

Análisis probabilístico de avenidas máximas con poblaciones mezcladas

Daniel Francisco Campos Aranda

Coordinación Regional de Infraestructura Hidráulica del Norte, SARH

En la República Mexicana es común que los registros hidrométricos de avenidas máximas anuales de cuencas con una superficie superior a los 1 000 km², muestren la presencia de dos poblaciones, ya sea porque el valor de su coeficiente de asimetría es demasiado alto o muy pequeño, o bien porque al representar sus valores en un papel de probabilidad log-normal o Gumbel, se obtiene una curva marcadamente ascendente. Esto último se debe a que unas avenidas se originaron durante la primavera y el verano, y otras durante el otoño; es decir, que unas se derivan de eventos no ciclónicos y otras de eventos asociados con la ocurrencia de ciclones, los cuales en general producen las mayores crecientes. El procedimiento de análisis hidrológico de ese tipo de muestras consiste en probar estadísticamente la presencia de dos poblaciones y, en caso de obtenerse un resultado positivo, continuar con la aplicación de criterios específicos para el procesamiento probabilístico, como son la segregación y el análisis de las poblaciones, o la construcción de curvas de frecuencia combinada. El objetivo de este trabajo consiste en describir los procedimientos de análisis probabilístico de muestras con dos poblaciones mezcladas y aplicarlos a un caso real.

En ciertas localidades, el análisis probabilístico de los registros de avenidas máximas anuales puede requerir de un tratamiento especial, cuando éstas son causadas por diferentes tipos de fenómenos hidrológicos y/o cuando al representarse en un papel de probabilidad log-normal señalan un cambio de curvatura pronunciado (Morris, 1982).

Los registros hidrométricos de avenidas de las áreas costeras son ejemplos de muestras con poblaciones mezcladas, pues algunas de ellas fueron originadas por los ciclones y otras por las tormentas convectivas, que caen principalmente durante la primavera y el verano. La ocurrencia de estas avenidas da lugar a dos estaciones y, según indican Browzin *et al* (1972), las condiciones específicas del terreno de la cuenca en cada época contribuyen también a modificar sus características. En consecuencia, la serie anual de avenidas máximas (muestra) puede no ser homogénea, y

en caso de que se pruebe estadísticamente la existencia de dos poblaciones, la muestra debe segregarse para analizar probabilísticamente cada población por separado. Lo anterior constituye el procedimiento de análisis propuesto por Browzin *et al*, mediante el cual, en la mayor parte de los casos estudiados, las series segregadas conducen a estimaciones más amplias de avenidas asociadas a un periodo de retorno.

Morris (1982), en cambio, propone obtener una curva de frecuencias para cada población segregada e integrarlas en una *curva de frecuencia de poblaciones combinadas*.

Kite (1977) hace referencia al enfoque propuesto por Stoddart y Watt en 1970 para analizar las avenidas de las cuencas del sur de Ontario, Canadá, las cuales tienen dos tipos de crecientes: aquéllas que se deben nada más a la precipitación, que por lo general ocurren en el verano, y las derivadas de la fusión de nieve, que se presen-

tan en el invierno o en la primavera. Este enfoque combina los dos anteriores, por lo cual también se expone en este artículo.

Segregación de una serie de avenidas máximas anuales

En ocasiones, un coeficiente de asimetría demasiado grande o muy pequeño puede indicar la presencia de poblaciones mezcladas; se consideran coeficientes de asimetría anormales los mayores de 0.70 y los menores de -0.40 (Morris, 1982). Por otra parte, un cambio acentuado en la curva de frecuencias lo causan en general varios eventos máximos que se apartan en forma significativa de la tendencia que sigue el resto de los datos. Esta clase de eventos la producen frecuentemente diferentes tipos de fenómenos hidrológicos, como ciclones en un registro de avenidas originadas por lluvia, precipitaciones en una serie de avenidas de deshielo o tormentas convectivas en un registro de avenidas asociadas a tormentas invernales.

El objetivo primario de un análisis por cada población es mejorar el ajuste entre una distribución de probabilidades derivada analíticamente y los valores representados en papel de probabilidad. Al respecto, el *US Water Resources Council* (1977) ha establecido un procedimiento para ponderar los grandes eventos con un periodo de retorno mayor que la amplitud del registro, basado en la extensión (histórica) del periodo de observaciones. Sin embargo, algunas veces el ajuste histórico no es muy confiable, o bien no existe información suficiente para llevarlo a cabo, por lo cual se debe recurrir a segregar la serie y obtener una curva de frecuencias de poblaciones combinadas, o aplicar otros enfoques.

Un factor esencial en el análisis de avenidas asociadas a precipitaciones ciclónicas y no ciclónicas es el tamaño de la cuenca. En una cuenca pequeña, la lluvia originada por un fenómeno no ciclónico frecuentemente iguala la magnitud de las avenidas derivadas de precipitaciones ciclónicas y, conforme el tamaño de la cuenca aumenta, decrece la probabilidad de que la magnitud de un evento ocasionado por precipitación no ciclónica iguale a la de uno causado por lluvia ciclónica. El Centro de Ingeniería Hidrológica del *US Army Corps of Engineers* ha encontrado que en cuencas menores de 1 300 km² (500 millas cuadradas) no se requiere de un análisis especial de los eventos ciclónicos, lo cual no ocurre cuando se segre-

gan eventos asociados con lluvias y con fusión de nieve (Morris, 1982).

Conviene señalar que la segregación de un registro tiene dos desventajas básicas; la primera consiste en la dificultad de identificar los eventos de cada población, y la segunda, en el reducido número de avenidas asociadas a una de las poblaciones.

Pruebas estadísticas de homogeneidad

Existen varias pruebas estadísticas para corroborar la presencia de eventos procedentes de poblaciones diferentes. El procedimiento general para dichas pruebas es el siguiente (Kite, 1977):

- Se segrega la muestra para determinar las dos series de datos por contrastar.
- Se establece la hipótesis nula (H_0): no existe diferencia entre las dos series de eventos.
- Se adopta un nivel de significancia (α) para la prueba, comúnmente un 5%.
- Se selecciona una prueba estadística y se aplica para aceptar o rechazar H_0 .

Para llevar a cabo el paso uno se pueden seguir dos criterios. En el primero se identifican los eventos asociados con ciclones para integrar la población ciclónica y con los restantes queda formada la población no ciclónica; en el segundo se divide el año en estaciones o épocas.

El *US Water Resources Council* (1977) señala que *la separación por periodos del calendario, en vez de separar por eventos, no es considerada hidrológicamente razonable, a menos que los eventos en los periodos separados sean claramente causados por diferentes condiciones hidrometeorológicas*. Además, subraya que, si la separación no obedece a un criterio objetivo y con total fundamento hidrológico, el registro debe de ser tratado como procedente de una sola población. Así, por ejemplo, Browzin *et al* (1972) analizan las avenidas ocurridas en varios ríos del Estado de Texas, EUA, y para tomar en cuenta las distintas condiciones de suelos en las cuencas y el origen diferente de las precipitaciones, establecen dos épocas de avenidas (la de primavera, que comprende de abril a junio, y la de otoño que va de septiembre a noviembre).

Por otra parte, existen dos tipos de pruebas estadísticas: las paramétricas y las no paramétricas. Las primeras se relacionan con ciertas hipótesis a las que debe apegarse la muestra, las cua-

les no se cumplen en el caso de un registro de avenidas. Por ello se recurre a pruebas no paramétricas, como las de Mann-Whitney, Kruskal-Wallis, Terry y Kolmogorov-Smirnov.

Prueba de Mann-Whitney

Esta, también conocida como Prueba U, es una de las más confiables. La muestra total, de tamaño N , se ordena en forma progresiva de acuerdo con la magnitud de los datos, y se evalúa (R_1), que es la suma del número (rango) asignado a cada evento de la serie de tamaño n_1 . Por ejemplo, la serie puede ser no ciclónica, con lo cual el tamaño de la serie ciclónica será: $n_2=N-n_1$; las estadísticas se evalúan con base en las siguientes fórmulas (Kite, 1977):

$$U_1 = n_1 n_2 + \frac{n_1}{2}(n_1 + 1) - R_1 \quad (1)$$

$$u_2 = n_1 n_2 - u_1 \quad (2)$$

El menor valor de U_1 ó U_2 se denomina U y, en función de él, se calcula la estadística Z , como sigue:

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{1/12[n_1 n_2 (N+1)]}} \quad (3)$$

El valor absoluto de Z se compara con 1.96; si Z es mayor, se rechaza H_0 , ya que los datos provienen probablemente de dos poblaciones.

Con objeto de comprobar el cálculo de U , se evalúa:

$$u_2 = n_1 n_2 + \frac{n_2}{2}(n_2 + 1) - R_2 \quad (4)$$

El valor de U_2 debe ser igual a U , cuando el valor de R_2 se determina en el cuadro 1 de la manera siguiente: $R_2=R_N-R_1$.

Prueba de Kruskal-Wallis

Esta prueba es sumamente útil para verificar si k muestras independientes proceden de poblaciones diferentes. La estadística H se calcula por medio de la siguiente expresión (Siegel, 1976):

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1) \quad (5)$$

Para el caso de avenidas, $k=2$ y R_i es la suma del número (rango) que corresponde a cada serie de tamaño n_i , por lo cual $N = n_1 + n_2$. Cuando $n_i > 5$, H se distribuye según una X^2 con $(i-1)$ grados de libertad, por lo que el valor crítico de H es $X_c^2(1, \alpha) = 3.841$ y, si $H > 3.841$, se rechaza H_0 .

Prueba de Terry

En esta prueba los datos de la muestra se ordenan en forma ascendente y, para cada evento de la población no ciclónica de tamaño n_1 , se obtiene el i -ésimo valor esperado del estadístico de orden $E(Z_{Ni})$ conforme a las tablas de Harter (1961) para población normal estandarizada. Así, por ejemplo, en el caso de una muestra de tamaño $N = n_1 + n_2 = 55$, el valor para el rango 2 es 1.8972, para el rango 16 es 0.5749 y para el rango 39 es -0.5222.

La suma de los valores esperados $E(Z_{Ni})$ para la serie no ciclónica estará representada por c , y la estadística de la prueba será T (Kite, 1977), que se expresa como:

$$T = c/S \quad (6)$$

En la expresión anterior, S está definida por:

$$S = \sqrt{\frac{n_1 n_2}{N(N-1)} \sum_{i=1}^N [e(Z_N)]^2} \quad (7)$$

Si T es mayor de 1.96, se rechaza H_0 , y en consecuencia, la muestra procede de dos poblaciones.

Prueba de Kolmogorov-Smirnov

Consiste en representar gráficamente las distribuciones de probabilidad acumulada de las dos series y evaluar la máxima desviación entre ellas (D). Para los casos en que n_1 ó $n_2 \leq 40$ y $n_1 \neq n_2$, la estadística X^2 se compara con un valor correspondiente a dos grados de libertad [$X_c^2(2, \alpha) = 5.99$], es decir:

$$X^2 = 4D \frac{n_1 n_2}{n_1 + n_2} \quad (8)$$

Si $X^2 > 5.99$, se rechaza H_0 . Una explicación más amplia sobre esta prueba se puede encontrar en Siegel (1976).

Curva de frecuencias de poblaciones combinadas

El procedimiento siguiente, descrito por Morris (1982), es bastante lógico, objetivo (cuando se

1. Tabla auxiliar para el cálculo de la suma de rangos^(R)

Núm. de orden	Suma						
1	1	26	351	51	1326	76	2926
2	3	27	378	52	1378	77	3003
3	6	28	406	53	1431	78	3081
4	10	29	435	54	1485	79	3160
5	15	30	465	55	1540	80	3240
6	21	31	496	56	1596	81	3321
7	28	32	528	57	1653	82	3403
8	36	33	561	58	1711	83	3486
9	45	34	595	59	1770	84	3570
10	55	35	630	60	1830	85	3655
11	66	36	666	61	1891	86	3741
12	78	37	703	62	1953	87	3828
13	91	38	741	63	2016	88	3916
14	105	39	780	64	2080	89	4005
15	120	40	820	65	2145	90	4095
16	136	41	861	66	2211	91	4186
17	153	42	903	67	2278	92	4278
18	171	43	946	68	2346	93	4371
19	190	44	990	69	2415	94	4465
20	210	45	1035	70	2485	95	4560
21	231	46	1081	71	2556	96	4656
22	253	47	1128	72	2628	97	4753
23	276	48	1176	73	2701	98	4851
24	300	49	1225	74	2775	99	4950
25	325	50	1275	75	2850	100	5050

dispone de la información necesaria) y sencillo; consiste en los pasos que se describen a continuación:

- Se identifican los eventos de la muestra que están asociados con la ocurrencia de ciclones. Esta es la parte más difícil del procedimiento, ya que la extensión que abarca un ciclón no se conoce con exactitud, por lo que únicamente se puede recurrir a las trayectorias ciclónicas y a la información obtenida de los moradores de la zona. Debido a la naturaleza diferente de las dos series de eventos, no es necesario probar estadísticamente que corresponda a dos poblaciones.
- Se ajusta un modelo o función de distribución de probabilidades a la serie de eventos no ciclónicos, la cual es en general la que comprende el mayor número de eventos, por lo que puede ser tratada como una serie común. En cambio, la serie ciclónica está integrada por pocos eventos que, además, no ocurren cada año, por lo cual constituyen una serie discontinua; debido a ello se requiere de un procedimiento especial para procesarla estadísticamente, co-

mo es el ajuste de una ecuación de regresión modificada. Primero, se obtiene la posición gráfica de cada evento ciclónico mediante la fórmula de Weibull:

$$P(X > x) = P(x) = m/N \tag{9}$$

En la expresión anterior, m es el número asignado al evento, en un orden del mayor al menor, y N es el total de datos de la muestra ($N = n_1 + n_2$). A continuación, se estima la desviación normal (K) que corresponde a $P(x)$, por medio de una tabla de distribución normal estándar, o bien de la aproximación siguiente (Abramowitz y Stegun, 1970):

$$K = t - \frac{2.30753 + 0.27061t}{1 + 0.99229t + 0.04481t^2} \tag{10}$$

$$t = \sqrt{\text{Ln}\{1/[p(x)^2]\}} \tag{11}$$

La ecuación de la curva de regresión modificada es la siguiente (Morris, 1982):

$$X = \bar{X} + S \cdot K \quad (12)$$

En la ecuación anterior:

$$\bar{X} = (\sum X_i/n_2) - S(\sum K_i/n_2) \quad (13)$$

$$X_i = \log Q \quad (14)$$

$$S = S_x/S_k \quad (15)$$

En dichas expresiones, Q es el gasto del evento ciclónico, K_i, el valor calculado mediante la ecuación (10) correspondiente a P(x), n₂, el número de eventos ciclónicos, y S_x y S_k, las desviaciones estándar de los valores X_i y K_i, respectivamente.

- Se combinan las dos curvas de frecuencias por medio de la siguiente expresión (US Army Corps of Engineers, 1975):

$$P_c = P_1 + P_2 - P_1 \cdot P_2 \quad (16)$$

Donde:

- P_c probabilidad de excedencia de la curva de frecuencias correspondiente a poblaciones combinadas para un gasto de avenida seleccionado.
- P₁ probabilidad de excedencia asociada a un gasto de avenida seleccionado, obtenida de la curva de frecuencias de la serie no ciclónica.
- P₂ probabilidad de excedencia asociada a un gasto de avenida seleccionado, obtenida de la curva de regresión modificada de la serie ciclónica.

Segregación en épocas

Cuando no es posible identificar los eventos derivados de ciclones, conviene establecer dos épocas de ocurrencia de avenidas (en la República Mexicana podrían corresponder a cada semestre del año), o bien definir las épocas que se consideren más convenientes desde el punto de vista hidrológico. El procedimiento completo se compone de los pasos siguientes.

- Mediante el uso de la serie anual de avenidas máximas, se comprueba estadísticamente la

presencia de dos poblaciones, con base en las pruebas estadísticas de homogeneidad especificadas en este artículo.

- En caso de que se detecten dos poblaciones (una para cada época definida), se integra a cada una la muestra de avenidas máximas anuales correspondiente, en la cual se sustituyen los eventos excluidos de una época por los mayores del año estudiado, a condición de que sean de su periodo o época (véase cuadro 2).
- Se procesa probabilísticamente cada serie, para lo cual se ajusta una función de distribución de probabilidades.
- En el procedimiento propuesto por Browzin *et al* (1972), sólo se consideran los mayores valores obtenidos en cada una de las tres muestras (la de primavera, la de otoño y la original). En cambio, en el procedimiento sugerido por Stoddart y Watt, se obtiene una curva de frecuencias anual asintótica a las otras dos. Estos autores determinan el periodo de retorno (T_A) de una avenida máxima anual seleccionada arbitrariamente como sigue (Kite, 1977):

$$T_A = \frac{T_{NC} \cdot T_C}{(T_{NC} + T_C - 1)} \quad (17)$$

En la expresión anterior, T_{NC} es el periodo de retorno al que corresponde el valor seleccionado de avenida máxima, si ésta es originada por precipitaciones no ciclónicas; lo mismo ocurre en el caso de T_C, para precipitaciones ciclónicas.

Ejemplo numérico

Datos generales. Para estimar la magnitud de la avenida de diseño de la presa de almacenamiento La Fragua, que en la actualidad se construye en el río San Rodrigo (en el municipio de Zaragoza, Coahuila), se procesó probabilísticamente el registro de avenidas máximas anuales de la estación hidrométrica Cerca de El Moral, cuya área drenada es aproximadamente de 1 800 km². Aunque esta estación ha cambiado de emplazamiento, se logró integrar un registro único de 55 años, cuyos valores se muestran en el cuadro 2; con base en ellos se construyó el histograma de la ilustración 1, en la que se observa la probable ocurrencia de dos poblaciones. Los parámetros estadísticos correspondientes se presentan en el cuadro 3, en el cual se aprecia que el coeficiente de asimetría es bastante alto (C_s=2.708). La información hidrométrica fue obtenida de los Boletines Hidrométricos números 35 y 36, de la Comisión Internacional de Límites y Aguas (CILA).

Análisis probabilístico de avenidas máximas con poblaciones mezcladas

2. Estación hidrométrica "Cerca de El Moral"

m	Año	Gastos máximos anuales (m ³ /s)		Gasto máximo complementario (m ³ /s)		Gastos máximos no ciclónicos ordenados (m ³ /s)	Gastos máximos ciclónicos ordenados (m ³ /s)	
		Valor	Fecha de ocurrencia	Valor ordenado	valor			Fecha de ocurrencia
1	1932	2300.0	Sep.	7.1	4.2	Abr.	0.2	0.6
2	1933	66.0	Oct.	10.4*	7.6	Ene.	0.4	1.1
3	1934	32.2	Sep.	12.0*	16.2	May.	0.6	1.1
4	1935	1280.0	May.	12.0*	160.0	Sep.	2.1	1.2
5	1936	47.0	Jun.	13.7*	3.6	Sep.	2.3	1.8
6	1937	10.4	May. 30	15.2*	4.2	Sep. 12	4.0	2.6
7	1938	272.0	Jul. 23	15.6*	14.3	Jun. 10	4.2	2.9
8	1939	31.3	May. 13	16.8	5.2	Sep. 12	4.3	3.6
9	1940	345.3	Ago. 13	22.0	33.1	Mar. 24	7.2	4.2
10	1941	303.0	Sep. 18	24.2*	22.5	Abr. 29	7.3	5.2
11	1942	24.2	Abr. 21	30.0*	23.7	Sep. 7	7.6	6.1
12	1943	305.0	Sep. 18	31.3*	10.4	Abr. 29	10.4	7.1
13	1944	145.0	Ago. 30	32.2	12.2	May. 19	10.4	8.0
14	1945	362.0	Oct. 9	33.0	13.6	May. 12	12.0	8.7
15	1946	248.0	Jun. 24	40.1	35.8	Sep. 28	12.0	10.0
16	1947	66.4	May. 10	47.0*	8.0	Ago. 3	12.1	15.1
17	1948	33.0	Sep. 10	53.0*	4.3	Jun. 23	12.2	16.8
18	1949	210.0	Ago. 7	66.0	85.5	Abr. 28	13.2	21.8
19	1950	12.0	May. 11	66.4*	1.1	Sep. 6	13.6	22.0
20	1951	53.0	May. 16	69.0*	6.1	Oct. 28	13.7	23.7
21	1952	13.7	May. 1	75.0*	0.6	Dic. 29	14.3	25.6
22	1953	309.0	Sep. 2	78.6	4.0	Abr. 11	15.1	32.2
23	1954	132.0	Jun. 27	82.8	25.6	Sep. 21	15.2	33.0
24	1955	12.0	Jun. 17	92.4*	8.7	Ago. 15	15.6	35.0
25	1956	15.6	May. 27	94.2	15.1	Sep. 6	16.2	40.1
26	1957	283.0	May. 27	96.2	1.2	Oct. 9	16.9	66.0
27	1958	840.0	Sep. 28	105.0	44.9	Jun. 18	21.7	78.6
28	1959	16.8	Ago. 26	132.0*	15.1	May. 22	21.8	82.8
29	1960	133.0	Oct. 16	132.7	22.0	Abr. 27	22.0	94.2
30	1961	1970.0	Jun. 17	133.0	10.0	Jul. 17	22.5	96.2
31	1962	30.0	Abr. 27	145.0	1.8	Jul. 1	23.4	105.0
32	1963	96.2	Oct. 21	193.2	2.3	Abr. 4	24.0	132.6
33	1964	695.0	Sep. 24	210.0	96.2	May. 11	24.2	133.0
34	1965	22.0	Ago. 3	233.5	7.2	May. 18	30.0	145.0
35	1966	75.0	May. 20	237.6	21.8	Sep. 10	31.3	160.0
36	1967	555.0	Sep. 15	240.0	0.2	Abr. 2	33.1	187.2
37	1968	105.0	Jul. 3	248.0*	21.7	May. 11	36.3	193.2
38	1969	416.4	Abr. 11	272.0	187.2	Ago. 27	37.8	210.0
39	1970	78.6	Sep. 28	283.0*	7.3	Jun. 29	44.9	233.5
40	1971	193.2	Ago. 11	303.0	37.8	Jun. 28	47.0	237.6
41	1972	94.2	Ago. 4	305.0	21.8	Jun. 15	53.0	240.0
42	1973	314.4	Oct. 5	309.0	13.2	May. 2	66.4	272.0
43	1974	237.6	Sep. 23	314.4	12.1	Mar. 14	69.0	303.0
44	1975	2382.0	Jul. 18	345.3	98.4	Abr. 7	75.0	305.0
45	1976	1668.0	Jul. 12	362.0	207.6	May. 7	85.5	309.0
46	1977	69.0	May. 22	382.8	1.1	Jul. 1	92.4	314.4
47	1978	240.0	Nov. 5	416.4*	2.1	Jun. 7	96.2	345.3
48	1979	15.2	Jun. 5	555.0	2.9	Jul. 20	98.4	362.0
49	1980	382.8	Ago. 11	695.0	23.4	May. 26	132.0	382.5
50	1981	92.4	Abr. 23	840.0	2.6	Jul. 7	207.6	555.0
51	1982	7.1	Sep. 17	1280.0*	0.6	Feb. 25	248.0	695.0
52	1983	132.6	Sep. 18	1668.0	0.4	Ene. 2	283.0	840.0
53	1984	233.5	Oct. 11	1870.0*	16.9	May. 19	416.4	1668.0
54	1985	40.1	Sep. 29	2300.0	24.0	Jun. 27	1280.0	2300.0
55	1986	82.8	Oct. 11	2382.0	36.3	Jun. 26	1870.0	2382.0

Curva de frecuencia de poblaciones combinadas. Primero, se identificaron los eventos ciclónicos, con base en la información disponible de la zona, en los Boletines Hidrométricos número 34, tomos II y III de la Secretaría de Recursos Hidráulicos y en el Atlas de Huracanes de la Secretaría de Programación y Presupuesto. Los eventos fueron los siguientes:

año	Q(m³/s)	m	Q(m³/s)	X=log Q	Q(x)=m/N	K
1932	2300.0	1	2300.0	3.36173	0.01786	2.1009715
1954	132.0	2	840.0	2.92428	0.03571	1.8028750
1958	840.0	3	555.0	2.74429	0.05357	1.6107286
1967	555.0	4	237.6	2.37585	0.07143	1.4643420
1974	237.6	5	132.0	2.12057	0.08929	1.3439018

De acuerdo con lo anterior:

$$S = S_x/S_k = \frac{0.48239}{0.29807} = 1.61839 \quad (15)$$

$$\bar{X} = 2.70534 - 1.61839(1.66456) = 0.01143 \quad (13)$$

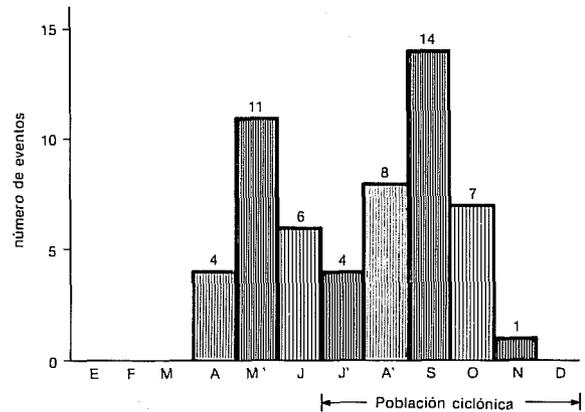
$$X = 0.01143 + 1.61839 \cdot K \quad (12)$$

En el caso de los eventos restantes, se ajustó una función de distribución de probabilidades; para ello, de los modelos probados se seleccionó el que arrojó el menor error medio cuadrático entre los datos y los valores pronosticados con ese fin para las mismas probabilidades asociadas a cada dato como posición gráfica. Los modelos que cumplieron este requisito fueron la función de distribución gamma de tres parámetros y la general de valores extremos (Raynal, 1985). En el cuadro 4 se concentran los resultados obtenidos mediante la ecuación (16) para 17 valores de avenidas máximas seleccionados; los resultados correspondientes se muestran en la ilustración 3.

3. Parámetros estadísticos de la serie de gastos máximos anuales (estación hidrométrica "Cerca de El Moral")

Parámetro	Valor	Parámetro	Valor
Media (\bar{x})	327.8 m³/s	Media logarítmica ($\log x$)	2.07721
Desviación típica (s)	545.0 m³/s	Índice de variabilidad (Iv)	0.64279
Coficiente de variación (Cv)	1.663	$\sum (L_n x_i)$	263.0627
Coficiente de asimetría (Cs)	2.708	Coficiente de oblicuidad (g)	+ 0.09455

1. Histograma de gastos máximos anuales (estación hidrométrica "Cerca de El Moral")



Segregación de la muestra. En este caso se consideró conveniente segregar la muestra por semestre, para no incluir avenidas máximas importantes que ocurrieron en julio y agosto. En el cuadro 2 se señalan con un asterisco los eventos no ciclónicos de la muestra ya ordenada en forma progresiva de acuerdo con sus magnitudes. Los valores característicos de cada prueba estadística aplicada son los siguientes:

Prueba de Mann-Whitney:

$$\begin{aligned} n_1 &= 21N = 55 \quad U_1 = 513 \\ n_2 &= 34R_1 = 432 \quad U_2 = 201 \\ |Z| &= 2.703 > 1.96 \therefore \text{se rechaza } H_0 \end{aligned}$$

Prueba de Kruskal-Wallis:

$$H = 7.304 > 3.84 \therefore \text{se rechaza } H_0$$

Prueba de Terry:

$$\begin{aligned} c &= 8.9165 \quad \sum \{E(Z_{N_i})\}^2 = 52.398056 \\ S &= 3.5492 \quad t = c/S = 2.5123 > 1.96 \therefore \text{se rechaza } H_0 \end{aligned}$$

4. Cálculos relativos al ejemplo numérico

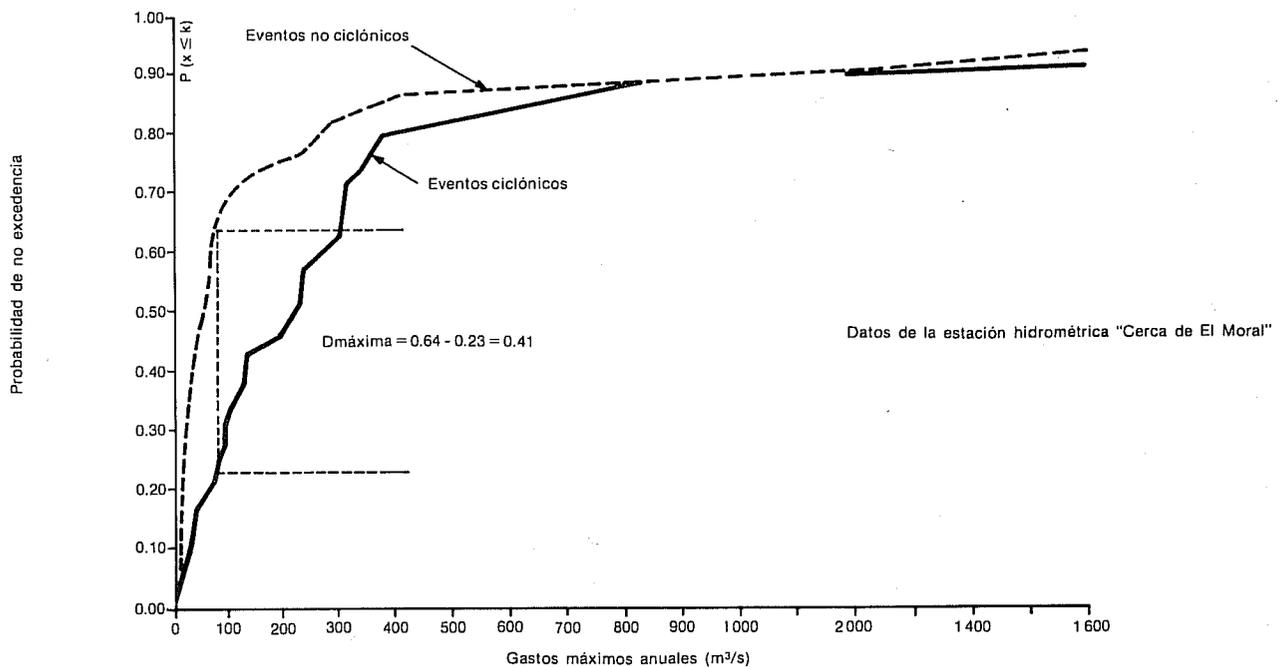
Avenida máxima (X) (m³/s)	Términos de la ecuación 16			Periodo de retorno (Tr) (Años)	Avenida máxima (X) (m³/s)	Términos de la ecuación 16		
	P ₁	P ₃	P _c			T _{NC} (Años)	T _C (Años)	T _A (Años)
10	0.9750	0.280	0.982	1.018	10	2.63	1.23	1.13
20	0.8700	0.210	0.897	1.115	20	3.23	1.35	1.22
40	0.7100	0.170	0.759	1.318	40	4.00	1.59	1.39
70	0.5800	0.125	0.633	1.580	70	5.00	1.89	1.60
100	0.5000	0.110	0.555	1.800	100	5.55	2.13	1.77
150	0.4100	0.090	0.463	2.160	150	7.14	2.56	2.10
200	0.3700	0.080	0.420	2.380	200	9.09	3.03	2.48
300	0.2800	0.060	0.323	3.096	300	12.50	4.00	3.23
500	0.1800	0.047	0.219	4.570	500	22.22	6.25	5.06
800	0.1000	0.040	0.136	7.350	800	40.00	11.80	9.29
1000	0.0700	0.034	0.102	9.800	1000	55.55	15.40	12.23
1500	0.0300	0.027	0.056	17.860	1500	125.00	28.60	23.43
2000	0.0140	0.020	0.034	29.400	2000	250.00	58.82	47.77
3000	0.0035	0.017	0.020	50.000	3000	770.00	250.00	188.90
4000	0.0015	0.013	0.014	68.960	4000	2500.00	625.00	516.16
5000	0.0005	0.011	0.012	86.980	5000	5000.00	2000.00	1428.00

Prueba de Kolmogorov-Smirnov:

En la ilustración 2 se presentan las curvas de frecuencias acumuladas de cada población, de la cual se obtuvo $D=0.41$, por lo que:
 $\chi^2=8.729 > 5.99 \therefore$ se rechaza H_0

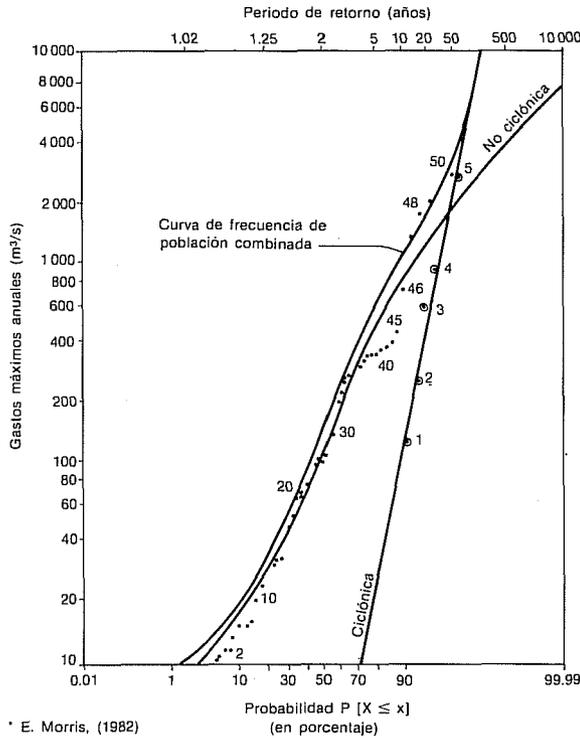
Una vez probada la existencia de dos poblaciones, se integran las muestras de avenidas máximas anuales no ciclónicas y ciclónicas, que también se concentran en el cuadro 2; la información requerida para ese fin se obtuvo de los boletines de la CILA. Cada muestra se procesó por medio

2. Curvas de frecuencias acumuladas para la prueba Kolmogorov-Smirnov



3. Curva de frecuencia de poblaciones combinadas*

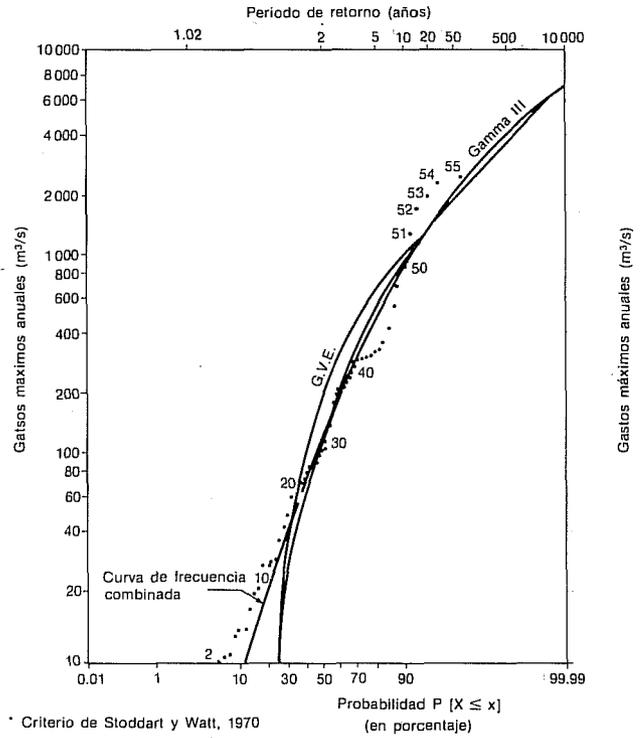
Proyecto: La Fragua Estado: Coahuila Estación: Cerca de El Moral
 Datos procesados: Qmáx. anuales Cálculo: DFCA Fecha: marzo de 1987



* E. Morris, (1982)

4. Curva de frecuencia combinada*

Proyecto: La Fragua Estado: Coahuila Estación: Cerca de El Moral
 Datos procesados: Qmáx. anuales Cálculo: R G de L Fecha: marzo de 1987



* Criterio de Stoddart y Watt, 1970

de una función de distribución de probabilidades, para lo cual se adoptó, como ya se indicó, el modelo que condujo al menor error medio cuadrático; los resultados de ciertos periodos de retorno se concentraron en el cuadro 5.

Por último, la aplicación del criterio de Stoddart y Watt (ecuación 17) permitió obtener los resultados que se muestran también en el cuadro 4 y en la ilustración 4.

Conclusiones

Las principales ventajas de los enfoques expuestos son las siguientes: primera, permiten tomar en cuenta la génesis de las avenidas, y segunda, dan la posibilidad de emplear más información relativa a las avenidas registradas.

En relación con los resultados obtenidos para el ejemplo numérico expuesto, cabe aclarar que el procedimiento utilizado para integrar una curva

5. Gastos máximos estimados (m³/s) (Estación Hidrométrica "Cerca de El Moral")

Población	Periodo de retorno (años)				Distribución aplicada
	100	500	1000	10000	
No ciclónica (abril a junio)	1460	2570	3120	5230	Gamma de 3 parámetros General de valores extremos
	1230	2110	2600	4950	
Ciclónica (julio a noviembre)	2340	3750	4420	6850	Gamma de 3 parámetros General de valores extremos
	2050	3300	3980	6950	
Máximos anuales	2540	3850	4450	6600	Gamma de 3 parámetros General de valores extremos
	2290	3530	4170	6860	

de frecuencias de poblaciones combinadas tiene el grave inconveniente de que exige definir, para un número normalmente reducido de eventos de alto valor (serie discontinua), la curva de regresión modificada, que en el papel log-normal da como resultado una recta de pendiente muy pronunciada, lo cual arroja valores muy elevados para los grandes periodos de retorno, como se observa en la ilustración 3.

Aunque la división por semestres no es muy ortodoxa, se considera que los resultados obtenidos por medio de ese procedimiento (véase cuadro 5) son confiables.

Por último, con respecto al criterio de Stoddart y Watt, en la ilustración 4 se observa que la curva de frecuencias combinada coincide prácticamente, en su porción de altos periodos de retorno, con el mejor ajuste logrado en la muestra mezclada y que el ajuste mejora con los datos de baja magnitud, lo cual significa que también es confiable.

Referencias

- Abramowitz, M., e I. Stegun. *Handbook of Mathematical Functions*, capítulo 26 "Probability Functions", Dover Publications, Inc., novena edición, EUA, pp. 925-996, 1970.
- Browzin, B.S., C.A. Baumbusch y M. G. Pavlides. Significance of the Genesis of Floods on Probability Analysis, Procedimientos del II Simposio Internacional en Hidrología, sesión V, "Regional Flood Frequency Analysis", Water Resources Publications, EUA, pp. 450-461, 1972.
- Comisión Internacional de Límites y Aguas. *Boletines Hidrométricos*, números 35 y 36, México.
- Harter, H.L. "Expected Values of Normal Order Statistics", *Biometrika*, vol. 48, números 1 y 2, pp. 151-165, 1961.
- Kite, G. W. *Frequency and Risk Analysis in Hydrology*, capítulo 2, "Data", Water Resources Publications, Fort Collins, Colorado, EUA, pp. 4-25, 1977.
- Morris, E. C. "Mixed-Population Frequency Analysis" Training Document 17 of the Hydrologic Engineering Center of the US. Army Corps of Engineers, Davis, California, EUA, pp. 43, 1982.
- Raynal, V.. *Apuntes del Curso de Hidrología Superficial*, capítulo 5 "Análisis Estadístico en Hidrología", DEPFI, UNAM, México, 1985.
- Siegel, S. *Estadística no paramétrica aplicada a las ciencias de la conducta*, capítulo 6 "El caso de dos muestras independientes" y capítulo 8 "El caso de K muestras independientes", Ed. Trillas, México, 1976.
- Secretaría de Programación y Presupuesto. *Atlas de huracanes en el Océano Pacífico y en el Atlántico*, DETENAL, México, 1979.
- US Water Resources Council. "Guidelines for Determining Flood Flow Frequency". Bulletin 17A of Hydrology Committee, EUA, 1977.
- US Army Corps of Engineers. "Hydrologic Frequency Analysis", *Hydrologic Methods for Water Resources Development*, vol. 3, The Hydrologic Engineering Center, EUA, 1975.