

Análisis estadístico de las series anuales de máximas lluvias diarias en la cuenca del Guadalhorce

JAVIER FERRER POLO (*)

RESUMEN. Se presenta el análisis estadístico de las series anuales de máximas lluvias diarias en 47 estaciones meteorológicas situadas en la cuenca del río Guadalhorce, aplicando metodologías usualmente empleadas en estudios de avenidas. Tres leyes de distribución han sido consideradas como posibles candidatas en el análisis: valores extremos generalizados (GEV), log-Pearson tipo 3 (LP3) y valores extremos con dos componentes (TCEV). Cada una de ellas se ha aplicado con un específico método de estimación de parámetros y un determinado esquema de uso combinado de datos locales y regionales, encontrándose sólo pequeñas diferencias en los resultados. Asimismo se ha llevado a cabo la comparación de las tres distribuciones candidatas utilizando los siguientes procedimientos: índices de ajuste y análisis mediante simulación del comportamiento estadístico y de la robustez del método.

ABSTRACT. Frequency analysis of daily rainfall annual maximum series has been performed in 47 meteorological stations located in Guadalhorce basin, applying approaches usually used in flood studies. Three candidate distributions have been considered: General Extreme Value (GEV), Log-Pearson type 3 (LP3) and Two Component Extreme Value (TCEV) with selected parameter estimation methods and schemes of joint use of at-site and regional data and only small differences have been found in results. Comparison of three procedures has been carried out in three ways: numerical indices of agreement, behaviour analysis and robustness simulation.

INTRODUCCION

El estudio hidrológico de las avenidas del río Guadalhorce (3.100 km²), situado en el Sureste de la costa mediterránea, ha incluido el empleo de modelos hidrológicos que requieren la determinación de las propiedades estadísticas de las máximas lluvias en la zona.

La mayor parte de las estaciones meteorológicas de la cuenca sólo disponen de registros de precipitación diaria, habiéndose seleccionado, tras un análisis previo, 47 estaciones con más de 30 años de datos y una longitud media de 40 años.

La modelación estadística de series de máximos (WMO 1989) requiere la elección de: a) tipo de modelo, b) ley de distribución de la población, c) método de estimación de parámetros y cuantiles, d) esquema de uso combinado de datos locales y regionales.

La elección del tipo de modelo en la práctica se reduce al empleo de series de máximos anuales (AM) o de series parciales o sobre un umbral (POT). La mayor parte de los trabajos de investigación y de los estudios reales utilizan el enfoque de AM (Cunnane, 1987), más fácil de aplicar y usualmente más eficiente en un sentido estadístico, siendo éste el enfoque aquí seleccionado.

El presente artículo se centra en los aspectos b), c) y d) de la modelación estadística de las series anuales de máximas lluvias diarias en la cuenca del Guadalhorce, siguiendo los métodos más extendidos o más recientes y que son utilizados generalmente en el análisis estadístico de avenidas.

LEYES DE DISTRIBUCION CANDIDATAS

WMO (1989), en su Apéndice 6, revisa un informe del Secretariado de la WMO que resume las contestaciones a un cuestionario enviado a 55 agencias de 28 países en el que se solicitaba información sobre las leyes de distribución habitualmente empleadas en el análisis de valores extremos de lluvia y caudal. Los resultados, para el análisis de lluvia, muestran dos grupos principales de leyes:

A. Distribución de valores extremos (52 %): valores extremos generalizados (GEV) que incluye como casos particulares las leyes Gumbel (EV1), Fréchet (EV2) y Weibull (EV3).

B. Distribución Gamma (31 %): Pearson III (P3); Log-Pearson III (LP3) y Log-Normal (LN).

Por otra parte los trabajos de Rossi et al. (1984) sugieren que la ley de valores extremos con dos componentes (TCEV) puede reproducir las propiedades esta-

(*) Centro de Estudios Hidrográficos del CEDEX.

DISTRIBUCION	$f(x)$ o $F(x)$	PARAMETROS
GEV	$f(x) = \exp \left\{ - \left[1 - k \left(\frac{x-u}{\alpha} \right) \right]^{1/k} \right\}$	u, α, k
LP3	$f(x) = \frac{(\log x - u)^{k-1}}{k \alpha \Gamma(k)} \exp \left\{ - \left(\frac{\log x - u}{\alpha} \right)^k \right\}$	u, α, k
TCEV	$F(x) = \exp \{ - \alpha_1 x^{-\theta_1} - \alpha_2 x^{-\theta_2} \}$	$\alpha_1, \theta_1, \alpha_2, \theta_2$

TABLA 1. Funciones de distribución.

dísticas de las series AM observadas, mejor que otras distribuciones más tradicionales.

Como conclusión, en el presente trabajo se han seleccionado como posibles candidatas las siguientes leyes de distribución: GEV, LP3 (basado en el hecho de que la distribución LN es un caso particular con sesgo nulo en el espacio logarítmico) y TCEV, cuya formulación se resume en la tabla 1.

MÉTODOS DE ESTIMACION DE PARAMETROS Y CUANTILES

La estimación de los parámetros relacionados con el sesgo de la población requiere el empleo de métodos regionales debido a la elevada varianza existente en el sesgo de los datos locales. Estos métodos regionales asumen que las series locales en una región son similares en relación con alguna característica estadística (generalmente coeficiente de sesgo: C_s o coeficiente de variación: C_v).

La primera etapa del enfoque regional, y probablemente la más controvertida, es la definición de regiones homogéneas. USWRC (1981) sugiere criterios geográficos y «el uso de al menos 40 estaciones, o todas las estaciones en un radio de 100 millas» para determinar el sesgo regional (en el espacio logarítmico) de los caudales. Dicha normativa propone el uso de la media de los valores locales del sesgo, si no hay ninguna tendencia geográfica evidente como resultado de un análisis gráfico. Este análisis gráfico se ha llevado a cabo en la cuenca del Guadalquivir y no ha sido posible definir un plano de isolinias de sesgo de forma clara, por lo que se ha asumido finalmente una única región desde el punto de vista del sesgo.

Para contrastar la hipótesis de homogeneidad regional se han calculado el valor medio del C_s ($\bar{C}_s$) y el coeficiente de variación del C_s ($C_{v_{C_s}}$), obteniendo los siguientes resultados: $\bar{C}_s = 0,43$ y $C_{v_{C_s}} = 0,21$ que han sido considerados valores razonablemente bajos.

Una vez definida la región, es necesario seleccionar un esquema completo para la estimación de parámetros y cuantiles utilizando de forma combinada los datos lo-

cales y regionales. Está fuera del ámbito de este trabajo analizar y comparar los méritos de todos los posibles métodos y sólo se ha seleccionado para cada ley de distribución candidata aquel método más extendido o con mejores propiedades de acuerdo con recientes trabajos, según se expone en los siguientes párrafos, por lo que al referirnos a una determinada ley nos estamos refiriendo al método completo.

— La distribución GEV fue recomendada en Estudios de Avenidas realizados en Gran Bretaña (NERC, 1975) con un método de «índice de avenidas» para cada una de las 10 regiones en que se dividió el territorio. Este método asume que la variable $y = X/\bar{X}$ tiene, en todos los puntos, la misma ley de distribución, cuyos parámetros son estimados a partir de un análisis regional de los distintos conjuntos de datos locales. Los resultados en cada punto (X) se obtienen multiplicando los valores de y por la media de los datos locales ($\bar{X}$). Este método, no sólo asume un sesgo regional constante, sino también un coeficiente de variación regional constante.

El enfoque utilizado en este artículo, es el método de «índice de avenidas» propuesto por Hosking et al. (1985), basado en una media regional de los momentos ponderados probabilísticamente (PWM) calculados en cada punto expresados en forma adimensional. Este método será denominado GEV/PWM.

— La distribución LP3 es la recomendada por USWRC (1981) en estudios de avenidas, siguiendo la tradicional ecuación de log-momentos:

$$\log X_T = \bar{Y} + K_T S \quad (1)$$

donde el valor de la variable para una determinada probabilidad de excedencia X_T es función de K_T , denominando factor de frecuencia, de la media ($\bar{Y}$) y de la desviación típica (S), estas dos últimas estimadas en cada punto a partir de los logaritmos de los datos. El factor de frecuencia es calculado a partir de la probabilidad de excedencia: T y del valor regional del coeficiente de sesgo: C_s , obtenido por media regional de las estimaciones locales de C_s . Estas últimas estimaciones se obtienen siguiendo a Wallis et al. (1974), mediante:

$$C_s = \frac{N \sum (y - \bar{y})^3}{(H-1)(N-2)S^3} (1 + 8,5/N) \quad (2)$$

donde los valores de y son los logaritmos de los datos y N el número de elementos de la serie. Este método se denominará LP3/LMOM en este artículo.

— La distribución TCEV fue recomendada por Rossi et al. (1984) en análisis regional de avenidas con un método de estimación de cuantiles consistente en las siguientes tres etapas. En primer lugar los datos locales son estandarizados mediante:

$$y = \frac{x - u}{\alpha} \quad (3)$$

donde (u, α) son estimadores de los parámetros de la distribución EVI obtenidos por máxima verosimilitud (ML) a partir de las series originales en las cuales se han eliminado previamente los valores anómalamente altos o «outliers». Estos valores y se agrupan, asumiendo la hipótesis «año-estación» en una única serie que resulta seguir una función de distribución análoga a los valores X originales:

$$F(y) = \exp(-a_1 e^{-y/\theta_1} - a_2 e^{-y/\theta_2}) \quad (4)$$

En la segunda etapa, se estiman por ML los parámetros regionales $(a_1, a_2, \theta_1, \theta_2)$ y se obtienen los cuantiles Y_T para los valores deseados de probabilidad de excedencia T . Finalmente, los valores locales de la variable X_T pueden obtenerse invirtiendo la ecuación (3):

$$X_T = u_K + a_K Y_T \quad (5)$$

donde (u_K, a_K) son los valores (u, α) obtenidos de la serie local K . Este método será llamado TCEV/ML.

RESULTADOS ALCANZADOS

En la tabla 2 se muestran para cada uno de los tres modelos alternativos los valores estimados para los parámetros regionales y los valores medios de los parámetros locales. Asimismo en la tabla 3 se recogen los cuantiles estimados para distintos períodos de retorno (relación $X-T$) a partir de dichos parámetros:

	GEV/PWM	LP3/LMOM	TCEV/ML
PARAMETROS REGIONALES	$u=0,79$ $\alpha=0,28$ $k=-0,156$	$C_1=0,501$	$a_1=0,414$ $a_2=0,517$ $\theta_1=0,651$ $\theta_2=1,627$
PARAMETROS LOCALES (VALORES MEDIOS)	$\bar{x}=67,13$	$\bar{y}=1,78$ $s=0,17$	$u=54,47$ $\alpha=18,63$

TABLA 2. Estimación de parámetros.

T(AÑOS)	GEV/PWM	LP3/LMOM	TCEV/ML
2	60	58	60
10	104	102	105
50	154	150	153
100	180	173	174
500	250	236	224
1.000	287	267	243

TABLA 3. Estimación de cuantiles (mm).

MÉTODOS PARA COMPROBAR DISTRIBUCIONES Y ESQUEMAS DE ESTIMACION DE CUANTILES

Existe una amplia variedad de técnicas que han sido usadas para comparar los diferentes métodos de análisis de leyes de frecuencia y de realizar una adecuada selección entre ellos. Estas técnicas pueden agruparse según los dos enfoques siguientes:

A. Pruebas de capacidad descriptiva que analizan en qué medida el modelo elegido describe y reproduce algunos aspectos de los datos observados. Se llevarán a cabo dos de este tipo de pruebas: el cálculo de índices numéricos de ajuste y el análisis del comportamiento estadístico de la cola derecha de la distribución.

B. Pruebas de la capacidad predictiva que estudian en qué grado la distribución candidata puede estimar la relación $X-T$ incluso cuando la población no sigue la distribución asumida, con lo que se analiza su robustez.

INDICES NUMERICOS DE AJUSTE

Estos métodos tradicionales (Benson, 1968) se basan en alguna fórmula de gráfico que atribuye probabilidad a la muestra de datos previamente ordenada. En este caso se ha seleccionado la fórmula de Gringorten:

$$F_i = \frac{i - 0,44}{N + 0,12} \quad (6)$$

donde F_i es la probabilidad gráfica asignada, N el tamaño de la muestra e i el número de orden en sentido creciente.

Para cada período de retorno seleccionado (T), se obtiene el valor de la variable (X_T) mediante interpolación lineal (en papel de probabilidad de Gumbel) entre puntos observados adyacentes, y por último se calcula la distancia al cuantil estimado mediante el ajuste de la distribución seleccionada (X_T) según la expresión:

$$d_i = \frac{X_T - X_i}{X_i} \quad (7)$$

La media de los valores absolutos de estas distancias calculadas para cada estación K se considera un índice de ajuste para cada período de retorno y su suma para los NT períodos de retorno considerados proporciona un valor global del ajuste según (8) y (9)

$$D_T = \sum_k |d_T|_k / M \quad (8)$$

$$d = \sum_T D_T \quad (9)$$

Los resultados obtenidos se resumen en la tabla 4.

AÑOS	T = 2	T = 5	T = 10	T = 25	T = 50	D
GEV	0,055	0,052	0,083	0,124	0,173	0,098
LP3	0,042	0,048	0,066	0,091	0,134	0,076
TCEV	0,040	0,044	0,074	0,137	0,191	0,097

TABLA 4. Índices numéricos de ajuste.

ANÁLISIS DEL COMPORTAMIENTO DE LA COLA DERECHA DE LA DISTRIBUCIÓN

El comportamiento de la cola derecha de las distintas distribuciones es una de las características más interesantes a preservar en la modelación estadística de series de máximos y para su análisis, de forma análoga a Rossi et al. (1984), se han seleccionado los siguientes parámetros:

— Sesgo, según:

$$c_s = \frac{N \sum (X - \bar{X})^3}{(N-1)(N-2)s^3} \quad (10)$$

— Máximo valor estandarizado, según:

$$Y_s = \frac{X_n - \bar{X}}{s} \quad (11)$$

donde para cada muestra, X_n es el valor máximo, $\bar{X}$ es la media y s la desviación típica.

En primer lugar se ha procedido a la generación de 3.000 muestras de un tamaño $N = 40$ para cada una de las distribuciones candidatas (según los parámetros de la tabla 2) y al cálculo en cada muestra de los valores de C_s e Y_s mediante las expresiones (10) y (11).

El análisis del comportamiento estadístico se realiza mediante la comparación gráfica (en papel de probabilidad de Gumbel) de los valores muestrales de C_s e Y_s obtenidos por generación juntamente con los realmente observados en las series del Guadalquivir. Los resultados se muestran en las figuras 1 y 2 utilizando la fórmula de gráfico (6) y sustituyendo los puntos representativos de los valores generados por una línea suavizada. En estas figuras se ha supuesto asimismo los resultados correspondientes a una distribución EVI, observándose que esta última distribución no reproduce adecuadamente los dos parámetros considerados. Por contra, el análisis no resulta concluyente respecto a las tres distribuciones analizadas.

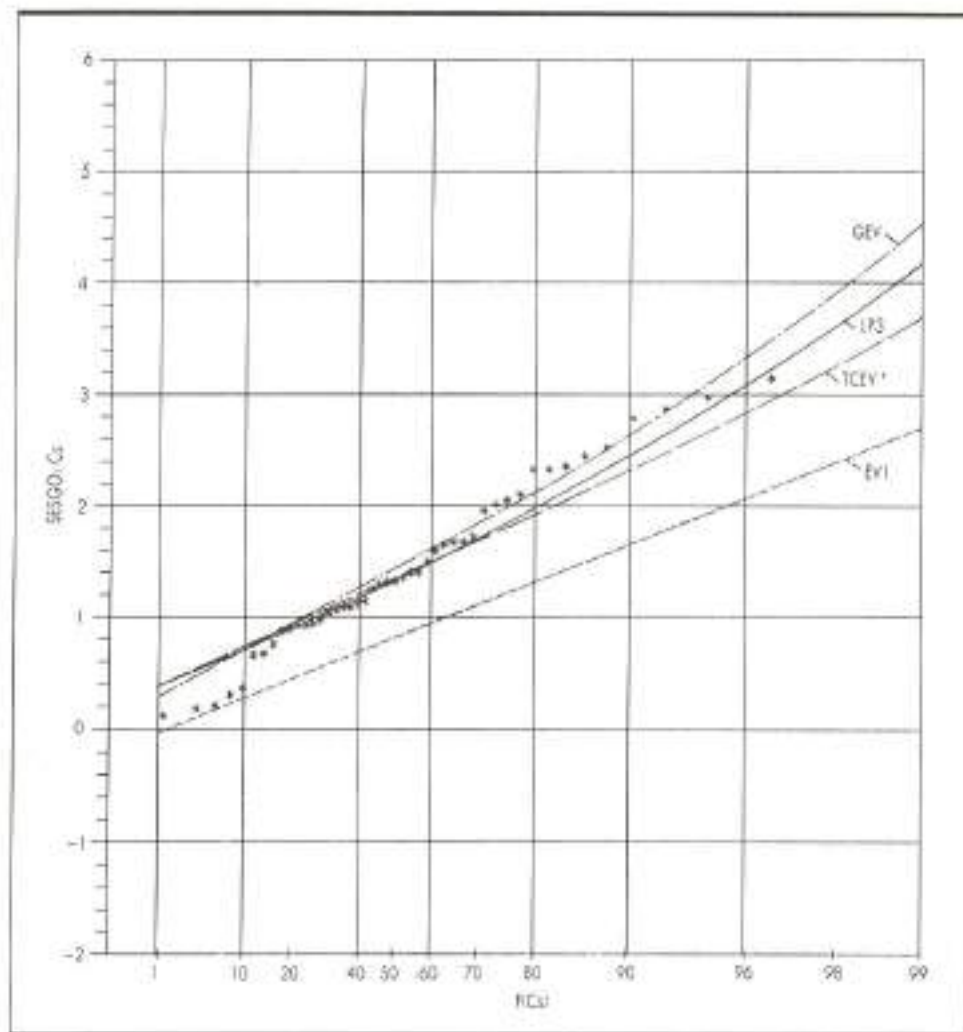


FIGURA 1. Valores observados y simulados de la función de distribución muestral del sesgo C_s .

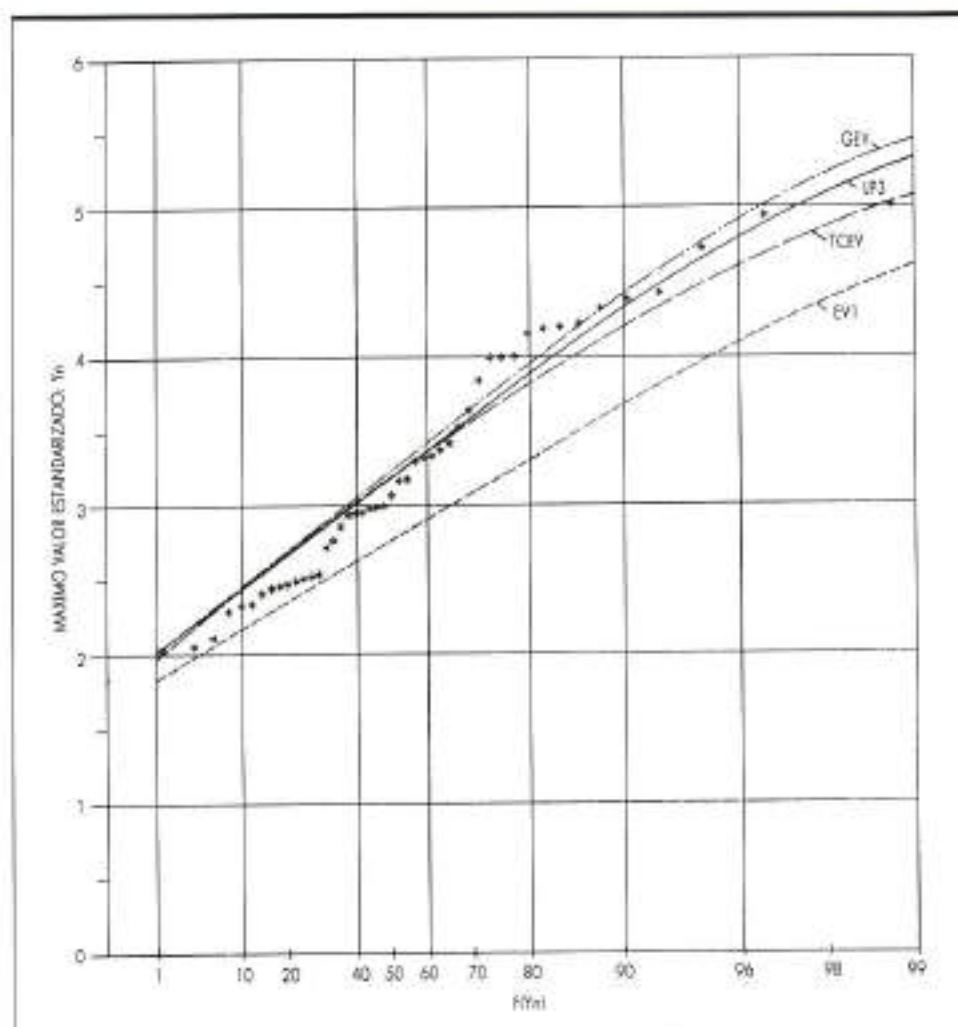


FIGURA 2. Valores observados y simulados de la función de distribución muestral del máx. valor estandarizado Y_n .

ANÁLISIS DE ROBUSTEZ

Las pruebas para comprobar la capacidad predictiva tratan de analizar la robustez (magnitud de sesgo y del error cuadrático medio) del modelo definido por la distribución y el procedimiento de ajuste, cuando se asumen cambios en la distribución de la población. El análisis se realiza mediante métodos de simulación (Lettenmaier et al., 1985), siguiendo los siguientes pasos:

A. Selección de una de las distribuciones candidatas como población con lo que son conocidos los verdaderos valores de los cuantiles: X_i (tabla 3).

B. Generación, por simulación de Montecarlo, de una región de datos: 47 series de 40 años de longitud utilizando los parámetros mostrados en la tabla 2.

C. Para cada uno de los tres métodos alternativos propuestos y para los períodos de retorno considerados (T) se procede al cálculo de los siguientes parámetros:

- Cuantiles estimados (X_T) para cada una de las 47 estaciones.
- Sesgo (b) y raíz cuadrada del error cuadrático medio (RMSE), expresados en forma unitaria mediante su cociente con el valor verdadero del cuantil, según las expresiones:

$$b_T = \left(\frac{\sum X_T}{N} - X_i \right) / X_i \quad (12)$$

$$RMSE_T = \sqrt{\left(\frac{\sum (X_T - X_i)^2}{N} \right)} / X_i \quad (13)$$

Los pasos b) y c) se repiten 100 veces para obtener valores medios y los resultados se muestran en las tablas 5 y 6.

PERIODO RETORNO (AÑOS)	POBLAC. GEV	POBLAC. TCEV	POBLAC. LP3	MAXIMO VALOR (+)	MINIMO VALOR (-)
GEV/PWM					
2	0,002	0,005	0,003	0,005	-
10	-0,001	-0,016	-0,007	-	-0,016
50	-0,005	-0,003	0,003	0,003	-0,005
100	-0,008	0,018	0,010	0,018	-0,008
500	-0,014	0,102	0,034	0,102	-0,014
1.000	-0,016	0,159	0,046	0,159	-0,016
LP3/LMOM					
2	0,000	0,007	0,000	0,007	0,000
10	0,010	-0,005	0,006	0,010	-0,005
50	0,000	0,001	0,015	0,015	0,000
100	-0,007	0,016	0,020	0,020	-0,007
500	-0,029	0,076	0,034	0,076	-0,029
1.000	-0,039	0,120	0,039	0,120	-0,039
TCEV/ML					
2	0,003	0,007	0,000	0,007	0,000
10	-0,005	-0,008	0,007	0,007	-0,008
50	-0,003	-0,004	-0,011	-	-0,011
100	-0,009	-0,002	-0,030	-	-0,030
500	-0,058	0,001	-0,093	0,001	-0,093
1.000	-0,089	0,009	-0,124	0,009	-0,124

TABLA 5. Sesgo en la estimación de cuantiles.

Una medida de bondad de un método desde el punto de vista del sesgo es comparar su mejor o peor valor cuando se aplica a una gama de poblaciones distintas. El criterio adoptado asume que el mejor procedimiento es aquel que conduce a sesgos positivos (resultados conservadores) de menor valor y el peor aquel que tiene sesgos negativos (resultados inseguros) y de mayor magnitud, todo ello con independencia de la hipótesis asumida respecto a la población. Estos resultados se muestran en la tabla 7, utilizándose información de la tabla 5.

Respecto al valor de RMSE, los resultados muestran que los modelos GEV/PWM y TCEV/ML son más eficientes que el LP3/LMOM incluso si la población sigue una distribución LP3. La tabla 8 muestra, utilizando información de la tabla 6, el mejor método definido como aquel con menor valor máximo de RMSE y el peor considerando el mayor valor de RMSE, todo ello con independencia de la hipótesis hecha respecto a la población.

CONCLUSIONES

Las metodologías usualmente aplicadas en el análisis de leyes de frecuencia de caudales de avenida son igualmente útiles para llevar a cabo la modelación estadística de series anuales de máximas lluvias diarias. Los méto-

PERIODO RETORNO (AÑOS)	POBLACION GEV	POBLACION TCEV	POBLACION LP3	VALOR MAXIMO
GEV/PWM				
2	0,073	0,070	0,074	0,074
10	0,073	0,071	0,073	0,073
50	0,074	0,070	0,074	0,074
100	0,073	0,073	0,075	0,075
500	0,074	0,127	0,084	0,127
1.000	0,074	0,178	0,090	0,178
LP3/LMOM				
2	0,062	0,062	0,062	0,062
10	0,108	0,100	0,104	0,108
50	0,152	0,142	0,148	0,152
100	0,170	0,161	0,167	0,170
500	0,207	0,223	0,210	0,223
1.000	0,223	0,264	0,228	0,264
TCEV/ML				
2	0,067	0,067	0,065	0,067
10	0,09	0,099	0,098	0,099
50	0,114	0,119	0,113	0,119
100	0,119	0,124	0,118	0,124
500	0,134	0,134	0,147	0,147
1.000	0,149	0,138	0,167	0,167

TABLA 6. RMSE de la estimación de cuantiles.

dos contrastados: GEV/PWM, LP3/LMOM y TCEV/ML, incluyendo determinados esquemas para utilizar de forma combinada los datos locales y regionales, han alcanzado resultados similares (tabla 3) con diferencias de alrededor del 15 % para períodos de retorno de 1.000 años. El orden de los métodos alternativos en sentido creciente de resultados ha sido: TCEV/ML, LP3/LMOM y GEV/PWM.

Respecto a la evaluación de índices numéricos de

PERIODO RETORNO (AÑOS)	MEJOR MAXIMO SESGO POSITIVO	PEOR MAXIMO SESGO NEGATIVO
2	GEV	TCEV
10	TCEV	GEV
50	GEV	TCEV
100	GEV	TCEV
500	LP3	TCEV
1.000	LP3	TCEV

TABLA 7. Distribución más y menos preferible desde el punto de vista del sesgo.

PERIODO RETORNO (AÑOS)	MENOR MAXIMO RMSE	MAYOR MAXIMO RMSE
2	IP3	GEV
10	GEV	IP3
50	GEV	IP3
100	GEV	IP3
500	GEV	IP3
1.000	TCEV	IP3

TABLA 8. Distribución más y menos preferible desde el punto de vista del RMSE.

ajuste entre las series observadas y los cuantiles estimados el método LP3/LMOM conduce a los mejores resultados con pequeñas diferencias entre los otros dos candidatos GEV/PWM y TCEV/ML. La restricción suplementaria impuesta en el método regional utilizado con la ley GEV (método de «índice de avenida» que asume valor constante del C_v) ha tenido pequeña repercusión en los resultados y sólo para 2,5 y 10 años de período de retorno conduce a los peores resultados.

Los tres métodos alternativos reproducen correctamente el comportamiento estadístico de la cola derecha de la distribución y tanto el sesgo como el máximo valor estandarizado de las muestras observadas son similares a aquellos calculados en las muestras generadas (figuras 1 y 2). Debe resaltarse que este análisis ha permitido comprobar que la distribución EVI (Gumbel) no parece una elección adecuada para el análisis de las lluvias máximas de la cuenca del Guadalupe.

El análisis de robustez, llevado a cabo mediante experiencias de simulación, muestra que los métodos GEV/PWM (para períodos de retorno medios) y LP3/LMOM (para altos períodos de retorno) tienen los menores sesgos positivos en la estimación de cuantiles, pero que el método TCEV/ML tiende a estimaciones sesgadas de forma negativa. El estudio de los valores de RMSE refleja la principal debilidad del empleo de método LP3/LMOM, que conduce al mayor valor de los tres métodos, debido a su alta varianza. Desde este punto de vista el método GEV/PWM es el mejor, con similares resultados por parte del método TCEV/ML, aunque la

hipótesis de C_v constante asumida en aquél es especialmente favorable debido a la homogeneidad regional implícita en las experiencias de simulación.

Como conclusión final, el método tradicional LP3/LMOM es un procedimiento fácil de aplicar con un uso geográfico de los datos regionales y ha conducido a resultados intermedios respecto a los obtenidos por los métodos GEV/PWM y TCEV/ML. Los criterios de selección utilizados en este trabajo han mostrado un buen comportamiento del modelo LP3/LMOM, comparado con métodos más recientes, con la importante excepción de la alta varianza observada en los resultados.

AGRADECIMIENTO

Los trabajos recogidos en este artículo han sido desarrollados por el CEDEX, como parte del Proyecto de Investigación y Desarrollo «Modelación Numérica de Procesos Hidrológicos», agradeciendo la asistencia y aliento recibido por el autor en el Centro de Estudios Hidrográficos.

REFERENCIAS

- BENSON, M. A. (1968). Uniform Flood Frequency Estimating Methods for Federal Agencies. Water Resour. Res. 4(5), 891-908.
- CUNNANE, C. (1987). Review of Statistical Models for Flood Frequency Estimation Paper presented at Inst. Symp. on Flood Frequency and Risk Analysis. Baton Rouge, La., Publ. in Singh, V. P. (Ed.), Hydrologic Frequency Modelling, 49-05, Reidel Publ. Co., Dordrecht.
- HOSKING, J. R. M.; WALLIS, J. R., y WOOD, E. F. (1985). An Appraisal of the Regional Flood Frequency Procedure in the U.K. Flood Studies Report. Hydrol. Sci. J., 30(1), 85-109.
- LETTENMAIER, D. F., y POTTER, K. W. (1985). Testing Flood Frequency Estimation Methods using a Regional Flood Generating Model. Water Resour. Res., 21(12), 1903-1914.
- NERC (1975). Flood Studies Report. Nat. Environ. Res. Council, London, vols. 1-5, págs. 1100.
- ROSSI, F.; FIORENTINO, M., y VERSACE, P. (1984). Two Component Extreme Value Distribution for Flood Frequency Analysis. Water Resour. Res., 20(7), 847-856.
- USWRC (1981). Guidelines for Determining Flood Flow Frequency. Bulletin 17B. Water Resources Council, Washington.
- WALLIS, J. R.; MATALAS, N. C., y SLACK, J. R. (1974). Just a Moment. Water Resour. Res., 10(2), 211-219.
- WORLD METEOROLOGICAL ORGANIZATION (1989). Statistical Distributions for Flood Frequency Analysis. WMO-Operational Hydrology Report, núm. 33, pág. 73.