

Mejores FDP en 19 series amplias de *PMD* anual del estado de San Luis Potosí, México
Best PDF in 19 large annual series of *MDP* from the San Luis Potosi state, Mexico

Daniel Francisco Campos-Aranda¹

¹Profesor jubilado de la Universidad Autónoma de San Luis Potosí, México, campos_aranda@hotmail.com

Autor para correspondencia: Daniel Francisco Campos-Aranda, campos_aranda@hotmail.com

Resumen

Todas las obras hidráulicas se planean y diseñan con base en las crecientes de diseño. Sin información hidrométrica, estas predicciones se estiman con métodos hidrológicos que transforman lluvias de diseño en los gastos buscados. La escasez de pluviógrafos origina que las lluvias de diseño se estimen a partir de los registros de precipitación máxima diaria (*PMD*) anual de los pluviómetros. El análisis probabilístico de las series de *PMD* anual es *idéntico* al de las crecientes; pero aún no se han definido funciones de distribución de probabilidades (FDP) adecuadas o que se deban aplicar bajo precepto, por lo cual es necesario probar varias. Primero se buscó la mejor FDP en el diagrama de cocientes *L*, que incluye seis modelos de tres parámetros de ajuste. Se realiza una *selección objetiva* al emplear la distancia absoluta ponderada en los 19 registros de *PMD* anual con más de 50 datos del estado de San Luis Potosí, México. Después se describen y aplican ocho índices de *habilidad descriptiva* (HD) a las ocho FDP que fueron contrastadas en cada uno de los 19 registros de *PMD*. Los resultados se concentran y analizan por áreas geográficas del estado: Altiplano Potosino y Zona Media. Se obtuvo que la FDP Wakeby es un modelo de

gran HD y por ello se sugiere que su aplicación se realice bajo precepto. También se definen las dos mejores opciones de FDP en cada uno de los 19 registros procesados de acuerdo con los ocho índices de HD. Por último, se realiza un contraste de predicciones con periodos de retorno de 50, 100, 500 y 1000 años, para explorar de manera somera la *habilidad predictiva* de las FDP encontradas como mejores opciones. En cada registro se aplican cuatro FDP: la obtenida según el diagrama de cocientes L, las dos mejores FDP según los ocho índices de HD y la distribución Wakeby. Se concluye que el uso del diagrama de cocientes L y la aplicación de los ocho índices de HD son adecuados y conducen a una buena aproximación, pues no se tuvo dificultad para seleccionar las predicciones adoptadas y la similitud que mostraron estas estimaciones en cada registro genera confianza en tales estimaciones.

Palabras clave: diagrama de cocientes L, error estándar de ajuste, error relativo estándar de ajuste, error absoluto medio, error absoluto máximo, criterio de información de Akaike, coeficiente de correlación de Q-Q, índices de concordancia y habilidad predictiva.

Abstract

All hydraulic works are planned and designed based on floods design. Without hydrometric information, these predictions are estimated using hydrological methods that yield the sought flows by means of design rains. Design rainfall is estimated based on pluviometer records of annual maximum daily precipitation (*MDP*) due to the shortage of pluviographs. The probabilistic analysis of the annual *MDP* series is *identical* to that of the floods; however, neither adequate probability distribution functions (PDFs) nor those that should be applied by precept have been defined so far, hence the need to try several. First, the best PDF was searched for using the L-ratio diagram, which includes six models with three fit parameters. An *objective selection* is made by using the weighted absolute distance, in the 19 annual *MDP* records with more than 50 data from the state of San Luis Potosi, Mexico. Then eight *descriptive ability* (DA) indexes are described and applied to the eight PDFs that were compared, in each of the 19 *MDP* records. The results are concentrated and analyzed for geographic areas of the state: Potosino Plateau and Middle Zone. Results show that Wakeby PDF is a model having high DA and for that reason, its application is suggested as precept. The two best PDF options are also highlighted in each of the 19 records processed, according to the eight DA indexes. Finally, a

comparison of predictions with periods of return of 50, 100, 500 and 1000 years is carried out to explore shallowly the *predictive ability* of the PDFs found as best options. In each registry four PDFs are applied, the one obtained according to the L-ratio diagram, the two best PDFs according to the eight DA indices and the Wakeby distribution. It is concluded that the use of the L-ratio diagram and the application of the eight DA indexes are adequate and lead to a good approximation, since it was not difficult to select the adopted predictions, besides the similarity of the predictions calculated in each register promotes confidence in such estimations.

Keywords: L-ratio diagram, standard error of fit, relative standard error of fit, mean absolute error, maximum absolute error, Akaike information criterion, Q-Q correlation coefficient, concordance indexes and predictive ability.

Recibido: 16/11/2017

Aceptado: 22/11/2018

Introducción

Las *crecientes de diseño* dimensionan hidrológicamente todas las obras hidráulicas en las diferentes etapas que atraviesan. En sitios de interés y sus cuencas respectivas, que no tienen información de gastos máximos anuales, las crecientes de diseño se deben estimar con base en *métodos hidrológicos* que transforman una lluvia de diseño en el hidrograma de respuesta o en el gasto pico buscado (Mujumdar & Nagesh-Kumar, 2012). Las lluvias de diseño proceden de las curvas intensidad-duración-frecuencia (IDF) que caracterizan la manera como llueve en la zona de estudio. La escasez de registros pluviográficos impide la construcción de las curvas IDF y por ello se recurre a su estimación, con base en los registros disponibles de las estaciones pluviométricas de mayor cobertura y más amplias (Teegavarapu, 2012; Johnson & Sharma, 2017).

Los registros de precipitación máxima diaria (*PMD*) anual se procesan probabilísticamente de forma *idéntica* a los de gasto máximo anual o crecientes, pasando por cuatro etapas: (1) búsqueda de registros, que incluye su depuración y la verificación de su homogeneidad; (2) elección de una *función de distribución de probabilidades* acumuladas o FDP, es decir, del modelo probabilístico que permitirá obtener las *predicciones* asociadas con bajas probabilidades de excedencia; (3) aplicación de uno o varios métodos de estimación de los parámetros de ajuste de la FDP, y (4) validación de la FDP adoptada y de sus predicciones (Rao & Hamed, 2000; Meylan, Favre, & Musy, 2012; Stedinger, 2017).

El *objetivo* del estudio consistió en seleccionar las mejores FDP que se deben aplicar en el análisis probabilístico de series de *PMD* anual. Primero se expone y aplica el diagrama de cocientes L, a través de la distancia absoluta ponderada para adoptar objetivamente la mejor FDP de las seis que incluye; después se sigue un enfoque basado en ocho índices de habilidad descriptiva para la selección entre las ocho FDP que se contrastan; se exponen y analizan los resultados concentrados del Altiplano Potosino y de la Zona Media del estado de San Luis Potosí, México, en las cuales se procesaron 9 y 10 registros de *PMD* anual, con 50 o más datos; por último, se realiza un contraste de predicciones que tienen periodos de retorno de 50, 100, 500 y 1000 años, y se formulan conclusiones respecto a la capacidad predictiva de las FDP seleccionadas.

Datos de *PMD* procesados

Series depuradas

Con base en el archivo en Excel actualizado hasta 2015 de las estaciones climatológicas del estado de San Luis Potosí, proporcionado al autor por la Dirección Local de la Comisión Nacional del Agua (Conagua), se seleccionaron todos los registros de precipitación máxima

diaria (*PMD*) anual con más de 40 valores y escasos datos faltantes, y se obtuvieron 100 series (Campos-Aranda, 2018). Después se llevó a cabo una ratificación de sus valores extremos mínimos y máximos con auxilio de la Conagua, para obtener las llamadas *series depuradas*.

Pruebas de homogeneidad aplicadas

A cada serie depurada se le aplicó el Test de Von Neumann, como prueba general, que detecta la no aleatoriedad por componentes determinísticas no especificadas y varias pruebas específicas: Anderson y Sneyers de persistencia, Kendall y Spearman de tendencia, Bartlett de inconsistencia en la dispersión y Cramer de cambios en la media. Estas pruebas se pueden consultar en WMO (1971), y Machiwal y Jha (2008). Se encontró que un total de 39 series resultaron o no aleatorias, o bien presentaron persistencia y/o tendencia.

Series por procesar

Las 39 series no homogéneas se eliminaron, así como las de amplitud menor de 50 datos; con tales restricciones se dispuso de 35 registros de *PMD* anual en el estado de San Luis Potosí. Nueve series pertenecen al Altiplano Potosino (AP), 10 a la zona media (ZM) y 16 a la Región Huasteca. En este estudio se procesaron las 19 series que se ubican en los climas áridos y semiáridos de las áreas geográficas AP y ZM, cuyas altitudes en general son superiores a los mil metros. En la Tabla 1 se exponen las estaciones procesadas, su altitud, amplitud de registro y valores mínimo y máximo. En la Tabla 2 se muestran sus estadísticos y cocientes de momentos L (ecuaciones (6) a (8)). Las primeras nueve estaciones climatológicas pertenecen al AP y las 10 restantes a la ZM. En la Figura 1 se muestra su ubicación geográfica en el estado de San Luis Potosí; México.

Tabla 1. Altitud, amplitud de registro y valores mínimos y máximos de las 19 series de precipitación máxima diaria (*PMD*) anual del estado de San Luis Potosí, México.

Núm.	Nombre de la estación:	Altitud ¹ (msnm)	Registro		<i>PMD</i>	
			Periodo	n^2	Mín	Máx
1	Cedral	1 702	1946-2014	66	19.0	315.8
2	Charcas	2 126	1961-2014	54	12.0	117.0
3	La Maroma	1 900	1965-2014	50	16.0	140.1
4	Los Filtros (SLP)	1 904	1949-2014	66	15.9	111.0
5	Matehuala	1 630	1961-2014	54	25.5	200.0
6	Mexquitic	1 749	1943-2014	72	12.0	107.0
7	Peñón Blanco	2 099	1950-2014	57	13.0	235.0
8	Santo Domingo	1 415	1961-2013	52	19.0	270.0
9	Vanegas	1 746	1964-2013	50	12.0	90.0
10	Armadillo de los Infante	1 636	1961-2013	52	22.0	133.0
11	Cárdenas	1 353	1946-2013	61	21.5	180.5
12	Lagunillas	908	1954-2013	53	30.0	210.0
13	Ojo de Agua	1 148	1960-2013	52	45.0	300.2
14	Ojo de Agua Seco	1 077	1961-2013	51	26.5	172.5
15	Paso de San Antonio	1 246	1958-2013	52	26.0	200.0
16	Rayón	1 258	1961-2013	51	33.5	330.0
17	Río Verde	987	1961-2013	52	27.0	126.3
18	San Francisco	1 066	1961-2013	50	12.0	135.0
19	San José Alburquerque	1 863	1961-2014	50	21.0	126.5

Tabla 2. Parámetros estadísticos y cocientes de momentos L de las 19 series de precipitación máxima diaria (*PMD*) anual del estado de San Luis Potosí, México.

Núm. est.	Parámetros estadísticos ³				Cocientes ⁴ de momentos L			
	$\bar{x}=I_1$	I_2	S	Cs	t_3	t_4	t_3^{\ln}	t_4^{\ln}
1	47.1	13.465	38.6	5.501	0.49274	0.41041	0.20245	0.20806
2	48.8	12.068	22.0	0.974	0.17331	0.18124	-0.06888	0.17927

3	46.6	10.807	21.3	2.006	0.27042	0.21953	0.05484	0.16449
4	43.0	8.448	15.7	1.315	0.13515	0.16117	-0.04764	0.13773
5	59.3	14.243	29.2	2.471	0.28588	0.23605	0.06420	0.13996
6	47.8	9.521	17.1	0.540	0.05475	0.15489	-0.14445	0.17245
7	47.8	15.395	40.2	3.669	0.51286	0.44109	0.19048	0.26367
8	57.6	16.564	37.8	3.685	0.31725	0.27714	0.02817	0.14591
9	37.6	10.013	18.4	1.028	0.22725	0.16424	-0.00150	0.14389
10	57.5	14.504	27.3	1.287	0.28601	0.18574	0.08089	0.14044
11	67.7	20.216	38.9	1.457	0.33709	0.18138	0.11049	0.12095
12	77.8	18.480	35.0	1.525	0.22455	0.15699	0.02475	0.10105
13	91.4	21.459	46.5	2.586	0.38619	0.31739	0.16556	0.20467
14	69.3	14.883	28.9	1.694	0.28166	0.21420	0.08873	0.16202
15	69.3	13.793	27.6	2.219	0.22726	0.23726	0.01936	0.18991
16	76.4	19.343	45.5	3.765	0.39215	0.34343	0.13435	0.21192
17	58.4	12.973	23.4	0.934	0.20926	0.09728	0.05010	0.06390
18	46.7	12.757	24.4	1.461	0.23966	0.24363	-0.03897	0.21540
19	50.1	11.466	21.7	1.434	0.19616	0.16114	0.00396	0.08689

Simbología:

¹metros sobre el nivel del mar.

²número de datos procesados.

³ \bar{x} = media aritmética, milímetros.

l_1, l_2 = momentos L de orden 1 y 2.

S = desviación estándar, en milímetros.

C_s = coeficiente de asimetría, adimensional.

⁴ t_3 = cociente L de asimetría, adimensional.

t_4 = cociente L de curtosis, adimensional.

t_3^{\ln} = t_3 de los logaritmos naturales de los datos, adimensional.

t_4^{\ln} = t_4 de los logaritmos naturales de los datos, adimensional.

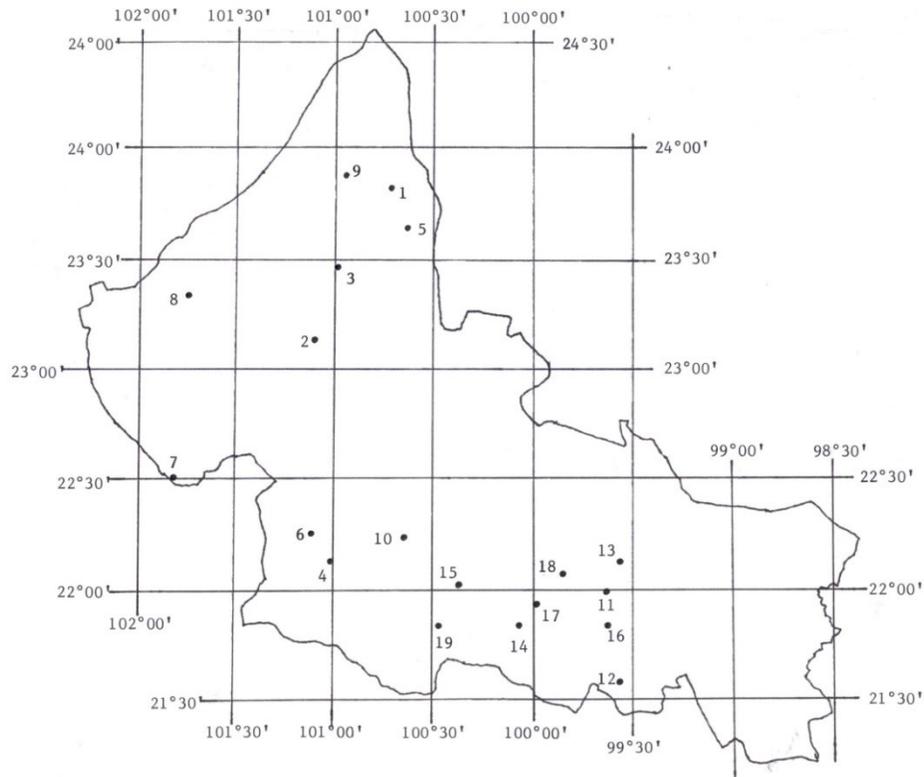


Figura 1. Localización geográfica de las 19 series de *PMD* anual procesadas del estado de San Luis Potosí, México.

Diagrama de cocientes L

Momentos L de la muestra

Los momentos L son combinaciones lineales de los momentos de probabilidad ponderada (b_r), por ello son robustos ante los valores dispersos de muestra. Su cálculo comienza por ordenar la serie disponible (x_i) de *PMD* anual de menor a mayor ($x_1 \leq x_2 \leq \dots \leq x_n$) y

entonces (Hosking & Wallis, 1997; Rao & Hamed, 2000; Asquith, 2011; Stedinger, 2017):

$$b_r = \frac{1}{n} \sum_{i=r+1}^n \frac{(i-1)(i-2)\dots(i-r)}{(n-1)(n-2)\dots(n-r)} x_i \quad (1)$$

En la expresión anterior, el número de orden r varía de 0 a 3 y n es el número de datos de la serie de *PMD* anual. Se deduce que b_0 es igual a la media aritmética. Los momentos L de la muestra (l) y sus respectivos cocientes (t) de similitud con los coeficientes de variación, asimetría y curtosis son:

$$l_1 = b_0 \quad (2)$$

$$l_2 = 2 \cdot b_1 - b_0 \quad (3)$$

$$l_3 = 6 \cdot b_2 - 6 \cdot b_1 + b_0 \quad (4)$$

$$l_4 = 20 \cdot b_3 - 30 \cdot b_2 + 12 \cdot b_1 - b_0 \quad (5)$$

$$t_2 = l_2/l_1 \quad (6)$$

$$t_3 = l_3/l_2 \quad (7)$$

$$t_4 = l_4/l_2 \quad (8)$$

Diagrama de cocientes L

Tiene en el eje de las abscisas a t_3 y en el de las ordenadas a t_4 . Las FDP de tres parámetros de ajuste son líneas curvas con las ecuaciones de tipo polinomio siguientes (Hosking & Wallis, 1997):

Logística Generalizada (LOG):

$$t_4^{\text{LOG}} = 0.16667 + 0.83333 \cdot t_3^2 \quad (9)$$

Pareto Generalizada (PAG):

$$t_4^{\text{PAG}} = 0.20196 \cdot t_3 + 0.95924 \cdot t_3^2 - 0.20096 \cdot t_3^3 + 0.04061 \cdot t_3^4 \quad (10)$$

Log-normal (LGN):

$$t_4^{\text{LGN}} = 0.12282 + 0.77518 \cdot t_3^2 + 0.12279 \cdot t_3^4 - 0.13638 \cdot t_3^6 + 0.11368 \cdot t_3^8 \quad (11)$$

Pearson tipo III (PT3):

$$t_4^{\text{PT3}} = 0.12240 + 0.30115 \cdot t_3^2 + 0.95812 \cdot t_3^4 - 0.57488 \cdot t_3^6 + 0.19383 \cdot t_3^8 \quad (12)$$

y General de Valores Extremos (GVE):

$$t_4^{\text{GVE}} = 0.10701 + 0.11090 \cdot t_3 + 0.84838 \cdot t_3^2 - 0.06669 \cdot t_3^3 + SF \quad (13)$$

siendo: $SF = 0.00567 \cdot t_3^4 - 0.04208 \cdot t_3^5 + 0.03763 \cdot t_3^6$

Utilizando los logaritmos de los datos en las ecuaciones (1) a (8) se obtienen los cocientes L logarítmicos y entonces se puede utilizar la expresión (12) para evaluar la FDP Log-Pearson tipo III. En la Figura 2 se muestra una porción del diagrama de cocientes de momentos L, procedente de Hosking y Wallis (1997).

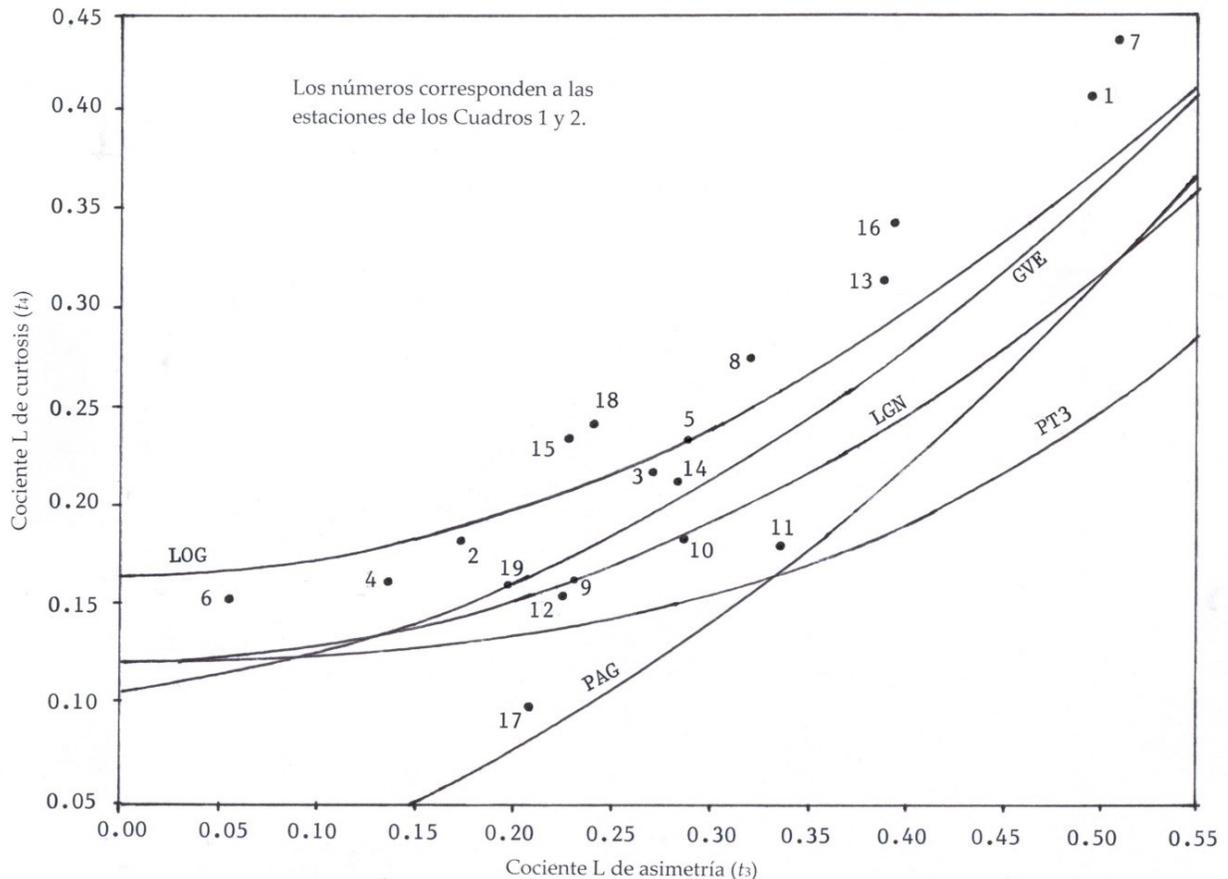


Figura 2. Diagrama de cocientes L y valores puntuales correspondientes a las 19 series de *PMD* anual procesadas del estado de San Luis Potosí, México.

Distancia absoluta ponderada

Uno de los enfoques recientes para la selección de la mejor FDP en los ámbitos local y regional consiste en llevar al diagrama de cocientes L los valores de la muestra (t_3 y t_4) y definir su cercanía a alguna de las curvas, para obtener el mejor modelo probabilístico. Lo anterior es

relativamente simple en los análisis locales, pero se complica en el enfoque regional, como han señalado Peel, Wang, Vogel y McMahon (2001), pues entonces se tiene una nube de puntos. Para evitar la subjetividad en la selección de la FDP, se ha propuesto evaluar la *Distancia Absoluta Ponderada (DAP)* con la expresión siguiente (Yue & Hashino, 2007):

$$DAP = \frac{\sum_{j=1}^{NE} n_j \cdot |t_4[t_3^{obs}(j)] - t_4^{obs}(j)|}{\sum_{j=1}^{NE} n_j} \quad (14)$$

donde NE es el número de estaciones que integran la región; n_j , el número de datos de cada registro de *PMD*; $t_3^{obs}(j)$ y $t_4^{obs}(j)$, los cocientes L de asimetría y curtosis de cada serie, y $t_4[t_3^{obs}(j)]$ es el valor teórico del cociente L de curtosis calculado con cada FDP (ecuaciones (9) a (13)), para el valor observado del cociente L de asimetría. Una FDP con el menor valor de la *DAP* es la mejor para los datos, locales ($NE = 1$) o regionales.

Funciones de distribución por contrastar

Con base en el diagrama de cocientes L se pueden probar seis FDP y como se aceptó *a priori* eliminar la que no fuera seleccionada al menos una vez en los 19 registros de *PMD* anual por procesar, no se contrastó la distribución Pearson Tipo III y entonces se probarán los modelos LOG, PAG, LGN, GVE y LP3.

Las ocho FDP que fueron contrastadas incluyen a los modelos Beta Kappa y Beta Pareto propuestos por Mielke y Johnson (1974), que no son populares en México, pero que fueron comparadas en un trabajo pionero de selección óptima de distribuciones con tres parámetros de ajuste en registros de *PMD* anual, el de Wilks (1993), con buenos resultados para la Beta- κ (BEK) en series de máximos y de valores arriba de un umbral para la Beta- P (BEP); Nguyen, El Outayek, Lim, y Nguyen (2017) también las incluyen en su contraste.

Por último, se incluyó a la FDP Wakeby (WAK) que tiene cinco parámetros de ajuste, fue propuesta por Houghton (1978), y permite modelar de manera separada los extremos izquierdo y derecho de la muestra. Nguyen *et al.* (2017) encuentran que la distribución Wakeby es la de mejor habilidad descriptiva.

Para evitar más variables involucradas en la selección de las mejores FDP en modelos que tienen varios métodos de estimación de sus parámetros de ajuste, se aplicó *exclusivamente* el de momentos L, que ha probado ser consistente y exacto. Tal método en los modelos GVE, LOG y PAG no se expone, pues ha sido descrito en varios artículos del autor, como en Campos-Aranda (2015; 2016). También se puede consultar en Hosking y Wallis (1997), Rao y Hamed (2000), y Stedinger (2017). Las FDP Log-normal (LGN) y Wakeby (WAK) se ajustaron con el método de momentos L expuesto por Hosking y Wallis (1997).

Respecto a la distribución Log-Pearson tipo III (LP3) se ajustó con el método de momentos en los dominios logarítmico y real (WRC, 1977; Bobée & Ashkar, 1991; Campos-Aranda, 2002), seleccionado el de menor error estándar de ajuste (Kite, 1977). Las distribuciones Beta se ajustaron con el método iterativo de máxima verosimilitud de Mielke y Johnson (1974), adoptando como valor inicial del parámetro de escala la media del registro de *PMD* y del parámetro de forma un valor de cinco. Se llevó el máximo de iteraciones a dos mil.

Índices de habilidad descriptiva

Gráficos de diagnóstico

Los gráficos *P-P* y *Q-Q* de probabilidad empírica contra calculada y de cantidad observada contra estimada se han popularizado (Coles, 2001; Wilks, 2011), y constituyen una manera simple y efectiva de comparar los resultados de una FDP contrastada. Su desventaja radica en la

apreciación subjetiva que se hace al comparar diversas FDP con tales gráficas, ya que no se dispone de un valor numérico (Nguyen *et al.*, 2017). Campos-Aranda (2015) ha expuesto tales gráficos y visualiza más útil la gráfica de $Q-Q$ para observar predicciones sobreestimadas (por quedar arriba de la recta a 45°) o subestimadas (por estar debajo de la recta a 45°).

Pruebas estadísticas

Estos tests como la J_i cuadrada, el de Kolmogorov-Smirnov o el de Anderson-Darling permiten justificar que una muestra se puede aceptar si proviene de una distribución específica que se esté analizando. En las dos primeras pruebas su desventaja radica en tener poca potencia y en la segunda de sólo ser aplicable para una FDP específica (Meylan *et al.*, 2012).

Índices de bondad de ajuste

Tienen la ventaja de ser de cálculo fácil y por lo común involucran la diferencia entre los valores observados x_i y los estimados con la FDP que se contrasta \hat{x}_i (Pandey & Nguyen, 1999; Zalina, Desa, Nguyen, & Kassim, 2002). Meylan *et al.* (2012) señalan dos desventajas de estos índices: (1) requieren de la aplicación de una fórmula empírica para estimar la probabilidad de cada dato; (2) no aportan un estimado de la probabilidad involucrada en el rechazo o aceptación de la FDP, y quizás ello es su ventaja al comparar varios modelos probabilísticos. En este trabajo se aplicaron ocho índices de bondad de ajuste. Nguyen *et al.* (2017) usan los cinco definidos en las ecuaciones (17) a (22), exceptuando la (19), e incluyen otro similar al AIC (ecuación (21)).

Fórmula empírica de probabilidad

Meylan *et al.* (2012) indican que todas las ecuaciones empíricas que asignan probabilidades se basan en ordenar la muestra o serie disponible en magnitudes ascendentes, haciendo posible asociar un orden i con cada dato y después hacen uso de la fórmula general siguiente, la cual asegura simetría respecto a la mediana:

$$p_{i:n} = \frac{i-c}{n+1-2c} \quad (15)$$

en la cual c es una cantidad constante y n es el tamaño del registro o serie de *PMD* anual.

Haktanir (1991), describe un hallazgo práctico de 1971 por Stipp y Young, que procesaron 37 registros anuales de crecientes de exactamente 20 datos cada uno, todos en EUA. Estimaron la probabilidad de cada evento máximo y mínimo de cada serie con base en la distribución Log-Pearson tipo III y después igualaron tales valores con el obtenido de la ecuación (15), encontrando que c tenía una magnitud aproximada de 0.40, por lo cual:

$$p_{i:n} = \frac{i-0.40}{n+0.20} \quad (16)$$

Haktanir (1991) también indica que años después Cunnane llega a la misma ecuación (16) de manera teórica e independiente, estableciendo que tal fórmula no es específica de una FDP, y que sus resultados son insesgados y tienen error cuadrado mínimo. Cunnane (1978) además encuentra, con argumentos estadísticos, que la popular fórmula de Weibull (Benson, 1962) sólo es aplicable para una distribución uniforme, por lo cual no es conveniente emplearla en series de valores hidrológicos extremos (crecientes, sequías, *PMD*, etc.).

Error estándar de ajuste (EEA)

Es el índice más común (Chai & Draxler, 2014). Se estableció a mediados de la década de 1970 (Kite, 1977), y se ha aplicado en México haciendo uso de la fórmula empírica de Weibull (Benson, 1962). Ahora se aplicará utilizando la fórmula de Cunnane (ecuación (16)), que de acuerdo con Stedinger (2017) conduce a probabilidades de no excedencia ($p_{i:n}$) aproximadamente insesgadas para muchas FDP, por ello fue adoptada por Nguyen *et al.* (2017). La expresión del *EEA* es:

$$EEA = \left[\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{(n-np)} \right]^{1/2} \quad (17)$$

en la cual x_i son los datos observados ordenados de menor a mayor; \hat{x}_i , los estimados para la probabilidad estimada con la ecuación (16) y la FDP que se contrasta; np , el número de parámetros de ajuste de la FDP, con cinco para la Wakeby y tres para el resto de las contrastadas; el numerador de esta ecuación es la suma de errores al cuadrado (*SEC*); el *EEA* tiene las unidades de la variable x_i .

Error relativo estándar de ajuste (EREA)

En el *EEA* todas las diferencias o residuos se elevan al cuadrado y ello implica dar mayor peso a los valores altos de la muestra; para reducir este impacto, se aplica el *EREA*, que es adimensional, su ecuación es (Pandey & Nguyen, 1999; Nguyen *et al.*, 2017):

$$EREA = \left[\frac{1}{(n-np)} \sum_{i=1}^n \left(\frac{x_i - \hat{x}_i}{x_i} \right)^2 \right]^{1/2} \quad (18)$$

Error absoluto medio (*EAM*)

Sus ventajas radican en tener las unidades de la variable, al igual que el *EEA*, y evitar que el impacto de los valores dispersos sea elevado al cuadrado, y por ello $EEA \geq EAM$ (Willmott & Matsuura, 2005). Su expresión es (Pandey & Nguyen, 1999; Nguyen *et al.*, 2017):

$$EAM = \frac{\sum_{i=1}^n |x_i - \hat{x}_i|}{(n-np)} \quad (19)$$

Error absoluto máximo (*EAMx*)

Este índice muestra al más grande de los errores o residuales absolutos, por ello Nguyen *et al.* (2017) han señalado que está relacionado con el estadístico de la prueba de Kolmogorov-Smirnov. Su ecuación es:

$$EAMx = \max(|x_i - \hat{x}_i|) \quad (20)$$

Criterio de información de Akaike (*AIC*)

Este índice emplea en su concepción original el valor máximo alcanzado en la función de verosimilitud durante el ajuste, con tal método, de la FDP que se contrasta. Para poder aplicar el índice, Nguyen *et al.* (2017)

emplean la suma de errores al cuadrado (*SEC*) como indicador de la calidad del ajuste y entonces su ecuación es:

$$AIC = 2 \cdot np + n \cdot \ln(SEC) + \frac{2(np+1) \cdot (np+2)}{(n-np-2)} \quad (21)$$

En estos primeros cinco índices, el valor menor de ellos indica el mejor ajuste de la FDP y su magnitud máxima el peor ajuste de la FDP contrastada. En los tres índices siguientes ocurre lo contrario, de manera que un valor máximo indica un buen ajuste de la FDP y viceversa.

Coeficiente de correlación de Q-Q (COC)

Este índice ha sido utilizado como el criterio principal de selección por Zalina *et al.* (2002) e indica la dependencia lineal que existe en el gráfico de Q-Q, por lo cual varía de 0 a 1. Los valores del COC cercanos a la unidad indican un buen ajuste de la FDP que se contrasta; su ecuación es:

$$COC = \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (\hat{x}_i - \hat{x}_i^m)}{[\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (\hat{x}_i - \hat{x}_i^m)^2]^{\frac{1}{2}}} \quad (22)$$

en donde \bar{x} y \hat{x}_i^m son los valores medios de los datos y de los valores estimados.

Índices de concordancia (d_1 , d_2)

De acuerdo con Legates y McCabe (1999), Willmott señaló desde inicios de los años de 1980 que el índice *COC* está limitado y es insensible a las diferencias y varianzas entre x_i y \hat{x}_i por usar valores estandarizados con sus medias, y por ello propone el *índice de concordancia* (d_2) adimensional, con la expresión:

$$d_2 = 1.0 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (|x_i - \bar{x}| + |\hat{x}_i - \bar{x}|)^2} \quad (23)$$

El numerador es la *SEC* y el denominador se llama el *error potencial*, por ser el valor máximo que la diferencia entre x_i y \hat{x}_i puede alcanzar en cada dato. Legates y McCabe (1999) señalan que el índice de concordancia (d_2) también resulta sensitivo a los valores dispersos por usar las diferencias entre x_i y \hat{x}_i al cuadrado, y por ello proponen el *índice de concordancia modificado* (d_1), en el cual el numerador y denominador no se tienen al cuadrado sino con valor absoluto, esto es:

$$d_1 = 1.0 - \frac{\sum_{i=1}^n |x_i - \hat{x}_i|}{\sum_{i=1}^n (|x_i - \bar{x}| + |\hat{x}_i - \bar{x}|)} \quad (24)$$

Tanto d_2 como d_1 varían de 0 a 1, con una interpretación igual al *COC*, por lo general $d_2 \geq d_1$.

Revisión de la habilidad predictiva

Enfoques disponibles

La *habilidad predictiva* de las FDP está relacionada con las *predicciones* realizadas a periodos de retorno (Tr) superiores al tamaño del registro (n), o bien a la comparación con los valores extremos observados en el

registro. Existen tres enfoques para probar o verificar tal habilidad predictiva, el primero es el más simple y consiste en el contraste numérico de las predicciones obtenidas con cada FDP para diversos Tr adoptados.

El segundo enfoque de verificación de la habilidad predictiva fue ejemplificado por Haktanir (1992). Utiliza muestras aleatorias generadas con modelos básicos (LP3, GVE y WAK), que emplean los parámetros de ajuste obtenidos en los registros adoptados como representativos de las regiones geográficas estudiadas. Las mejores FDP se adoptan con base en el menor *EREA*.

A partir de la década de 1990 (Wilks, 1993; Zalina *et al.*, 2002; Nguyen *et al.*, 2017) se ha propuesto un tercer enfoque de contraste, basado en muestras aleatorias generadas a partir del registro histórico mediante muestreo con remplazo (técnica *bootstrap*), el cual contrasta predicciones dentro del registro obtenidas con las FDP probadas con los valores extremos de tal muestra. Las mejores FDP son las que muestran menor dispersión.

Enfoque adoptado

Corresponde al primero y más simple, pues se cuenta con las mejores FDP de cada registro de *PMD* anual, según el diagrama de cocientes *L* y de acuerdo con los ocho índices de habilidad descriptiva aplicados. Consiste en adoptar una de las FDP contrastadas y sus predicciones, siguiendo una regla establecida *a priori*, por ejemplo, adoptar los valores más críticos o mayores en la mayoría de los Tr contrastados, siempre y cuando tales predicciones sean similares, lo cual implica exactitud y genera confianza en las magnitudes adoptadas bajo tal esquema subjetivo.

Resultados según diagrama de cocientes *L*

La evaluación de la *distancia absoluta ponderada* (ecuación (14)) en cada una de las seis FDP del diagrama de cocientes de momentos L (ecuaciones (9) a (13)), haciendo uso de los valores de la Tabla 2, aportó los tres valores mínimos que se tienen en la Tabla 3, definiendo así la mejor FDP y las dos subsecuentes en el nivel *local* de cada serie de *PMD* anual. En la Figura 2 se han indicado los valores de t_3 y t_4 de cada registro, tomados de la Tabla 2. La mayoría de tales puntos definen su proximidad a una FDP, excepto las estaciones: Los Filtros (núm. 4), Santo Domingo (núm. 8) y Cárdenas (núm. 11), con cercanía al modelo LP3 debido a sus valores de t_3^{ln} y t_4^{ln} .

Tabla 3. Tres valores mínimos de las DAP y sus FDP respectivas del diagrama de cocientes de momentos L en las 19 series de *PMD* anual procesadas del estado de San Luis Potosí, México.

Núm.	Estación	DAP / FDP		
		1ª opción	2ª opción	3ª opción
1	Cedral	0.0414	0.0511	0.0717
		LOG	GVE	LP3
2	Charcas	0.0105	0.0299	0.0350
		LOG	GVE	LGN
3	La Maroma	0.0081	0.0218	0.0394
		LOG	GVE	LGN
4	Los Filtros (SLP)	0.0146	0.0207	0.0238
		LP3	LOG	GVE
5	Matehuala	0.0013	0.0163	0.0296
		LOG	LP3	GVE
6	Mexquitic	0.0143	0.0297	0.0316
		LOG	LGN	PT3
7	Peñón Blanco	0.0552	0.0635	0.1078
		LOG	GVE	LGN
8	Santo Domingo	0.0233	0.0266	0.0517
		LP3	LOG	GVE

9	Vanegas	0.0011	0.0110	0.0215
		LGN	GVE	LP3
10	Armadillo de los Infante	0.0012	0.0160	0.0208
		LGN	LP3	GVE
11	Cárdenas	0.0053	0.0115	0.0132
		LP3	PAG	PT3
12	Lagunillas	0.0052	0.0169	0.0170
		LGN	GVE	PT3
13	Ojo de Agua	0.0264	0.0450	0.0733
		LOG	GVE	LP3
14	Ojo de Agua Seco	0.0102	0.0186	0.0292
		GVE	LOG	LGN
15	Paso de San Antonio	0.0276	0.0620	0.0674
		LOG	GVE	LP3
16	Rayón	0.0486	0.0666	0.0838
		LOG	GVE	LP3
17	Río Verde	0.0148	0.0401	0.0593
		PAG	PT3	LP3
18	San Francisco	0.0291	0.0622	0.0759
		LOG	GVE	LGN
19	San José Alburquerque	0.0002	0.0083	0.0258
		GVE	LGN	PT3

De acuerdo con el resumen por áreas geográficas de la Tabla 4, se concluye que la primera o *mejor* opción de FDP es la logística generalizada (LOG) con 10 selecciones; le siguen la Log-normal (LGN) y la Log-Pearson tipo III (LP3) con tres selecciones, y la menos adecuada resultó ser la Pearson tipo III (PT3) con ninguna selección.

Tabla 4. Conteo de la mejor selección de cada FDP del diagrama de cocientes L, en las 19 series de *PMD* anual del estado de San Luis Potosí, México.

FDP	Altiplano Potosino	Zona Media	Sumas
-----	--------------------	------------	-------

Logística generalizada (LOG)	6	4	10
Log-normal (LGN)	1	2	3
Log-Pearson tipo III (LP3)	2	1	3
General de valores extremos (GVE)	0	2	2
Pareto generalizada (PAG)	0	1	1
Pearson tipo III (PT3)	0	0	0
Sumas	9	10	19

Resultados según habilidad descriptiva

Observaciones generales

Respecto a la distribución Log-Pearson tipo III (LP3), se obtuvo un menor error estándar de ajuste en el dominio real en las estaciones Charcas, Los Filtros, Mexquitic y Santo Domingo del Altiplano Potosino, y en las estaciones Paso de San Antonio y San Francisco de la Zona Media. En las 13 estaciones restantes, el mejor ajuste se obtuvo en el dominio logarítmico.

En relación con la distribución Wakeby (WAK), en un total de nueve registros se obtuvo que el parámetro de ubicación (ξ) resultó ligeramente superior al valor mínimo de la serie, lo cual estrictamente es incorrecto. En estos nueve registros se ajustó la distribución Wakeby con $\xi = 0$, según procedimiento de Hosking y Wallis (1997), y se contrastaron sus resultados (índices de habilidad descriptiva y predicciones) contra los ajustes previos. Exclusivamente en la estación San José Albuquerque se encontró más adecuada según los índices *EAMx* y *COC*, así como predicciones menos dispersas, es decir, mejoró su habilidad predictiva.

Resultados en el Altiplano Potosino

Concentrado de índices numéricos

En la Tabla 5 se han concentrado los tres valores característicos (mínimo, medio y máximo) de cada índice de *habilidad descriptiva* (HD) obtenidos con cada una de las ocho FDP contrastadas, en las nueve estaciones pluviométricas del Altiplano Potosino del estado de San Luis Potosí, México. La Tabla 5 resume las ocho tabulaciones no expuestas de resultados de cada índice con las ocho FDP contrastadas en los nueve registros de *PMD* anual del Altiplano Potosino. Exclusivamente para los valores medios se indican con paréntesis circular el mejor valor de índice respectivo, indicando tal magnitud la mejor FDP regional. También se marcan con paréntesis rectangular los peores índices regionales.

Tabla 5. Valores característicos de los ocho índices de habilidad descriptiva (HD) en las nueve series de *PMD* anual procesadas del Altiplano Potosino del estado de San Luis Potosí, México.

Índice de HD	FDP contrastadas							
	BEK	BEP	LGN	GVE	LOG	PAG	LP3	WAK
EEA mín	2.46	1.85	2.10	2.32	1.67	2.74	2.25	1.45
EEA med	8.13	7.19	7.10	6.98	(6.37)	[8.35]	6.54	7.06
EEA máx	20.40	16.42	15.68	14.85	14.80	15.96	13.40	15.73
EREA mín	0.041	0.038	0.033	0.035	0.041	0.073	0.048	0.040
EREA med	0.066	(0.060)	0.069	0.065	0.065	[0.108]	0.078	0.065
EREA máx	0.125	0.110	0.149	0.124	0.122	0.156	0.138	0.131
EAM mín	1.473	1.227	1.268	1.364	1.182	2.334	1.650	1.397
EAM med	3.107	2.669	3.013	2.802	2.618	[3.886]	3.127	(2.555)
EAM máx	6.705	5.409	5.498	4.978	4.920	6.737	6.989	5.106
EAMx mín	9.6	7.8	7.4	5.3	6.6	2.9	7.7	5.4
EAMx med	50.5	46.9	46.1	45.3	43.4	[52.5]	(39.7)	45.3

<i>EAMx</i> máx	161.2	129.5	120.2	116.0	116.1	122.8	100.5	115.4
<i>AIC</i> mín	311.5	302.2	302.2	313.8	309.7	300.0	306.4	293.6
<i>AIC</i> med	441.9	426.1	430.5	432.0	(421.8)	[463.5]	424.9	423.1
<i>AIC</i> máx	678.2	649.5	643.4	636.2	635.8	645.7	620.7	639.2
<i>COC</i> mín	0.916	0.944	0.935	0.945	0.953	0.914	0.938	0.951
<i>COC</i> med	0.971	0.976	0.972	0.975	0.978	[0.959]	0.976	(0.980)
<i>COC</i> máx	0.997	0.995	0.996	0.996	0.995	0.989	0.995	0.997
d_2 mín	0.889	0.937	0.946	0.951	0.951	0.943	0.963	0.944
d_2 med	[0.967]	0.979	0.981	0.982	0.983	0.974	(0.985)	0.980
d_2 máx	0.997	0.997	0.998	0.997	0.997	0.995	0.997	0.998
d_1 mín	0.839	0.874	0.879	0.890	0.888	0.864	0.864	0.886
d_1 med	0.913	0.927	0.921	0.925	0.929	[0.897]	0.918	(0.933)
d_1 máx	0.956	0.957	0.953	0.951	0.957	0.923	0.952	0.965

Concentrado por estaciones pluviométricas

En la Tabla 6 se han integrado los resultados de las últimas columnas de cada tabulación no expuesta del índice analizado, es decir, se tienen las mejores FDP en cada estación de acuerdo con cada índice. Cuando dos o más FDP mostraron igual valor para el índice analizado en una cierta estación, la mejor FDP se escogió con base en el mejor valor promedio (último renglón de cada tabulación no expuesta).

Tabla 6. Mejores FDP según índices de habilidad descriptiva en las nueve series de *PMD* anual procesadas del Altiplano Potosino del estado de San Luis Potosí, México.

Estación	Índices de habilidad descriptiva								Mejores dos FDP*
	<i>EEA</i>	<i>EREA</i>	<i>EAM</i>	<i>EAMx</i>	<i>AIC</i>	<i>COC</i>	d_2	d_1	
Cedral	LP3	WAK	WAK	LP3	LP3	WAK	LP3	WAK	WAK(4), LP3(4)
Charcas	WAK	WAK	WAK	GVE	WAK	GVE	WAK	WAK	WAK(6), GVE(2)
La Maroma	BEP	BEP	BEP	BEP	BEP	BEP	BEP	BEP	BEP(8)

Los Filtros	BEK	LGN	WAK	BEK	BEK	BEK	BEK	WAK	BEK(5), WAK(2)
Matehuala	BEK	WAK	WAK	BEK	BEK	BEK	BEK	WAK	BEK(5), WAK(3)
Mexquitic	WAK	LGN	WAK	WAK	WAK	BEK	LOG	WAK	WAK(5), BEK(1)
Peñón Blanco	LP3	WAK	GVE	LP3	LP3	LP3	LP3	GVE	LP3(5), GVE(2)
Santo Domingo	BEK	WAK	WAK	LP3	BEK	BEK	BEK	LOG	BEK(4), WAK(2)
Vanegas	PAG	LP3	LP3	PAG	PAG	PAG	PAG	LP3	PAG(5), LP3(3)
Regional	LOG	BEP	WAK	LP3	LOG	WAK	LP3	WAK	WAK(3), LP3(2)

*Entre paréntesis el número de veces que ocurre.

En la Tabla 6 se observa que en todas las columnas aparece la distribución Wakeby (WAK), con una ocurrencia en los índices EAM_x y COC , hasta seis en el índice EAM , y cinco en los índices $EREA$ y d_1 . El modelo probabilístico Wakeby es el mejor en 31.9% de los casos. Estos resultados orientan a definir la FDP Wakeby como un modelo que siempre se debe aplicar al procesar registros de PMD anual en climas áridos y semiáridos. En el último renglón de la Tabla 6, la segunda opción de la FDP regional pueden ser las distribuciones LOG y LP3 con dos ocurrencias cada una; se escoge la segunda por ser mejor en relación con dos índices no correlacionados (EAM_x y d_2).

Suprimiendo la distribución Wakeby de la Tabla 6, se busca la siguiente mejor y entonces se forma la Tabla 7, cuya columna final indica las dos mejores FDP y su número de ocurrencias en cada registro procesado. Como resumen de resultados de la Tabla 7 para el Altiplano Potosino, se puede indicar que las FDP Beta son la mejor opción en cuatro estaciones, después siguen los modelos LP3 y LOG en dos estaciones cada uno y, por último, la distribución Pareto generalizada es la mejor opción de una estación.

Tabla 7. Mejores FDP (excluyendo la Wakeby), según índices de habilidad descriptiva en las nueve series de PMD anual procesadas del Altiplano Potosino del estado de San Luis Potosí, México.

Estación:	Índices de habilidad descriptiva								Mejores dos FDP*
	EEA	$EREA$	EAM	EAM_x	AIC	COC	d_2	d_1	
Cedral	LP3	LOG	LOG	LP3	LP3	LP3	LP3	LOG	LP3(5), LOG(3)
Charcas	LGN	BEK	LOG	GVE	LGN	GVE	LP3	LOG	LOG(2), GVE(2)

La Maroma	BEP	BEP(8)							
Los Filtros	BEK	LGN	BEP	BEK	BEK	BEK	BEK	BEP	BEK(5), BEP(2)
Matehuala	BEK	GVE	LOG	BEK	BEK	BEK	BEK	BEP	BEK(5), LOG(1)
Mexquitic	LOG	LGN	LOG	LOG	LOG	BEK	LOG	LOG	LOG(6), BEK(1)
Peñón Bco.	LP3	BEP	GVE	LP3	LP3	LP3	LP3	GVE	LP3(5), GVE(2)
S. Domingo	BEK	LOG	BEP	LP3	BEK	BEK	BEK	LOG	BEK(4), LOG(2)
Vanegas	PAG	LP3	LP3	PAG	PAG	PAG	PAG	LP3	PAG(5), LP3(3)

*Entre paréntesis el número de veces que ocurre.

En la Figura 3 se muestra el gráfico *Q-Q* de la estación Mexquitic, cuyos valores estimados de *PMD* anual se obtuvieron con la FDP Wakeby. Este ajuste corresponde a un *EEA* de 1.45 milímetros, que fue el mínimo encontrado en las estaciones del Altiplano Potosino (Tabla 5).

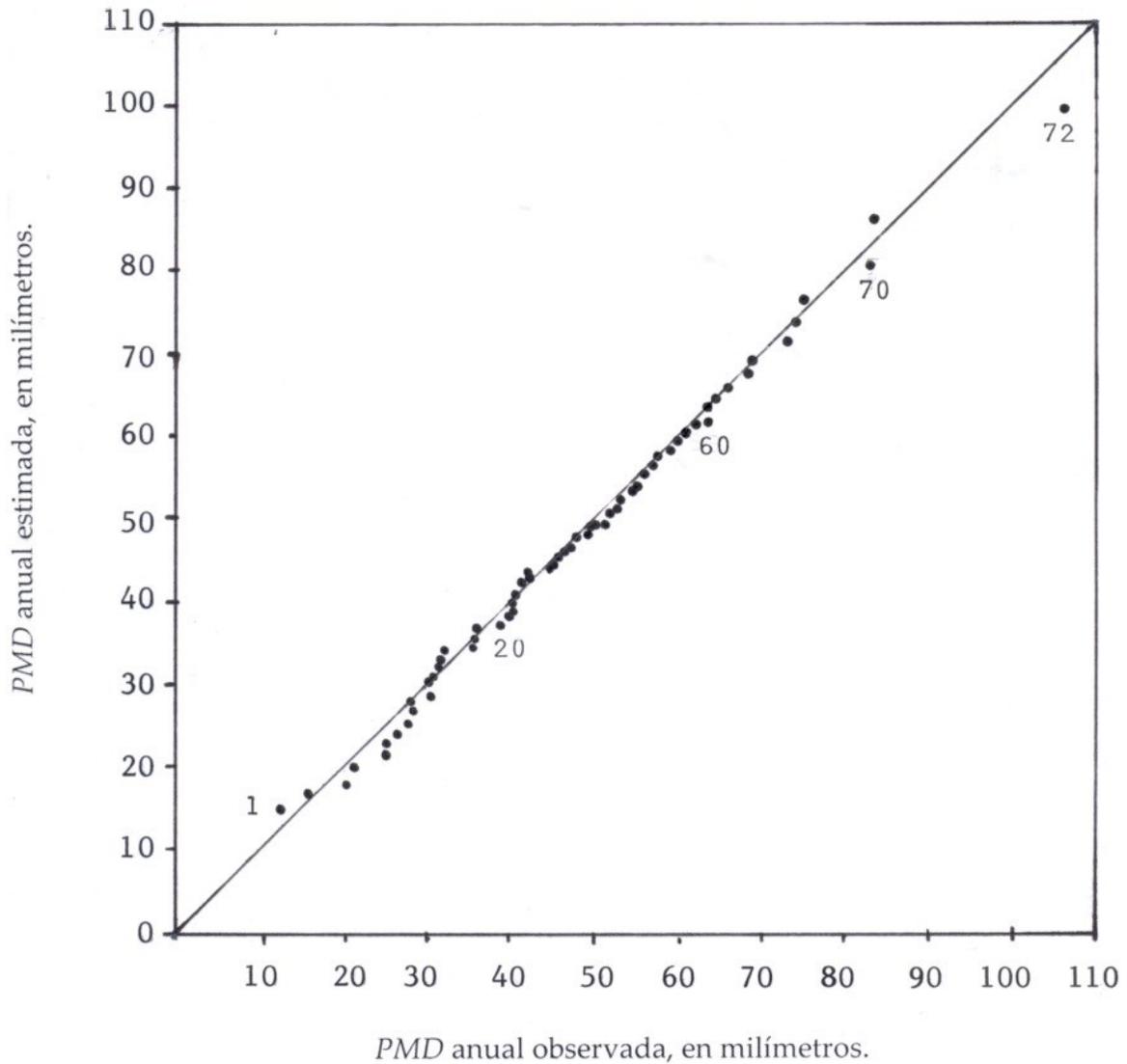


Figura 3. Gráfico Q-Q de la estación Mexquitic obtenido con la FDP Wakeby.

Resultados en la Zona Media

Concentrado de índices numéricos

En la Tabla 8 se han integrado los tres valores característicos (mínimo, medio y máximo) de cada índice de habilidad descriptiva (HD) obtenidos con cada una de las ocho FDP contrastadas en las 10 estaciones pluviométricas de la Zona Media del estado de San Luis Potosí, México. La Tabla 8 es similar a la Tabla 5.

Tabla 8. Valores característicos de los ocho índices de habilidad descriptiva (HD) en las 10 series de *PMD* anual procesadas de la Zona Media del estado de San Luis Potosí, México.

Índice de HD	FDP contrastadas							
	BEK	BEP	LGN	GVE	LOG	PAG	LP3	WAK
<i>EEA</i> mín	3.52	3.48	3.03	3.26	3.45	2.30	3.38	2.27
<i>EEA</i> med	[8.31]	8.19	6.46	6.41	6.57	7.24	(6.15)	7.47
<i>EEA</i> máx	15.52	20.78	14.40	13.34	13.00	15.64	12.37	22.68
<i>EREA</i> mín	0.043	0.043	0.045	0.048	0.044	0.034	0.045	0.0302
<i>EREA</i> med	0.068	0.070	(0.065)	0.063	0.065	0.086	0.063	[0.108]
<i>EREA</i> máx	0.093	0.125	0.098	0.091	0.091	0.161	0.092	0.571
<i>EAM</i> mín	2.361	2.382	2.112	2.312	2.424	1.768	2.225	1.653
<i>EAM</i> med	3.784	3.676	3.615	(3.483)	3.575	4.289	3.579	[5.064]
<i>EAM</i> máx	5.247	6.292	5.437	4.864	5.638	6.611	4.870	28.199
<i>EAMx</i> mín	8.8	8.4	8.7	7.7	8.3	6.1	10.6	5.8
<i>EAMx</i> med	[47.6]	46.5	32.1	31.8	32.6	35.4	29.3	(26.5)
<i>EAMx</i> máx	105.5	150.5	94.0	88.5	87.0	100.8	78.9	79.4
<i>AIC</i> mín	325.1	324.0	324.4	326.0	323.2	295.8	335.2	297.1
<i>AIC</i> med	[423.6]	417.4	396.9	396.9	400.3	406.8	(394.3)	401.9
<i>AIC</i> máx	579.3	624.6	511.6	529.3	543.8	484.8	518.0	514.4
<i>COC</i> mín	0.950	0.921	0.959	0.968	0.964	0.943	0.970	0.971
<i>COC</i> med	0.975	[0.974]	0.982	0.982	0.982	0.974	0.983	(0.987)
<i>COC</i> máx	0.991	0.991	0.992	0.991	0.991	0.995	0.991	0.992
<i>d₂</i> mín	0.964	0.944	0.971	0.975	0.976	0.965	0.979	0.800
<i>d₂</i> med	0.982	0.982	0.989	0.989	0.989	0.985	(0.990)	[0.973]
<i>d₂</i> máx	0.995	0.995	0.996	0.995	0.995	0.998	0.995	0.998

d_1 mín	0.886	0.894	0.896	0.895	0.893	0.881	0.901	0.453
d_1 med	0.920	0.922	0.924	(0.926)	0.923	0.911	0.924	[0.891]
d_1 máx	0.946	0.941	0.946	0.948	0.944	0.956	0.944	0.954

Concentrado por estaciones pluviométricas

En la Tabla 9 se indican para cada registro de *PMD* anual procesado cuál es la mejor FDP de acuerdo con cada índice de habilidad descriptiva. Se observa que la distribución Wakeby es la mejor opción en seis estaciones (Ojo de Agua a San Francisco); en total es la mejor FDP en 42.5% de los casos. Con base en los resultados de la Tabla 9 se concluye que el modelo Wakeby siempre se debe probar al analizar registros de *PMD* anual de climas cálido-subhúmedo. En el último renglón de la Tabla 9 se puede seleccionar como segunda mejor opción de FDP regional los modelos WAK y GVE; se escogió el primero debido a su mayor número de ocurrencias locales.

Tabla 9. Mejores FDP según cada índice de habilidad descriptiva en las 10 series de *PMD* anual procesadas de la Zona Media del estado de San Luis Potosí, México.

Estación	Índices de habilidad descriptiva								Mejores dos FDP*
	EEA	EREA	EAM	EAMx	AIC	COC	d_2	d_1	
Armadillo de los Infante	PAG	LGN	LGN	PAG	PAG	PAG	PAG	WAK	PAG(5), LGN(2)
Cárdenas	PAG	LGN	PAG	PAG	PAG	PAG	PAG	PAG	PAG(7), LGN(1)
Lagunillas	BEP	LGN	WAK	BEP	BEP	BEP	BEP	WAK	BEP(5), WAK(2)
Ojo de Agua	LP3	WAK	WAK	LP3	BEP	BEP	LP3	WAK	WAK(3), LP3(3)
Ojo de Agua Seco	WAK	WAK	WAK	BEP	GVE	GVE	BEK	WAK	WAK(4), GVE(2)
Paso de San Antonio	WAK	BEP	BEP	WAK	WAK	WAK	WAK	LOG	WAK(5), BEP(2)

Rayón	WAK	WAK	WAK	LP3	LP3	WAK	WAK	WAK	WAK(6), LP3(2)
Río Verde	WAK	WAK	WAK	PAG	PAG	WAK	PAG	WAK	WAK(5), PAG(3)
San Francisco	WAK	BEK	WAK	WAK	WAK	WAK	WAK	WAK	WAK(7), BEK(1)
San José Alburquerque	LP3	LP3	LP3	WAK	LP3	LP3	LP3	LP3	LP3(7), WAK(1)
Regional	LP3	LGN	GVE	WAK	LP3	WAK	LP3	GVE	LP3(3), WAK(2)

*Entre paréntesis el número de veces que ocurre.

Al eliminar a la distribución Wakeby de la Tabla 9 y buscar la siguiente mejor opción de FDP, se integra la Tabla 10, cuyos resultados para los registros de *PMD* anual de la Zona Media posicionan en primer término y por orden descendente a las distribuciones Pareto generalizada en tres estaciones; Log-Pearson tipo III, Beta Pareto y Logística Generalizada en dos estaciones. La FDP general de valores extremos es mejor opción en una estación.

Tabla 10. Mejores FDP (excluyendo la Wakeby), según cada índice de habilidad descriptiva en las 10 series de *PMD* anual procesadas de la Zona Media del estado de San Luis Potosí, México.

Estación	Índices de Habilidad Descriptiva								Mejores dos FDP*
	<i>EEA</i>	<i>EREA</i>	<i>EAM</i>	<i>EAMx</i>	<i>AIC</i>	<i>COC</i>	d_2	d_1	
Armadillo de los Infante	PAG	LGN	LGN	PAG	PAG	PAG	PAG	LGN	PAG(5), LGN(3)
Cárdenas	PAG	LGN	PAG	PAG	PAG	PAG	PAG	PAG	PAG(7), LGN(1)
Lagunillas	BEP	LGN	LP3	BEP	BEP	BEP	BEP	LP3	BEP(5), LP3(2)
Ojo de Agua	LP3	LOG	BEP	LP3	BEP	BEP	LP3	BEP	BEP(4), LP3(3)
Ojo de Agua Seco	GVE	BEK	GVE	BEP	GVE	GVE	BEK	GVE	GVE(5), BEK(2)
Paso de San Antonio	LOG	BEP	BEP	LOG	BEP	LOG	LOG	LOG	LOG(5), BEP(3)

Rayón	LP3	LOG	BEP	LP3	LP3	LOG	LP3	LOG	LP3(4), LOG(3)
Río Verde	PAG	PAG(8)							
San Francisco	LOG	BEK	LOG	GVE	LOG	LOG	GVE	LOG	LOG(5), GVE(2)
San José Albuquerque	LP3	LP3(8)							

*Entre paréntesis el número de veces que ocurre.

Resultados según habilidad predictiva

FDP aplicadas

Para cada estación climatológica o registro de *PMD* anual se escogieron cuatro FDP por contrastar. La primera corresponde a la mejor opción de la Tabla 3, es decir, es la FDP más adecuada según resultados del diagrama de cocientes L. Las siguientes dos FDP por aplicar fueron las obtenidas como mejores opciones de acuerdo con los ocho índices de habilidad descriptiva, las cuales se concentraron en la Tabla 7 y Tabla 10. Por último, se aplicó la FDP Wakeby debido a su gran capacidad descriptiva, que fue mostrada en la Tabla 6 y Tabla 9; por ello, se sugiere de aplicación bajo precepto. En la Tabla 11 y Tabla 12 de predicciones calculadas y adoptadas se han destacado con negritas las siguientes tres estaciones: La Maroma, Río Verde y San José Albuquerque, debido a que en ellas sólo se contrastan tres FDP, pues la Tabla 7 y Tabla 10 reportan sólo una mejor FDP en los ocho índices de habilidad descriptiva aplicados.

Cuando alguna de las dos mejores FDP de la Tabla 7 o Tabla 10 coincidía con la primera distribución aplicada, se cambió esta última por su segunda y/o tercera opción de la Tabla 3. La opción que tiene la primera FDP aplicada se indicó entre paréntesis rectangular en la Tabla 11 y Tabla 12 de *predicciones* calculadas y seleccionadas. Para las dos

mejores FDP de la Tabla 7 y Tabla 10 se indica en paréntesis circular el número de índices de habilidad descriptiva en que son las mejores. Lo mismo se indica para la distribución Wakeby, pero tal dato procede de la Tabla 6 y Tabla 9.

Tabla 11. Predicciones de cuatro periodos de retorno obtenidas con las FDP indicadas en cada una de las nueve series de *PMD* anual del Altiplano Potosino del estado de San Luis Potosí, México (se indican entre paréntesis las predicciones adoptadas).

Estación mejor FDP	P_M (P_M/P_{50})	Periodos de retorno en años			
		50	100	500	1 000
Cedral	315.8				
GVE [2]	2.21	143	192	384	520
LP3 (5)	2.10	(151)	(206)	(431)	(596)
LOG (3)	2.24	141	191	400	555
WAK (4)	2.32	136	191	443	2326
Charcas	117.0				
LGN [3]	1.10	106	119	147	202
LOG (2)	1.07	(109)	(126)	(174)	(199)
GVE (2)	1.07	109	121	150	162
WAK (6)	1.07	109	122	150	198
La Maroma	140.1				
LOG [1]	1.30	108	129	197	236
BEP (8)	1.25	(112)	(136)	(213)	(258)
WAK (0)	1.29	109	126	172	279
Los Filtros	111.0				
LP3 [1]	1.32	84	93	114	123
BEK (5)	1.32	(84)	(95)	(127)	(143)
BEP (2)	1.35	82	91	117	129
WAK (2)	1.34	83	93	119	176
Matehuala	200.0				
LP3 [2]	1.39	144	170	245	285
BEK (5)	1.29	(155)	(191)	(309)	(380)
LOG (1)	1.41	142	171	266	323

WAK (3)	1.48	135	170	316	1211
Mexquitic	107.0				
LGN [2]	1.24	86	91	104	124
LOG (6)	1.22	88	97	117	127
BEK (1)	1.32	81	89	111	122
WAK (5)	1.20	(89)	(98)	(120)	(163)
Peñón Blanco	235.0				
LOG [1]	1.51	156	216	473	667
LP3 (5)	1.39	(169)	(233)	(490)	(676)
GVE (2)	1.49	158	217	457	631
WAK (1)	1.51	156	221	510	2521
Santo Domingo	270.0				
LP3 [1]	1.61	168	201	293	340
BEK (4)	1.70	(159)	(198)	(327)	(406)
LOG (2)	1.72	157	195	322	400
WAK (2)	2.03	133	175	400	2770
Vanegas	90.0				
LGN [1]	1.00	(90)	(102)	(132)	(197)
PAG (5)	1.06	85	92	103	107
LP3 (3)	1.01	89	101	130	143
WAK (0)	1.01	89	99	118	146

Tabla 12. Predicciones de cuatro periodos de retorno obtenidas con las FDP indicadas en cada una de las 10 series de *PMD* anual de la Zona Media del estado de San Luis Potosí, México (se indican entre paréntesis las predicciones adoptadas).

Estación mejor FDP	P_M (P_M/P_{50})	Periodos de retorno en años			
		50	100	500	1 000
Armadillo de los Infante	133.0				
LP3 [2]	0.95	140	163	226	258
PAG (5)	0.99	135	150	180	191
LGN (3)	0.96	(139)	(162)	(220)	(354)
WAK (1)	0.96	138	156	196	267
Cárdenas	180.5				

LP3 [1]	0.95	(191)	(231)	(345)	(406)
PAG (7)	0.97	186	215	282	311
LGN (1)	0.95	191	228	330	586
WAK (0)	0.97	187	216	285	418
Lagunillas	210.0				
LGN [1]	1.21	173	196	251	368
BEP (5)	1.16	(181)	(214)	(317)	(375)
LP3 (2)	1.19	176	201	265	295
WAK (2)	1.22	172	200	289	579
Ojo de Agua	300.2				
LOG [1]	1.31	(229)	(290)	(510)	(656)
BEP (4)	1.29	233	293	498	625
LP3 (3)	1.27	236	292	474	583
WAK (3)	1.29	233	298	533	1623
Ojo de Agua Seco	172.5				
LOG [2]	1.11	(155)	(185)	(283)	(341)
GVE (5)	1.12	154	180	252	289
BEK (2)	1.11	155	185	274	325
WAK (4)	1.11	155	178	234	356
Paso de S. Antonio	200.0				
GVE [2]	1.41	142	161	208	230
LOG (5)	1.39	(144)	(167)	(238)	(277)
BEP (3)	1.42	141	163	228	263
WAK (5)	1.36	147	172	246	463
Rayón	330.0				
GVE [2]	1.63	203	254	427	533
LP3 (4)	1.57	210	265	447	558
LOG (3)	1.63	(202)	(257)	(461)	(596)
WAK (6)	1.63	203	267	515	1834
Río Verde	126.3				
LP3 [3]	1.00	(126)	(142)	(184)	(204)
PAG (8)	1.07	118	125	137	141
WAK (5)	1.07	118	126	140	153

San Francisco	135.0				
LGN [3]	1.18	114	131	172	260
LOG (5)	1.15	117	139	208	247
GVE (2)	1.17	115	134	181	204
WAK (7)	1.13	(119)	(141)	(198)	(342)
San J. Alburquerque	126.5				
GVE [1]	1.17	108	121	153	167
LP3 (8)	1.13	(112)	(127)	(167)	(186)
WAK (1)	1.00	127	137	156	182

Estadísticas obtenidas

En la Tabla 1 se observa que la mayoría de los registros procesados de *PMD* anual tienen una amplitud de 50 años o más, debido a lo cual se calculó el cociente entre el valor máximo del registro (P_M) y la predicción de un periodo de retorno de 50 años (P_{50}). Este cociente se indica en las columnas 2 de la Tabla 11 y Tabla 12; cuando es cercano a la unidad indica que el registro no tiene valores dispersos extremos (*outliers*) que se apartan de la tendencia natural de los datos. En cambio, cuando excede a 1.50 existe presencia de uno o más valores dispersos, caso de las cuatro estaciones: Cedral, Peñón Blanco, Santo Domingo y Rayón.

En las cuatro estaciones citadas, la distribución Wakeby, debido a la extraordinaria flexibilidad que le otorgan sus cinco parámetros de ajuste, conduce a predicciones muy elevadas en el periodo de retorno de mil años, como se observa al compararlas con las obtenidas con las otras FDP contrastadas. En ninguno de los casos citados se adoptó la FDP Wakeby, por considerarse exageradas sus predicciones, al no coincidir con las de los otros tres modelos probabilísticos contrastados en tal estación. Nguyen *et al.* (2017) también encuentran que la distribución Wakeby tiene baja habilidad predictiva, al mostrar gran variabilidad en sus predicciones.

En la Tabla 11 y Tabla 12 de predicciones calculadas y adoptadas en las estaciones del Altiplano Potosino y de la Zona Media se están tomando

en cuenta de manera implícita las habilidades descriptivas y predictivas de cada una de las FDP contrastadas; por ello, las conclusiones siguientes se consideran globales del estudio.

Se obtuvo que en 12 estaciones los valores adoptados proceden de las dos FDP que fueron mejor opción de acuerdo con los ocho índices de habilidad descriptiva. En cinco estaciones, las predicciones adoptadas fueron calculadas con las FDP mejores opciones según el diagrama de cocientes L , y sólo en dos estaciones se adoptaron las predicciones calculadas con la distribución Wakeby.

Como ya se indicó, exclusivamente en tres estaciones: La Maroma, Río Verde y San José Alburquerque se obtuvo una concordancia total en los ocho índices de habilidad descriptiva para las FDP Beta- P , Pareto generalizada y Log-Pearson tipo III, respectivamente. Dichas estaciones se han destacado con negritas tanto en la Tabla 11 como en la Tabla 12.

Por áreas geográficas, en el Altiplano Potosino de nueve registros procesados, la FDP Beta- κ fue el modelo adoptado en tres estaciones y la distribución Log-Pearson tipo III en dos estaciones. En la Zona Media de 10 registros procesados, tuvo preponderancia de adopción la FDP Logística Generalizada con cuatro estaciones y le siguió la distribución LP3 con tres estaciones.

Conclusiones

La distribución Wakeby, ajustada con el método de momentos L , es un modelo de excelente habilidad descriptiva, y por ello *se sugiere* que sea aplicado bajo precepto en los análisis probabilísticos de registros de *PMD* anual de los climas áridos y semiáridos del Altiplano Potosino (AP) y del clima cálido-subhúmedo de la Zona Media (ZM) del estado de San Luis Potosí, México.

Las distribuciones Beta- κ y Beta- P , ajustadas con el método de máxima verosimilitud, son modelos no aplicados en México *que se sugiere sean probados*, pues para cuatro registros de *PMD* anual del AP (Tabla 7) y

dos de la ZM (Tabla 10) conducen a los mejores índices de habilidad descriptiva.

En relación con las distribuciones que se aplican bajo precepto en EUA e Inglaterra, se obtuvo (Tabla 7 y Tabla 10) lo siguiente: (1) la FDP Log-Pearson tipo III resultó ser la mejor opción en dos estaciones del AP y ZM; (2) la FDP General de Valores Extremos sólo en una estación de la ZM fue mejor opción; (3) la FDP Logística Generalizada fue mejor opción en dos estaciones del AP y de la ZM; (4) en el AP destaca como segunda mejor opción la Logística Generalizada, y en la ZM los modelos Log-Normal y Log-Pearson tipo III.

Respecto a la FDP Pareto Generalizada, que por lo común se aplica junto con los modelos LOG y GVE, fue mejor opción en una estación del AP y tres de la ZM. Estos resultados ratifican la aplicación sistemática o bajo precepto de las distribuciones LP3, GVE, LOG y PAG en series de *PMD* anual de climas áridos, semiáridos y cálido-subhúmedos.

Respecto a las *predicciones calculadas* (Tabla 11 y Tabla 12) en los periodos de retorno de 50, 100, 500 y 1000 años, en general muestran valores similares y ello genera confianza en los valores adoptados. Exclusivamente se encontró dispersión en las predicciones de la FDP Wakeby en las estaciones o registros de *PMD* anual con valor extremo disperso (*outlier*), como es el caso de las estaciones: Cedral, Peñón Blanco, Santo Domingo y Rayón.

En relación con las *predicciones adoptadas* (Tabla 11 y Tabla 12), se concluye que los procedimientos de búsqueda de la mejor FDP por aplicar a los registros de *PMD* anual, basados en el diagrama de cocientes L y en los ocho índices de habilidad descriptiva, son adecuados y conducen a una buena aproximación, pues no existió dificultad para seleccionar las predicciones adoptadas.

Referencias

- Asquith, W. H. (2011). L-moments. Chapter 6. In: *Distributional analysis with L-moments statistics using the R environment for statistical computing* (pp. 87-122). Lubbock, USA: United States Geological Services.
- Benson, M. A. (1962). Plotting positions and economics of engineering planning. *Journal of Hydraulics Division*, 88(6), 57-71.

- Bobée, B., & Ashkar, F. (1991). *The Gamma Family and derived distributions applied in Hydrology*. Littleton, USA: Water Resources Publications.
- 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. (2015). Ajuste de las distribuciones GVE, LOG y PAG con momentos L depurados (1,0). *Tecnología y ciencias del agua*, 6(4), 153-167.
- Campos-Aranda, D. F. (2016). Ajuste de las distribuciones GVE, LOG y PAG con momentos L de orden mayor. *Ingeniería, Investigación y Tecnología*, 17(1), 131-142.
- Campos-Aranda, D. F. (2018). Estimación estadística actualizada de la PMP en el estado de San Luis Potosí, México. *Tecnología y ciencias del agua*, 9(6), 32-70.
- Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7, 1247-1250.
- Coles, S. (2001). Model diagnostics. Theme 2.6.7. In: *An introduction to statistical modeling of extreme values* (pp. 36-44). London, England: Springer-Verlag.
- Cunnane, C. (1978). Unbiased plotting positions. A review. *Journal of Hydrology*, 37(3-4), 205-222.
- Haktanir, T. (1991). Statistical modeling of annual maximum flows in Turkish rivers. *Hydrological Sciences Journal*, 36(4), 367-389.
- Haktanir, T. (1992). Comparison of various flood frequency distributions using annual flood peaks of rivers in Anatolia. *Journal of Hydrology*, 136(1-4), 1-31.
- Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis. An approach based on L-moments*. Cambridge, England: Cambridge University Press.
- Houghton, J. C. (1978). Birth of a parent: The Wakeby distribution for modeling flood flows. *Water Resources Research*, 14(6), 1105-1109.

- Johnson, F., & Sharma, A. (2017). Design rainfall. Chapter 125. In: Singh, V. P. (ed.). *Handbook of applied hydrology*, 2nd ed. (pp. 125.1-125.13). New York, USA: McGraw-Hill Education.
- 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.
- Legates, D. R., & McCabe, G. J. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research*, 35(1), 233-241.
- Machiwal, D., & Jha, M. K. (2008). Comparative evaluation of statistical tests for time series analysis: Applications to hydrological time series. *Hydrological Sciences Journal*, 53(2), 353-366.
- Mujumdar, P. P., & Nagesh-Kumar, D. N. (2012). *Floods in a changing climate. Hydrologic modeling*. Cambridge, United Kingdom: International Hydrology Series (UNESCO) and Cambridge University Press.
- Meylan, P., Favre, A. C., & Musy, A. (2012). *Predictive hydrology. A frequency analysis approach*. Boca Raton, USA: CRC Press.
- 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.
- 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.
- Pandey, G. R., & Nguyen, V. T. V. (1999). A comparative study of regression based methods in regional flood frequency analysis. *Journal of Hydrology*, 225(1-2), 92-101.
- Peel, M. C., Wang, Q. J., Vogel, R., & McMahon, T. A. (2001). The utility of L-moment ratio diagrams for selecting a regional probability distribution. *Hydrological Sciences Journal*, 46(1), 147-155.
- Rao, A. R., & Hamed, K. H. (2000). Probability weighted moments and L-moments. Chapter 3 (pp. 53-72). In: *Flood frequency analysis*. Boca Raton, USA: CRC Press.

-
- Stedinger, J. R. (2017). Flood frequency analysis. Chapter 76. In: Singh, V. P. (ed.). *Handbook of applied hydrology*, 2nd ed. (pp. 76.1-76.8). New York, USA: McGraw-Hill Education.
- Teegavarapu, R. S. V. (2012). *Floods in a changing climate. Extreme precipitation*. Cambridge, United Kingdom: International Hydrology Series (UNESCO) and Cambridge University Press.
- 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.
- 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.
- Wilks, D. S. (2011). *Statistical Methods in the Atmospheric Sciences*. Theme 4.5: Qualitative assessments of the goodness fit, pp. 112-116. San Diego, USA: Academic Press (Elsevier). Third edition. 676 p.
- WMO, World Meteorological Organization. (1971). *Climatic change*. Annexed III (pp. 58-71) (Technical Note No. 79). Geneva, Switzerland: Secretariat of the World Meteorological Organization.
- WRC, Water Resources Council. (1977). *Guidelines for determining flood flow frequency*. Bulletin # 17A of the Hydrology Committee (Revised edition). Washington, DC, USA: Water Resources Council.
- Yue, S., & Hashino, M. (2007). Probability distribution of annual, seasonal and monthly precipitation in Japan. *Hydrological Sciences Journal*, 52(5), 863-877.
- Zalina, M. D., Desa, M. N. M., Nguyen, V. T. V., & Kassim, A. H. M. (2002). Selecting a probability distribution for extreme rainfall series in Malaysia. *Water Science and Technology Journal*, 45(3), 63-68.