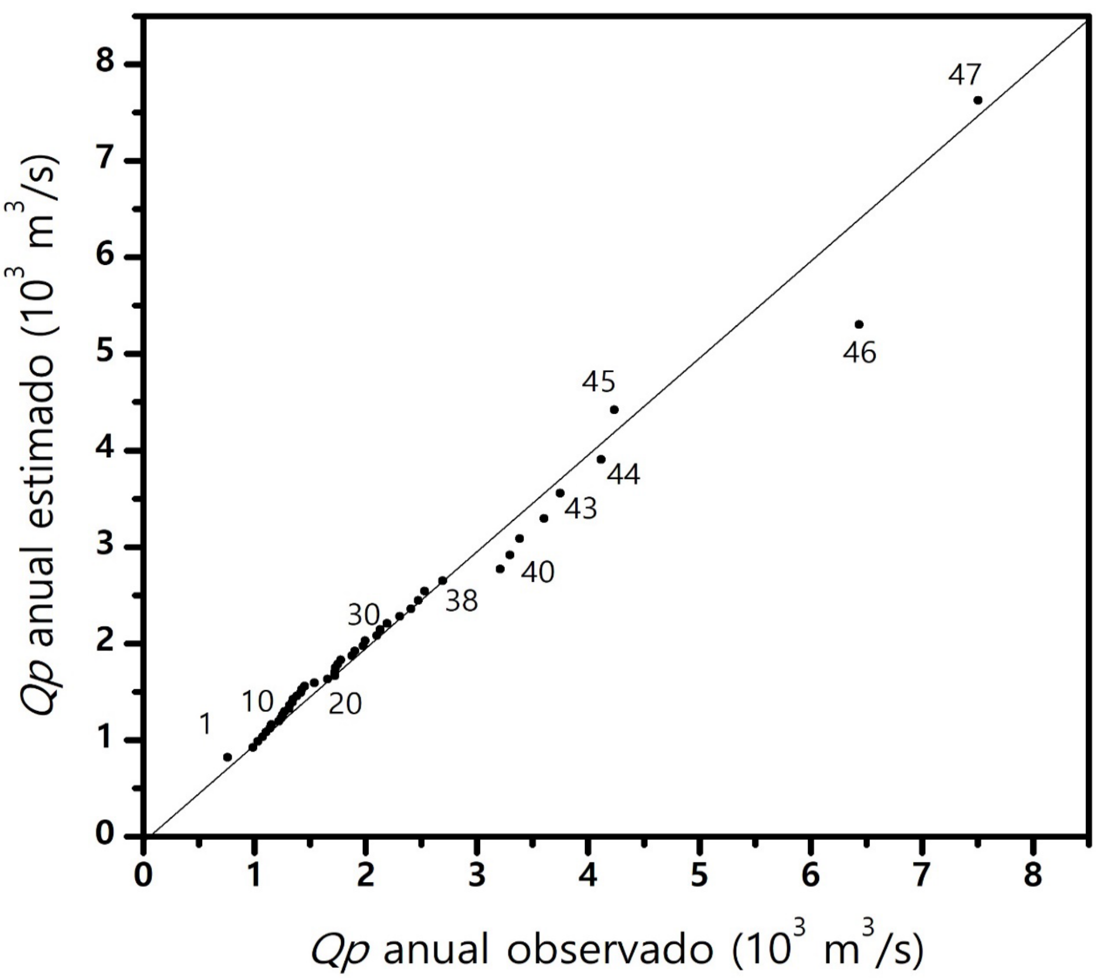

Figure 1. *Q-Q* graph of record 7 of annual *Qp* in the Malpaso dam obtained with the BEK distribution.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

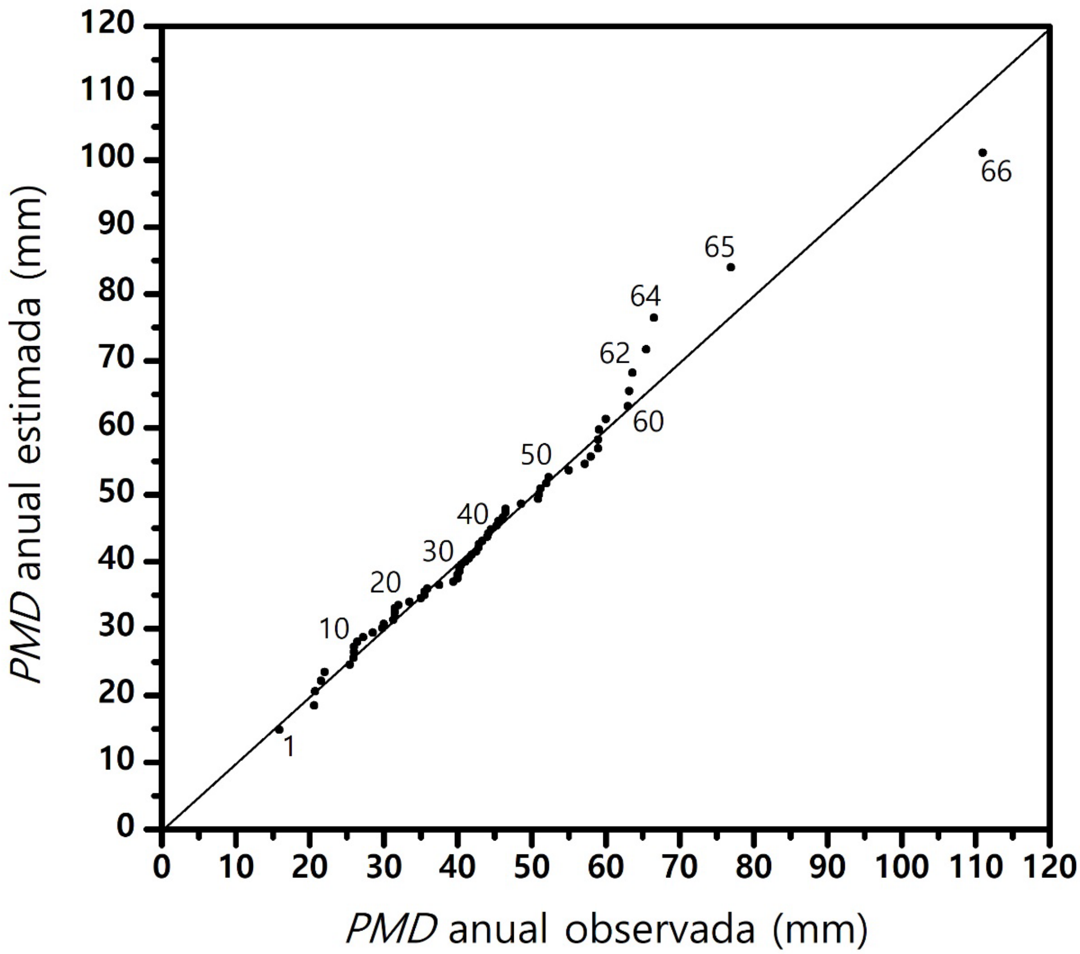

Figure 2. *Q-Q* graph of record 9 of annual *PMD* at Los Filtros station obtained with the BEP distribution.

In Figure 1, up to data number 38, the BEK model reproduces the sample data exactly and from there, it underestimates the following six values and data 46; finally, it slightly overestimates the 45th and last. In contrast, in Figure 2, only a lack of accuracy is detected in the last six data, but not severe.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

OPEN ACCESS
Check for updates

Conclusions

The contrasted BEK and BEP distributions have the following two advantages: (1) they show great versatility to represent records of extreme hydrological events due to their dense right tail and the use of two shape parameters and (2) their maximum likelihood fitting method it is efficient, simple and computationally uncomplicated. Due to the above mentioned reasons, its routine inclusion in the frequency of extreme hydrological events analysis is recommended.

In the eleven records processed, the predictions of the first three return periods (*Tr* < 100 years), are quite similar. In general, the PDF adopted for having a lower *EEA* and *EAM* value, leads to representative predictions in the last four high return periods (*Tr* ≥ 500 years). The two previous conditions give rise to confidence in all the calculated and adopted *predictions*.

The observations deduced from Table 5 of results (errors and predictions), allow us to suggest the routinely application of the BEK and BEP distributions to *complement* those of of general application (LP3, GVE and LOG); especially for selecting the predictions to be adopted in the three extreme return periods of 1,000, 5,000 and 10,000 years. The later was verified with the BEK distribution that was adopted in two *Qp* and two *Vol* registers (from 4 to 7) of the eight processed ones and the BEP model was adopted in two *PMD* registers (9 and 11) of the three analyzed.

It should be noted that these Conclusions are based on the results of the 11 records processed and therefore, they may give the impression

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

that the BEK and BEP distributions are better than those of general application, but this is not the case; rather, only practical and feasible options should be considered for extreme hydrological data frequency analyses.

Acknowledgment

Comments and corrections suggested by anonymous referees C, D, E and the Editor are gratefully appreciated. These corrections allowed to refine the text, adapt it to the Mexican hydrological context and helped to justify the proposal for the routine application of the beta-kappa and beta-Pareto distributions, in the analysis of extreme data frequencies: floods (*Qp*, *Vol*) and maximum daily rainfall.

References

Aldama, A. A., Ramírez, A. I., Aparicio, J., Mejía-Zermeño, R., & Ortega-Gil, G. E. (2006). *Seguridad hidrológica de las presas en México.* Jiutepec, México: Instituto Mexicano de Tecnología del Agua.

Benson, M. A. (1962). Plotting positions and economics of engineering planning. *Journal of Hydraulics Division*, 88(6), 57-71. DOI: 10.1061/jYCEAj.0001293

Bobée, B., & Ashkar, F. (1991). Chapter 1. Data requirements for hydrologic frequency analysis. In: *The Gamma Family and derived distributions applied in Hydrology* (pp. 1-12). Littleton, USA: Water Resources Publications.

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

Campos-Aranda, D. F. (1998). Distribución de probabilidades β-ρ: descripción y aplicación en hidrología superficial. En: *XV Congreso Nacional de Hidráulica (AMH)* (pp. 965-971), del 13 al 16 de octubre, Oaxaca, Oaxaca, México.

Campos-Aranda, D. F. (2002). Contraste de seis métodos de ajuste de la distribución Log-Pearson tipo III en 31 registros históricos de eventos máximos anuales. *Ingeniería Hidráulica en México,* 17(2), 77-97.

Campos-Aranda, D. F. (2019). Mejores FDP en 19 series amplias de *PMD* anual del estado de San Luis Potosí, México. *Tecnología y ciencias del agua*, 10(5), 34-74. DOI: 10.24850/j-tyca-2019-05-02

Coles, S. (2001). Theme 2.6.7: Model diagnostics. In: *An introduction to statistical modeling of extreme values* (pp. 36-44). London, UK: Springer-Verlag.

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

Domínguez, M., R., & Arganis, M. L. (2012). Validation of method to estimate design discharge flow for dam spillways with large regulating capacity. *Hydrological Sciences Journal*, 57(3), 460-478. DOI: 10.1080/02626667.2012.665993

Hosking, J. R. M., & 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.

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

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.

Mason, S. J., Waylen, P. R., Mimmack, G. M., Rajaratnam, B., & Harrison, J. M. (1999). Changes in extreme rainfall events in South Africa. *Climatic Change*, 41(2), 249-257. DOI: 10.1023/A:1005450924499

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

Mielke, P. W. (1973). Another family of distributions for describing and analyzing precipitation data. *Journal of Applied Meteorology*, 12(2), 275-280. DOI: 10.1175/1520-0450(1973)012<0275:AFODFD>2.0.CO;2

Mielke, P. W., & Johnson, E. S. (1974). Some generalized Beta distributions of the second kind having desirable application features in hydrology and meteorology. *Water Resources Research*, 10(2), 223-226. DOI: 10.1029/WR010i002p00223

Murshed, M. S., Kim, S., & Pak, J. S. (2011). Beta-κ distribution and its application to hydrologic events. *Stochastic Environmental Research and Risk Assessment*, 25(7), 897-911. DOI: 10.1007/s00477-011-0494-4

Nguyen, T. H., El Outayek, S., Lim, S. H., & Nguyen, T. V. T. (2017). A systematic approach to selecting the best probability models for annual maximum rainfalls - A case study using data in Ontario (Canada). *Journal of Hydrology,* 553, 49-58. DOI: 10.1016/j.jhydrol.2017.07.052

2024, Instituto Mexicano de Tecnología del Agua.
Open Access bajo la licencia CC BY-NC-SA 4.0
(https://creativecommons.org/licenses/by-nc-sa/4.0/)

Oberhettinger, F. (1972). Chapter 15. Hypergeometric functions. In: Abramowitz, M., & Stegun, I. A. (eds.). *Handbook of mathematical functions* (pp. 555-566). New York, USA: Dover Publications.

Öztekin, T. (2007). Wakeby distribution for representing annual extreme and partial duration rainfall series. *Meteorological Applications*, 14(4), 381-387. DOI: 10.1002/met.37

Rao, A. R., & Hamed, K. H. (2000). Theme 1.8. Tests on hydrologic data. Chapters 7, 8 and 9. In: *Flood frequency analysis* (pp. 12-21, 207-321). Boca Raton, USA: CRC Press.

Stedinger, J. R. (2017). Chapter 76. Flood frequency analysis. In: Singh, V. P. (ed.). *Handbook of applied hydrology* (2nd ed.). (pp. 76.1-76.8). New York, USA: McGraw-Hill Education.

Strupczewski, W. G., Markiewicz, I., Kochanek, K., & Singh, V. P. (2008). Short walk into two-shape-parameter flood frequency distributions. In: Singh, V. P. (ed.). *Hydrology and hydraulics* (pp. 669-716). Highlands Ranch, USA: Water Resources Publications.

Teegavarapu, R. S. V., Salas, J. D., & Stedinger, J. R. (2019). Chapter 1. Introduction. In: *Statistical analysis of hydrologic variables* (pp. 1-4). Reston, USA: American Society of Civil Engineers.

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

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)

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

Wilks, D. S. (1993). Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series. *Water Resources Research*, 29(10), 3543-3549. DOI: 10.1029/93WR01710

Wilks, D. S. (2011). Theme 4.5. Qualitative assessments of the goodness fit. In: *Statistical methods in the atmospheric sciences* (3rd ed.) (pp. 112-116). San Diego, USA: Academic Press (Elsevier).

2024, Instituto Mexicano de Tecnología del Agua. Open Access bajo la licencia CC BY-NC-SA 4.0 (https://creativecommons.org/licenses/by-nc-sa/4.0/)