

UNIVERSITAT DE VALÈNCIA

Facultat de Física



Estimación del riesgo de lluvias fuertes  
en la zona de Murcia.

Análisis comparativo entre las distribuciones  
de Gumbel y SQRT-ET.

TESIS DE LICENCIATURA

Ramón Garrido Abenza

1996

UNIVERSITAT DE VALÈNCIA  
Facultat de Física  
Departament de Termodinàmica



Estimación del riesgo de lluvias fuertes  
en la zona de Murcia.  
Análisis comparativo entre las distribuciones  
de Gumbel y SQRT-ET.

Memoria presentada por  
RAMÓN GARRIDO ABENZA  
para optar al grado de  
Licenciado en Física.

*“Il est impossible que l'improbable n'arrive jamais”.*

**A Rosa María, Román y Dalia,  
y al siempre pequeño Bruno.**

AGRADECIMIENTOS.

D. VICENTE CASELLES MIRALLES, Profesor Titular de Física Aplicada de la Facultat de Física de la Universitat de València,

CERTIFICO:

Que la presente Memoria *Estimación del riesgo de lluvias fuertes en la zona de Murcia. Análisis comparativo entre las distribuciones de Gumbel y SQRT-ET*, ha sido realizada bajo mi dirección en el Departament de Termodinàmica de esta Facultat por Ramón Garrido Abenza, y constituye su Tesis de Licenciatura para optar al grado de Licenciado en Física.

Y para que así conste, en cumplimiento de la legislación vigente, firmo el presente certificado en Burjassot, a siete de junio de mil novecientos noventa y seis.



Vicente Caselles

## AGRADECIMIENTOS.

Deseo expresar mi agradecimiento a las personas que han facilitado la ejecución de este trabajo.

En primer lugar al Dr. Vicente Caselles, del Departamento de Termodinámica de la Universidad de Valencia y director de este trabajo, por sus valiosas orientaciones y su actitud animadora.

Al Dr. Joaquín Meliá, también del Departamento de Termodinámica, quien despertó mi interés por el estudio de los fenómenos atmosféricos.

A Franciscó Cabezas, de la Confederación Hidrográfica del Segura, que propició mi primer contacto con la distribución SQRT-ET.

Al Dr. Antonio Viguera, del Departamento de Matemática Aplicada y Estadística de la Universidad de Murcia, por sus enseñanzas sobre métodos numéricos.

A mis compañeros del Centro Meteorológico Territorial de Murcia, a su director el Dr. Luis Sánchez, por la facilidades concedidas, y de un modo especial a Plácido García, Juan Manuel Martín y Francisco Martínez, por toda la ayuda prestada.

# ÍNDICE.

---

	<u>pág.</u>
<b>RESUMEN.</b> .....	1
<b>CAPÍTULO 1. INTRODUCCIÓN.</b> .....	3
1.1. Marco del trabajo y objetivo. ....	4
1.2. Concepto y mecanismo básico de lluvias fuertes. ....	12
1.3. Descripción de la zona de estudio. ....	15
1.4. Conclusiones parciales .....	16
<b>CAPÍTULO 2. METODOLOGÍA.</b> .....	18
2.1. Distribuciones de valores extremos. ....	19
2.1.1. El postulado de estabilidad. ....	19
2.1.2. El proceso de Poisson. ....	22
2.2. La distribución de Gumbel. ....	23
2.2.1. Orígenes de la distribución de Gumbel. ....	23
2.2.2. Métodos de estimación de parámetros de la distribución de Gumbel. ....	26
2.2.2.a. Método de los momentos. ....	26
2.2.2.b. Método de los mínimos cuadrados. ....	27
2.2.2.c. Método de Chow. ....	28

---

---

	<u>pág.</u>
2.2.2.d. Método de máxima verosimilitud. ....	30
2.2.2.e. Método de Kimball. ....	31
2.3. La distribución SQRT-ET o Raíz. ....	33
2.3.1. Origen de la distribución Raíz. ....	33
2.3.2. Estimación de parámetros de la distribución Raíz. Método de máxima verosimilitud. ....	35
2.4. Conclusiones parciales. ....	37
<b>CAPÍTULO 3. RESULTADOS Y DISCUSIÓN.</b> .....	<b>38</b>
3.1. Descripción de los datos utilizados. ....	39
3.2. Control de calidad de los datos. ....	39
3.3. Aleatoriedad de las series. ....	40
3.4. Bondad del ajuste general. ....	42
3.5. Fiabilidad de los máximos históricos. ....	49
3.6. Estabilidad de la distribución. ....	54
3.7. Conclusiones parciales ....	56
<b>CAPÍTULO 4. CONCLUSIONES.</b> .....	<b>57</b>
<b>BIBLIOGRAFÍA.</b> .....	<b>60</b>

---

## RESUMEN.

---

Desde la Organización de Naciones Unidas, y sus organismos especializados, se ha llamado la atención sobre la necesidad de disponer de evaluaciones realistas de la probabilidad de ocurrencia de los desastres naturales en general y de los derivados de las precipitaciones intensas en particular. De ordinario, el problema suele abordarse recurriendo a la distribución estadística de Gumbel pero, dado que sus resultados no son suficientemente satisfactorios en general, en los últimos años han surgido algunas alternativas, entre ellas la distribución SQRT-ET o Raíz, biparamétrica como la anterior, que se sustenta sobre una sólida base teórica. El objetivo de este trabajo es realizar un estudio comparativo de la aplicabilidad de ambas distribuciones en la provincia de Murcia, expuesta históricamente al riesgo de inundaciones, comprobar las diferencias existentes entre ellas y establecer la fiabilidad que debe concederse a los resultados que se derivan del empleo de cada una de ellas.

El primer capítulo de este estudio es una introducción en la que se enmarca el trabajo a desarrollar, pasando revista a las propuestas que se han venido planteando para aproximarse a la solución del problema de la valoración del riesgo de lluvias intensas. Se establece el objetivo del trabajo y sus límites, esbozando las herramientas que se emplearán. Se sigue con un breve repaso del mecanismo meteorológico que origina las precipitaciones fuertes y de una terminología de adjetivación de las lluvias, para concluir con una concisa descripción climatológica de la vertiente mediterránea, centrándonos principalmente en la Región de Murcia.

El núcleo metodológico del trabajo se muestra en el capítulo 2 y radica en las teorías estadísticas de valores extremos, apoyadas fundamentalmente en el postulado de estabilidad o en el proceso de Poisson. Se esboza la deducción de las distribuciones de Gumbel y Raíz y se recopilan y explicitan los métodos de estimación de sus parámetros, dispersos con anterioridad en la bibliografía.

---

En el capítulo 3, dedicado a los resultados y su discusión, se consideran las series anuales de lluvias máximas diarias en la provincia de Murcia. Imponiendo criterios de control de calidad y aleatoriedad de los datos, se selecciona la serie correspondiente al aeródromo de San Javier. Se calcula las probabilidades asociadas a las distribuciones de Gumbel y Raíz, evaluándose la bondad del ajuste general proporcionado, la estabilidad de los parámetros y la magnitud de la extrapolación efectuada con el máximo histórico.

Independientemente del método de estimación paramétrica empleado, la utilización de la distribución de Gumbel resulta insatisfactoria, quedando cuestionada su validez general para dar cuenta del riesgo de lluvias intensas. Por el contrario, el empleo de la distribución Raíz se muestra como un recurso más apropiado para la valoración de ese riesgo, pues ofrece un buen ajuste y genera mayor verosimilitud y estabilidad que la de Gumbel, proporcionando, además, una explicación admisible para el máximo histórico.

Por último, en el capítulo 4, se recopilan y enumeran las principales conclusiones derivadas de este trabajo.

## CAPÍTULO 1.

### INTRODUCCIÓN.

---

Casi todos los años, y principalmente en otoño, algún episodio de lluvias fuertes en el área mediterránea española causa daños de importancia, provocando inundaciones devastadoras y cobrándose, en ocasiones, vidas humanas. El problema es común en extensas regiones de todo el mundo, por lo que, para reducir las pérdidas de vidas humanas, bienes y desórdenes económicos y sociales causados por las inundaciones y otros desastres naturales, la Asamblea General de las Naciones Unidas aprobó en 1989 una resolución por la que adoptaba el Decenio Internacional para la Reducción de los Desastres Naturales (1990-1999). El Secretario General estableció un Comité Científico y Técnico que estimó que las medidas prácticas para la reducción de desastres cubrían tres esferas: la evaluación de riesgos, la adopción de medidas preventivas y los pronósticos a corto plazo (Lahore, 1991). La segunda de ellas sólo puede surgir adecuadamente tras un desarrollo suficiente de la primera. En la tercera se registran importantes avances en las ciencias atmosféricas, mientras que la posibilidad de elaboración de otros pronósticos, como los relativos movimientos sísmicos, se encuentra poco desarrollada. Dentro de la primera esfera, la meta principal para el año 2000, a alcanzar por todos los países, consiste en haber establecido evaluaciones nacionales de los riesgos de los desastres naturales. A nivel nacional, también la Comisión Interministerial de Ciencia y Tecnología española, a través de su Programa Nacional sobre el Clima, ha llamado la atención sobre la necesidad de disponer de dichas evaluaciones.

Entre los riesgos naturales asociados a las precipitaciones intensas, y al margen de los de efecto inmediato, figura también el de la desertización. Esta amenaza se extiende también sobre amplias zonas del planeta, incluyéndose España, principalmente su vertiente mediterránea

---

y, de manera aún más llamativa, la región del Sureste, habiéndose impulsado el desarrollo de numerosos proyectos de investigación (EFEDA, MEDALUS, LUCDEME, etc.). La importancia de las lluvias torrenciales en los procesos de desertización es bien conocida, constituyendo, probablemente, el principal factor de pérdida de la capacidad productiva de los suelos en aquellas regiones con precipitaciones escasas e irregulares (López, 1993). Los estudios acerca del riesgo de lluvias fuertes podrán, por tanto, facilitar también la confección de cartografías detalladas del peligro de desertización.

### **1.1. Marco del trabajo y objetivo.**

La posibilidad de que en amplias zonas de la Tierra se produzcan aguaceros de magnitud extraordinaria, que desencadenen inundaciones u otro tipo de daños, obliga a tener en cuenta ciertas cuestiones fundamentales en el diseño de planes de emergencia, en el proyecto de obras hidráulicas, como embalses, desagües, etc. y, de manera más general, en el de todas aquellas estructuras cuyo fallo podría poner en peligro vidas o bienes materiales. Las cuestiones fundamentales suelen estar basadas en la consecución de la mejor relación posible entre la seguridad y los costes económicos y la adopción de decisiones de la manera más objetiva. Tampoco es desdeñable el papel que las precipitaciones intensas desempeñan en otro tipo de actividades humanas, como la agraria, especialmente en la conservación de suelos, y como agente erosivo y modelador del paisaje, interviniendo, no sólo en el proceso de separación de partículas del suelo, sino también en el transporte provocado por la escorrentía.

A la hora de abordar el tema de las predicciones de lluvias fuertes o de inundaciones, los principios de Hidrodinámica y Termodinámica con los que debería relacionarse el análisis hidrológico están suficientemente bien establecidos, pero el medio natural en el que éstos tendrían que aplicarse, la atmósfera, es tan extenso, irregular y sólo parcialmente conocido que, a menudo, elementos esenciales del mismo no son, o no pueden ser, medidos directamente. Esta

---

problemática justifica que un análisis hidrológico pueda admitir dos tipos fundamentales de enfoque: mientras que una aproximación de carácter determinista utilizaría las leyes que describen los procesos físicos, otra de carácter estadístico o probabilista analizaría las frecuencias de ocurrencia de las variables hidrológicas involucradas, tomando como referencia lo ocurrido en el pasado.

Por otro lado, el estado actual de conocimientos o desarrollo no permite la realización de predicciones meteorológicas detalladas, deterministas en mayor o menor grado, que sean fiables para un plazo de tiempo que alcance más allá de unos pocos días. Admitida esta realidad, se impone la necesidad de intentar también describir ciertos sucesos hidrológicos en plazos más largos mediante un enfoque probabilista de evaluación de los riesgos. Esta posibilidad estaría en consonancia con la primera esfera de actuación propuesta por el anteriormente citado Comité Científico y Técnico. De esta manera, el análisis de las frecuencias de ocurrencia de lluvias y sus intensidades, junto con la aplicación de modelos matemáticos de síntesis de datos hidrológicos, se nos puede revelar de gran importancia, entre otras aplicaciones, en el diseño y construcción de las estructuras de control del agua.

Los análisis de las frecuencias de lluvias y la realización de inferencias son realizados, comúnmente, mediante la aplicación de distribuciones estadísticas. Es bien sabido que, bajo ciertas condiciones, un gran número de distribuciones (como la binomial, la de Poisson, la de Student, etc.) tienden a comportarse como una distribución normal, por lo cual ésta goza de un puesto de privilegio en la teoría estadística, siendo la de más frecuente utilización en las ciencias naturales.

Sin embargo, este marco *normal*, principalmente adecuado en aquellos fenómenos en los cuales los valores medios juegan un papel dominante, es inapropiado a la hora de resolver problemas de valores extremos en general, y de las lluvias más fuertes en particular, ni siquiera de una manera mínimamente aproximada.

---

Ilustrémoslo brevemente, con un ejemplo numérico. Consideremos una muestra constituida por los valores extremos anuales de precipitación registrados en un lugar, en el que se mide diariamente, durante medio siglo. Supongamos ahora que la media y la desviación típica de la muestra, extraída de una población que admitimos que se distribuye normalmente, son suficientemente representativas y que sus valores son 66 y 50 mm, respectivamente. En estas condiciones, la probabilidad de que la precipitación diaria máxima en un año alcance un valor de 330 mm es de  $6 \cdot 10^{-8}$ , o, lo que es lo mismo, dicha precipitación se presentaría una vez cada 17 millones de años, por término medio. Dado que los datos de partida no son ficticios, sino que corresponden a lo registrado en un observatorio del litoral mediterráneo, se concluye de manera incuestionable la necesidad de abordar el problema mediante otros procedimientos. En palabras de Castillo (1988, p. 8) “la estadística de los valores medios, que rindió culto supremo a la ley normal, no es válida para resolver el problema de extremos”.

Existe un número muy grande de modelos susceptibles de ser usados en el análisis de frecuencias de las precipitaciones más fuertes, tanto de base empírica como de tipo teórico (Sevruk y Geiger, 1981), pero el problema dista de estar solucionado de manera satisfactoria.

Los métodos empíricos renuncian a cuestionarse los fundamentos teóricos. Los ejemplos más frecuentes consisten en persistir en la utilización de la distribución de probabilidad normal, o, alternativamente, en la de la distribución *Gamma*, aplicando alguna transformación a los datos originales, generalmente logarítmica, con el propósito de mejorar los resultados. Las posibilidades de transformación son infinitas (Plate, 1979) y, aunque, en ocasiones, los resultados son satisfactorios, los procedimientos son demasiado casuísticos y no proporcionan una perspectiva global ni un mayor conocimiento de los fenómenos.

En oposición a esos métodos, es posible desarrollar unas teorías genéricas de análisis de los extremos de precipitación, mejor fundamentadas, desde varios puntos de vista, alternativos y complementarios, en función del rango de aplicación, caracterizado por el orden del período

---

medio de retorno de los extremos. Suele distinguirse entre los tres métodos de análisis a los que nos referiremos a continuación (W. M. O., 1983a).

En primer lugar, el análisis basado en las series de máximos anuales, con un rango comprendido entre  $10^1$  y  $10^3$  años, se caracteriza por el uso de un dato anual y es el utilizado más comúnmente debido, por un lado, a la facilidad de obtención de los datos y, por otro, a las bases teóricas, relativamente sencillas, que permiten la extrapolación más allá del alcance temporal de las observaciones.

El segundo tipo de análisis es el basado en las series de duración parcial, de rango entre  $10^{-1}$  y 10 años, y en él se considera todos los valores que superan un umbral, independientemente de la fecha en la cual se registraron. La teoría para la extrapolación temporal es complicada, debido a la posible autocorrelación de los sucesos, los cuales pueden sucederse directamente uno a otro en el tiempo. La elección del umbral introduce un nuevo parámetro en la descripción y puede ser objeto de controversia. Por otro lado, en este tipo de análisis se puede estar mezclando sucesos de origen físico muy diverso.

Por último, la modelización de series temporales, cuyo rango oscila entre  $10^{-3}$  y  $10^{-1}$  años, se basa en teorías estocásticas aplicadas a bases observacionales diarias u horarias, en función de la duración de los episodios lluviosos. Requieren una gran resolución temporal de las observaciones, no siempre disponible.

Por las razones apuntadas anteriormente, los métodos basados en las series de máximos anuales son los más usados y a ellos nos dedicaremos en este trabajo. El principal inconveniente que presentan es el hecho de que, al estar cada año representado por un solo suceso, cabe la posibilidad de que algún segundo máximo anual, necesariamente excluido, sea superior al máximo de otro año, el cual sí sería incluido. Los resultados podrían depender de la definición utilizada para el año (civil, hidrológico, agrícola, etc.). Sin embargo, dado que es habitual que los máximos pluviométricos de un lugar se registren en la misma época del año, el problema no

---

reviste una importancia relevante, pudiéndose suponer que todos los sucesos seleccionados están causados físicamente de manera bastante similar.

Admitido este tipo de análisis, la cuestión fundamental radica en el tipo de distribución de probabilidad más recomendable. Aunque se suele aconsejar que se tenga en consideración aspectos como la intuición y la facilidad de cálculo, éstos no deben de condicionar la selección.

Desde el punto de vista matemático, la respuesta se encuentra en la teoría de valores extremos, la cual puede decirse que comienza con Pierce en 1852, en el tema de la observación astronómica, se desarrolla en la década de los años 20, con los trabajos de Fréchet (1927, citado, entre otros, por Gumbel, 1958; Sevruk y Geiger, 1981 y Kotz *et al.*, 1983) y Fisher y Tippet (1928, *ídem*), y alcanza la madurez en 1958 al publicar Gumbel la obra más trascendental (*Statistics of extremes*), donde ya se incluye las distribuciones de valores extremos más conocidas: la de tipo I (o de Gumbel), la de tipo II (o de Fréchet) y la de tipo III (o de Weibull), así como la distribución General de Valores Extremos, establecida por Jenkinson (1955), que engloba a las tres anteriores.

A lo largo de los últimos 50 años, se han multiplicado las aplicaciones prácticas de las distribuciones de valores extremos, tanto en el estudio de fenómenos meteorológicos e inundaciones, como en problemas de ingeniería. En este último caso, con frecuencia se han reproducido leyes que anteriormente habían sido establecidas de una manera meramente experimental.

Sin duda alguna, como veremos, la distribución de valores extremos de Gumbel, también denominada doble exponencial, ha sido la más universalmente empleada, sobre todo en el análisis de extremos de precipitación. Se trata de una distribución de dos parámetros, los cuales han de ser estimados a partir de muestras de datos, y que supone un caso particular de la distribución de Jenkinson (1955), triparamétrica. Los trabajos de análisis de la precipitación en los que se utiliza la distribución de Gumbel son muy numerosos y, a veces, muy ambiciosos en

---

cuanto a cobertura territorial. En este sentido, limitándonos sólo a España, cabe destacar los célebres trabajos de Elías (1963) y, más recientemente, el Atlas Nacional de España, editado por el Instituto Geográfico Nacional (1992). En ocasiones, también se ha utilizado para la elaboración de climatologías de la intensidad de la lluvia propiamente dicha, definida como volumen de agua por unidad de tiempo y superficie, previas a la construcción de las curvas Intensidad-Duración-Frecuencia, ampliamente utilizadas en hidrología (Redaño *et al.*, 1986).

En un informe de la Organización Meteorológica Mundial (Sevruk y Geiger, 1981) se pasa revista a 94 trabajos realizados en todo el mundo sobre el análisis de los máximos anuales de precipitación. La primera conclusión que se obtiene es que la distribución de Gumbel se ha empleado, aproximadamente en la mitad de las ocasiones (47 %), seguida de la distribución *Log-normal* (20 %). En general, se encuentran argumentos para recomendar el uso de la primera (existencia de una base teórica, buenas estimaciones de los extremos de precipitación, aunque con fracasos cuando el coeficiente de variación, es decir, la razón entre la desviación típica y la media, pasa del 50%), frente al de la segunda (ausencia de justificación teórica del uso de las transformaciones logarítmicas y peores resultados proporcionados en conjunto).

La distribución de Gumbel, cuando proporciona resultados satisfactorios, es también capaz de explicar porqué la utilización de la distribución *Log-normal* pudiera resultar adecuada. Es posible comprobar (Gumbel, 1958) que cuando el coeficiente de variación toma aproximadamente el valor 0,36, la descripción proporcionada por ambas distribuciones es muy similar. En esto podría estar basado el empleo tan extendido de la distribución *Log-normal*, a pesar de carecer de fundamentos teóricos mínimamente adecuados. Algo parecido sucedería con el empleo de la distribución *Gamma*, para valores del coeficiente de variación en torno a 0,57.

Según el citado informe, las ocasiones en las que se ha recurrido a la distribución generalizada de valores extremos de Jenkinson junto con aquellas en las cuales se ha recurrido al empleo mixto de las distribuciones de los tipos I y II representan un 6% del total, mientras que

---

las de tipo III sólo se han utilizado en el 1 % de los casos. Otro conjunto de distribuciones diversas alcanzaría un 10 % de los trabajos.

La conclusión general del informe es que no parece haber una distribución que sea satisfactoria en todos los casos, aunque una posible solución momentánea pueda pasar por el uso de la distribución de valores extremos generalizada. Sin embargo, la recomendación final es que se impulse la investigación apoyada en fundamentos físicos, fomentando la colaboración entre matemáticos, físicos e hidrólogos, pues “se necesita una solución urgentemente en muchos países” (Sevruk y Geiger, 1981, p. 10).

Los insatisfactorios resultados a que ha conducido, en ocasiones, el análisis basado en el empleo de la distribución de Gumbel no siempre han podido ser explicados con el argumento de una inadecuada estimación de los parámetros (Garrido, 1992) o de una base de datos corta (Etoh *et al.*, 1987). También, en más casos de los admisibles, ha puesto de manifiesto la existencia de puntos aislados (*outliers*), es decir, datos de la muestra que se apartan considerablemente de la ley de distribución de probabilidad que se supone rige a la población, y que aparentemente son inconsistentes con el resto de los datos. Esto puede suceder, por ejemplo, con los máximos históricos de la serie o en el caso de otras lluvias excepcionalmente grandes. En general, la existencia de estos puntos aislados puede deberse a varios motivos:

- i) La variedad de causas físicas que originan la lluvia, tales como tormentas aisladas, complejos convectivos, frentes, precipitaciones orográficas, etc. En teoría, las observaciones deberían de realizarse bajo condiciones físicas idénticas.
  - ii) La no aleatoriedad de las series climatológicas, es decir, la existencia de tendencias a largo plazo o periodicidades en el clima. Implícitamente, siempre se está suponiendo que la población de la cual se extraen los máximos debe permanecer constante en el tiempo, es decir, sin que tengan lugar cambios climáticos durante el período de registro de las observaciones ni sean esperados en el futuro.
-

- iii) La incorrección de los datos, tanto por causas técnicas diversas, tales como deformaciones debidas al campo de viento (con pérdidas variables, pero en el orden del 5% en las cantidades de precipitación medidas con un pluviómetro Hellmann) o pérdidas por evaporación o por humectación de los elementos sensores, todas ellas de poca importancia en la medida de extremos de precipitación, como por errores humanos.
- iv) La aplicación a los datos de funciones de distribución de probabilidad erróneas o de procedimientos de estimación de sus parámetros inadecuados.

En este trabajo ignoraremos las dos primeras causas, que caerían dentro del campo de la Climatología Dinámica y de los estudios acerca del Cambio Climático, respectivamente. Efectuaremos controles de calidad de los datos que eliminen, en parte, las fuentes de error, de manera que nuestro objetivo se centra fundamentalmente en la última de las causas enumeradas, es decir, trataremos de establecer si la utilización de leyes de distribución, como pudiera ser la Gumbel en los casos anteriormente citados, es o no apropiada. Por supuesto, antes de llegar a un hipotético rechazo de la distribución de Gumbel, habría que hacer un repaso exhaustivo de los métodos de estimación de los parámetros, para que el fallo descriptivo pueda ser achacable con propiedad al empleo de dicha distribución.

Aunque es de esperar que los resultados mejorasen, respecto de los proporcionados por la distribución de Gumbel, si pasásemos a utilizar la distribución triparamétrica de Jenkinson, en virtud del principio de simplicidad, antes de recurrir a una distribución triparamétrica habría que agotar las posibilidades descriptivas de las biparamétricas. Aunque la utilización de tres parámetros sea capaz de proporcionarnos un mejor ajuste de los datos, del mismo modo que un polinomio de segundo grado siempre se aproxima más a los datos experimentales que una recta, cabe la posibilidad de que suponga una peor descripción cualitativa del fenómeno, como sería en el caso de cualquier ley física de tipo lineal. Además, si estamos interesados especialmente en los valores más altos de la precipitación para una posible extrapolación temporal más allá de la longitud de los registros históricos, como en la cola final de la distribución estadística es

---

difícil que haya una base experimental, no hay que descartar que la extrapolación basada en un modelo triparamétrico pudiera llegar a ser mucho peor que la de uno biparamétrico. Por último, con el incremento del número de parámetros aumenta la fluctuación o dependencia de la estimación respecto de la muestra estadística, ampliándose los márgenes de error para intervalos de confianza fijos y disminuyendo la estabilidad de la distribución. No parece que estos aspectos se tengan suficientemente en cuenta en algunas investigaciones, incluso relativamente recientes, en las cuales se llega a trabajar con mezclas de distribuciones de valores extremos que necesitan del concurso de hasta 6 parámetros (Upadhyay y Surinder, 1990).

La búsqueda de una distribución biparamétrica alternativa llevó a Etoh *et al.* (1987), tras mostrar las deficiencias de la distribución de Gumbel al estudiar las lluvias extremas en Japón, a proponer el empleo de otra distribución, a la que denominaremos SQRT-ET o Raíz, con un sólido soporte teórico y experimental. En España, también se ha aplicado esta distribución alternativa, con resultados muy alentadores (Garrido y Caselles, 1995), los cuales se pretende confirmar ahora, ampliando la base de datos histórica.

Por ello, en este trabajo, nos ocuparemos únicamente de la distribución de Gumbel y de la distribución Raíz, estableciendo la bondad de cada una de ellas mediante índices numéricos objetivos, definidos sobre la base de medidas de ajuste general, de fiabilidad de los puntos aislados y de estabilidad en la estimación de los parámetros.

## **1.2. Concepto y mecanismo básico de lluvias fuertes.**

Básicamente, el mecanismo meteorológico de generación de lluvias fuertes consiste en la existencia en la atmósfera de intensas corrientes ascendentes de aire cálido y húmedo, propiciado frecuentemente por un largo recorrido marítimo del viento, desde los niveles bajos

---

hasta los altos, formando nubes de gran desarrollo vertical, como cumulonimbos o sistemas convectivos.

Para la creación de esa corriente ascendente es necesaria la convergencia del aire cálido y húmedo en los niveles bajos, sin que existan inversiones de temperatura. Este marco se refuerza si existe divergencia en los niveles altos, lo que actuaría como un mecanismo de succión del aire, y también si hay un forzamiento del ascenso del aire debido a obstáculos orográficos.

Una gran parte de la existencia de problemas derivados de estas precipitaciones, como inundaciones, avenidas, etc., dependen de las condiciones del terreno y de las estructuras de control de volumen de agua existentes. Por ello, la adjetivación de las lluvias, atendiendo a su intensidad, cabe efectuarla desde dos perspectivas distintas, una de tipo absoluto, ligada estrechamente a las condiciones meteorológicas, y otra de tipo relativo, relacionada con la climatología y, en conexión con ella, a las consecuencias de la precipitación.

Limitándonos a las precipitaciones de carácter líquido, el primer punto de vista es función, únicamente, de la intensidad de la lluvia, entendida como la altura de agua medida en la unidad de tiempo sobre una superficie unitaria. La unidad de tiempo puede oscilar desde 10 minutos (según la definición de la Organización Meteorológica Mundial para el cálculo de intensidades máximas) hasta varios días, aunque lo más habitual pueda ser la hora (de utilización en labores de predicción meteorológica) o el día (para la medida en la mayoría de las estaciones meteorológicas).

Para la elaboración de las predicciones meteorológicas, un *Manual de Términos Meteorológicos* (I. N. M., 1992) define las lluvias *fuertes* como “aquellas cuya intensidad es mayor que 15 y menor o igual que 30 mm/h”, las *muy fuertes* como aquellas cuya “intensidad es mayor que 30 y menor o igual que 60 mm/h”, reservando el término de *torrenciales* “para intensidades mayores de 60 mm/h”. Sin embargo, salvo en observatorios principales o

---

automatizados, no se dispone de información horaria, de manera que la mayor parte de los datos disponibles corresponden a precipitaciones diarias, especialmente si se desea trabajar con series temporales largas. Estas definiciones, aun siendo objetivas, no reflejan suficientemente la posibilidad de que surjan problemas asociados a la lluvia, pues estos dependen, como dijimos, de las condiciones del terreno y de las estructuras de control del agua.

El segundo punto de vista en la adjetivación de las lluvias puede hacer intervenir, de un modo indirecto, las consecuencias. Como primera aproximación, según el diccionario de la Real Academia Española (R. A. E., 1992), las lluvias *torrenciales* serían aquellas que produjeran una “corriente o avenida impetuosa de aguas”, siendo bastante difícil de establecer si se cumple o no esa condición. Sin embargo, es posible analizar, en términos objetivos, las frecuencias de ocurrencia de la lluvia en cada lugar, otorgando una probabilidad a cada intensidad. Relacionado con esa probabilidad estaría el concepto de período de retorno, o intervalo de tiempo que, por término medio ha de transcurrir para que acontezca una determinada intensidad de precipitación. De este modo, en ingeniería hidráulica suele utilizarse un período de retorno de 500 años en el diseño de algunas estructuras, de manera que podrían considerarse como *fuertes* o *muy fuertes* a aquellas precipitaciones que se acercasen al correspondiente valor de diseño, reservando el término de *torrenciales* para aquellas que, con una probabilidad de ocurrencia anual inferior a 1/500, lo superasen, poniendo en grave riesgo la funcionalidad de la estructura. En otro tipo de aplicaciones el período de retorno se fija de acuerdo con el riesgo que se considera asumible.

De este modo, en zonas especialmente preparadas ante el riesgo de inundaciones, por la solidez de sus estructuras, una precipitación fuerte o torrencial será más intensa, en términos absolutos, que en regiones más desprotegidas. Este criterio relativo responde en mayor medida que el anterior a la idea de prevención de desastres naturales y es generalizable tanto a otro tipo de fenómenos meteorológicos adversos, entre los que cabría destacar los vientos fuertes, como a otros riesgos naturales, como los movimientos sísmicos.

---

### **1.3. Descripción de la zona de estudio.**

El clima de la vertiente mediterránea española, la cual puede considerarse que se extiende “desde el Ampurdán hasta Almería” (Font, 1983, p. 175) está determinado fundamentalmente por su cercanía al Mediterráneo, un mar relativamente cálido y prácticamente cerrado, que se encuentra rodeado de tierras de relieve accidentado. Esta configuración proporciona a la región unas características climáticas esencialmente benignas, con veranos cálidos e inviernos templados, con escasas heladas. La cercanía al mar provoca un aporte de humedad y tiene un efecto termorregulador, favoreciendo la generación de brisas. Sin embargo, la región no escapa a violentos temporales de viento, ocasionalmente repentinos, ni a la amenaza de la desertización.

Desde el punto de vista de la precipitación, una de las características más significativas es la existencia de un máximo otoñal, cuando las aguas marítimas están más cálidas, con otro máximo, en ocasiones secundario, en la primavera, y la extrema sequedad de los veranos. En general, la precipitación presenta un gradiente latitudinal positivo, salvo las anomalías debidas al relieve, que pueden provocar valores anuales medios por encima de 1.000 mm. La zona más seca, con precipitaciones inferiores a 300 mm anuales, se localiza en la zona del Sureste. En toda la vertiente la variabilidad interanual de la precipitación es muy alta.

Pero hay otra característica, más conocida por sus consecuencias devastadoras, y que es consecuencia directa de que en la región mediterránea se puedan reunir las circunstancias del marco de generación de lluvias fuertes que vimos en el apartado anterior, principalmente en otoño, cuando el viento sopla de Levante, con largo recorrido marítimo, y las condiciones en los niveles medios y altos de la atmósfera son las apropiadas. Los fuertes aguaceros que se desencadenan hacen que, además de la gran variabilidad interanual, la variabilidad diaria sea excepcional. En ocasiones, la cantidad de agua recogida en un día puede llegar a ser superior al

---

valor medio anual. Como ejemplo más llamativo cabe destacar los 817 mm recogidos en Oliva (Valencia) en un sólo día (I. G. N., 1992).

Tanto la variabilidad del régimen pluviométrico como el riesgo de desertización se acentúan en la Cuenca del Segura, fundamentalmente en la provincia de Murcia y en comarcas colindantes de las de Alicante y Almería, donde largos períodos de tiempo seco suelen acabar bruscamente con inundaciones catastróficas. Esta zona se sitúa geológicamente dentro del ámbito de las Cordilleras Béticas, que la protegen de los vientos del suroeste y del oeste y, junto con las últimas estribaciones del Sistema Ibérico, de los del norte. Está comunicada con la meseta de la Mancha a través del portillo de la Cordillera Subbética, permaneciendo sólo abierta al mar por el sureste. Al margen de la latitud y de las oscilaciones de la circulación general de la atmósfera, esta configuración orográfica tan aislante es determinante para la existencia o no de la precipitación. Los vientos atlánticos, cuando llegan, suelen haber perdido su humedad característica, tras desecarse al atravesar la Península Ibérica, de manera que a sotavento de las Sierras de Cazorla y del Segura, ya en la provincia de Murcia, no suelen proporcionar más que precipitaciones de escasa cuantía. Han de ser los vientos de Levante los que originen, cuando coinciden con alguna o varias circunstancias favorables, las lluvias más fuertes. La historia de Murcia, como la de otras regiones mediterráneas, está jalonada por esas lluvias.

#### **1.4. Conclusiones parciales.**

Hemos visto cómo el problema de las lluvias fuertes preocupa a la Organización de Naciones Unidas y a su organismo especializado, la Organización Meteorológica Mundial, debido a las repercusiones humanas, sociales y económicas del fenómeno. Las regiones expuestas al desencadenamiento de mecanismos de generación de fuertes precipitaciones, como la Mediterránea en general o la de Murcia en particular, deben contar con análisis adecuados de

---

vulnerabilidad, que permitan el desarrollo de planes mitigadores de los efectos desastrosos de tales lluvias.

La evaluación de los riesgos de ocurrencia de esas lluvias, en plazos de tiempo largos, ha de fundamentarse en la teoría estadística de valores extremos y, por motivos prácticos, en el uso de series de máximos anuales. La distribución estadística más comúnmente utilizada, la de Gumbel, no siempre ha proporcionado una descripción adecuada de la realidad, por lo que han surgido trabajos con el objeto de encontrar una alternativa, como pudiera ser la distribución Raíz. Establecer una comparación entre la aplicabilidad de ambas, en el ámbito de la Región de Murcia, constituirá el objetivo central de este trabajo.

## CAPÍTULO 2.

### METODOLOGÍA.

---

El núcleo metodológico del análisis de frecuencias de las precipitaciones intensas radica en las teorías estadísticas de valores extremos. Estas teorías dan lugar a diversas distribuciones de valores extremos, en función, fundamentalmente, de las hipótesis que se establecen para la población de la cual se extraen los máximos, siendo, como vimos, la distribución de Gumbel la de aplicación más extendida para el caso de datos hidrometeorológicos.

En aquellas situaciones en las que resultan insatisfactorios los resultados proporcionados por esa distribución, es preciso asegurarse de que el fracaso es inherente a ella, y no al método de estimación de sus parámetros, los cuales pueden ofrecer resultados significativamente distintos entre sí. En todo caso, siempre ha de prestarse una atención especial al método de máxima verosimilitud, recomendado por la Organización Meteorológica Mundial, a pesar de la complejidad de cálculos que pueda conllevar, puesto que es la manera de extraer la máxima información de unas series de datos que, por limitaciones naturales, no pueden ampliarse todo lo que quisiéramos.

En la búsqueda de distribuciones alternativas, dentro del marco biparamétrico, habrá que intentar seleccionar aquellas que, como la distribución Raíz, posean una sólida base teórica y el mayor significado físico posible.

---

## 2.1. Distribuciones de valores extremos.

En el capítulo anterior, entre otros aspectos, hemos esbozado una panorámica general de las distribuciones de valores extremos, fijando nuestro interés en las distribuciones de Gumbel y Raíz. Sin embargo, conviene cerrar el marco en el que se inscribe este trabajo estructurando con algo más de detalle, aunque sea brevemente, las distribuciones de valores extremos, así como sus fundamentos teóricos.

Hay dos procedimientos estandarizados de generación de distribuciones de valores extremos. El desarrollo del primero de ellos, basado en el postulado de estabilidad, está prácticamente concluido, al haberse encontrado la solución general de dicho postulado, en el cual no interviene la forma de la distribución inicial de probabilidades, de la cual se iría extrayendo cada máximo, anual por ejemplo. La otra alternativa, que en ocasiones puede conducir a las mismas soluciones, recurre al proceso de Poisson (Martín, 1981; Castillo, 1988), pudiendo generar tantas distribuciones de valores extremos como distribuciones iniciales de la precipitación, de las que extraemos los máximos, seamos capaces de concebir. Al margen de estos dos procedimientos existen multitud de derivaciones particulares de distribuciones de valores extremos, imponiendo ciertos requisitos a las distribuciones iniciales, alcanzándose, con frecuencia, resultados completamente equivalentes.

### 2.1.1. *El postulado de estabilidad.*

El postulado de estabilidad, que da origen a tres tipos de distribuciones de valores extremos, conocidos como las tres asíntotas de Fisher-Tippet, fue establecido por Fréchet en 1927. Consideremos  $n$  muestras independientes de tamaño  $N$  tomadas de la misma población, cada una de ellas con un valor máximo. La distribución del máximo de las  $nN$  observaciones tenderá asintóticamente, para valores muy grandes de la variable, a la misma expresión que la distribución del valor máximo en las muestras de tamaño  $N$ , con tal de que dicha asíntota exista.

---

Puesto que una transformación lineal no debe cambiar la forma de la distribución, la probabilidad de que el valor máximo sea inferior a  $x$ , es decir, la probabilidad  $F^N(x)$  de que las  $N$  observaciones sean inferiores a  $x$ , debe ser igual a la probabilidad de una función lineal de  $x$ , es decir:

$$F^N(x) = F(a_N x + b_N) \quad (1)$$

siendo los parámetros  $a_N$  y  $b_N$  dos parámetros dependientes de  $N$ , representando  $F(x)$  la función de distribución de probabilidad acumulada (f. d. a.), o sea, la probabilidad de que un valor sea inferior a  $x$ . Esa expresión se conoce con el nombre de *postulado de estabilidad* (Gumbel, 1958) o *ecuación funcional de Fisher-Tippet* (Essenwanger, 1986). En 1943 Gnedeko demostró que las únicas distribuciones que satisfacen dicho postulado son las siguientes:

1ª) *Primera asíntota* ( $a_N=1$ ):

$$F(x) = \exp \{ - \exp [ - a(x-u)] \} \quad (2)$$

con  $a$  y  $u$  parámetros positivos, no estando la variable  $x$  limitada. A la distribución definida por la ec. (2) se le suele denominar también *distribución de Gumbel, de tipo I* o *doble exponencial*.

2ª) *Segunda asíntota* ( $a_N \neq 1$  y  $F(0)=0$ ):

$$F(x) = \exp \left\{ - \left( \frac{v}{x} \right)^k \right\} \quad (3)$$

con  $v$  y  $k$  parámetros positivos y  $x \geq 0$ . A veces se la reconoce con el nombre de *distribución de tipo II*.

3ª) *Tercera asíntota* ( $\alpha \neq 1$  y  $F(0) \neq 0$ ):

$$F(x) = \exp \left\{ - \left( \frac{x}{\nu} \right)^k \right\} \quad (4)$$

con  $k$  positivo,  $\nu$  negativo y  $x \leq 0$ . Esta sólo difiere de la anterior en el signo del parámetro  $\nu$  y de la variable  $x$ . Se la conoce también como *distribución de tipo III*.

Tanto la segunda como la tercera asíntota pueden derivarse de la primera mediante una transformación logarítmica. Para aquellas también existen sendas versiones generalizadas de tres parámetros, introduciendo un umbral, denominándoseles distribuciones de Fréchet y de Weibull, respectivamente, que incluyen a la primera asíntota como caso particular. El conjunto de las tres asíntotas puede ser escrito en la forma general triparamétrica de Jenkinson (1955), que es la solución general del postulado de estabilidad:

$$F(x) = \exp \{ - [ 1 - k\alpha(x-u) ]^{1/k} \} \quad (5)$$

con el parámetro  $\alpha$  positivo. Cuando  $k > 0$  ( $k < 0$ ) existe un límite superior (inferior) y se abarca la tercera (segunda) asíntota, mientras que cuando  $k \rightarrow 0$  nos lleva a la primera asíntota. Posteriormente, Jenkinson (1969 y 1975) sugirió un procedimiento iterativo para los estimadores máximo-verosímiles de los tres parámetros. Por otro lado, la posible existencia de ese límite superior está propiciando su aplicación para la estimación de la precipitación máxima probable (Nobilis *et al.*, 1991) y, en otros campos de la climatología, para la detección de un posible cambio climático (Almarza y López, 1995).

Estas distribuciones asíntóticas pueden obtenerse también mediante otras formas alternativas. La primera surgiría para distribuciones iniciales de tipo exponencial, cuya definición detallaremos posteriormente, la segunda para las de tipo Pareto y la tercera para las que verifican un conjunto de hipótesis acerca de las derivadas de la función densidad de probabilidad (f. d. p.) establecidas por Von Mises (Gumbel, 1958).

---

### 2.1.2. El proceso de Poisson.

El método de obtención de distribuciones de valores extremos basado en el proceso de Poisson es más intuitivo y posee un mayor significado físico. Supongamos un proceso en el cual, a lo largo del tiempo  $t$ , se va produciendo una serie de sucesos, siendo  $N(t)$  la función que los cuenta. Se dice que es un *proceso de Poisson* (Martín, 1981, Kotz *et al.*, 1983) cuando se cumple que:

- i) La variable  $N(t)$  es estacionaria, es decir, la probabilidad  $P_k(t)$  de que en un intervalo de tiempo  $t$  se hayan producido  $N(t) = k$  sucesos sólo depende del intervalo de tiempo  $t$ .
- ii)  $N(t)$  es independiente del número de sucesos ocurridos en cualquier intervalo anterior a  $t$ , es decir, la historia del proceso en un tiempo anterior no influye en el intervalo  $t$ .
- iii) La probabilidad  $P_f(dt)$  de que en un intervalo  $dt$  ocurra un suceso es proporcional a  $dt$  (*hipótesis de regularidad*), es decir,  $P_f(dt) = \lambda dt$ , con la constante de proporcionalidad  $\lambda > 0$ .

Estas hipótesis conducen a una ecuación diferencial para  $P_k(t)$  que tiene por solución:

$$P_k(t) = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \quad (6)$$

El parámetro  $\lambda$  representa la frecuencia media de sucesos por unidad de tiempo o probabilidad del suceso. Si consideramos, por simplificar, un período de tiempo unitario, por ejemplo un año, la probabilidad  $P(k)$  de que suceso se repita  $k$  veces será:

$$P_k(1) \equiv P(k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (7)$$

siendo  $\lambda$  la frecuencia media o probabilidad del suceso. Sea  $H(x)$  la f. d. a. inicial, la cual representa la probabilidad de que un valor aleatorio no supere el valor  $x$ . Por tanto, para  $k$  sucesos tendremos que la probabilidad de que ninguno de ellos supere el valor  $x$  será  $[H(x)]^k$ . Por consiguiente, la f. d. a. de los máximos será (Castillo, 1988):

$$F(x) = \sum_{k=0}^{\infty} P(k) [H(x)]^k = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} [H(x)]^k \quad (8)$$

donde hemos sustituido el valor de  $P(k)$  dado por (7). Esta expresión puede ponerse también, sacando el parámetro  $\lambda$  del sumatorio y utilizando el desarrollo en serie de Taylor de la función exponencial, como:

$$F(x) = e^{-\lambda [1 - H(x)]} \quad (9)$$

que nos da de manera universal la distribución de los máximos en función del parámetro  $\lambda$ , el cual se interpreta físicamente como el número medio de precipitaciones por año, y de la distribución inicial, la cual puede responder a múltiples formas, pero que siempre debería tener un sentido físico o responder a una evidencia experimental.

## 2.2. La distribución de Gumbel.

### 2.2.1. Orígenes de la distribución de Gumbel.

La función de distribución de probabilidad de Gumbel, también denominada doble exponencial, primera asíntota de Fisher-Tippet o función de distribución de valores extremos de tipo I (Sevruk y Geiger, 1981), ha sido utilizada frecuentemente en climatología para determinar las probabilidades de ocurrencia de valores extremos de diversas variables, entre ellas la precipitación. Se trata de una función biparamétrica, para la que existen diversos

métodos de derivación, entre los cuales cabe citar, al margen del basado en el postulado de estabilidad, el fundamentado en el desarrollo en serie de la f. d. p., así como los métodos de Cramer y de Von Mises (Gumbel, 1958). En general, en ellos se supone, fundamentalmente, que las muestras de las observaciones, de las cuales se toma cada máximo (anual, por ejemplo) son:

- i) muy numerosas.
- ii) del mismo tamaño.
- iii) estadísticamente independientes.
- iv) se distribuyen de acuerdo a lo que se denomina una distribución  $h(x)$  de *tipo exponencial*, entendiéndose por tal aquella distribución ilimitada que asintóticamente, es decir, para grandes valores de la variable  $x$ , verifica que

$$\frac{-h'(x)}{h(x) [1 - H(x)]} = 1 + |\epsilon(x)| \quad (10)$$

siendo  $H(x)$  la correspondiente f. d. a., con

$$\lim_{x \rightarrow \infty} |\epsilon(x)| = 0 \quad (11)$$

Entre este tipo de distribuciones podríamos citar, por ejemplo, la exponencial, la *Gamma* o la logística. Así mismo, se supone que únicamente estamos interesados en los grandes valores de la variable aleatoria  $x$ .

Por homogeneidad con la derivación que efectuaremos posteriormente en el apartado 2.3.1. para la distribución Raíz, vamos a deducir formalmente la expresión de la distribución de Gumbel mediante la hipótesis de que la ocurrencia de las lluvias máximas responden a un proceso de Poisson, lo cual es tradicionalmente aceptado para fenómenos poco frecuentes y su

---

aplicación al caso de las precipitaciones está suficientemente contrastada (Carrera y Martínez, 1991).

Si suponemos que la distribución inicial  $H(x)$  es la distribución exponencial, es decir si:

$$H(x) = 1 - e^{-ax} \quad (15)$$

siendo la f. d. p.:

$$h(x) = a e^{-ax} \quad (16)$$

tendríamos mediante (9) que la f. d. a. de los máximos es:

$$F(x) = \exp \{ - \lambda \exp [-ax] \} \quad (17)$$

En esta deducción formal no ha sido necesario suponer que el tamaño de las muestras de las que se extrae cada máximo es constante, pero la restricción impuesta al tipo de distribución inicial, necesariamente exponencial, es más restrictiva que cuando, en otras derivaciones, sólo se exige que sea de *tipo exponencial*.

Esta distribución en los parámetros  $\lambda$  y  $a$  suele escribirse, haciendo  $u = \frac{\log \lambda}{a}$ , en la forma más conocida:

$$F(x) = \exp \{ - \exp [ - a(x-u) ] \} \quad (18)$$

de manera que la f. d. p. será:

$$f(x) = a e^{-a(x-u)} F(x) \quad (19)$$


---

siendo  $u$  un parámetro de posición y  $a$  un parámetro de escala. La estimación de estos parámetros puede efectuarse mediante distintos métodos, a partir de las muestras de datos disponibles. Esta estimación diversa puede provocar diferencias significativas en el cálculo de probabilidades asociadas a la distribución, por lo que resultará conveniente realizar un estudio comparativo de los resultados que proporcionan los distintos métodos de estimación. En el apartado siguiente se muestra los cinco métodos de estimación más universalmente aceptados (momentos, mínimos cuadrados, Chow, máxima verosimilitud y Kimball) y que serán a los que se recurrirá en este trabajo.

### ***2.2.2. Métodos de estimación de parámetros de la distribución de Gumbel.***

#### **2.2.2.a. Método de los momentos.**

El método de los momentos es el que presenta un cálculo más cómodo y quizás, debido a ello, sea uno de los más empleados. En él se atribuye a los parámetros valores deducidos de los momentos de primer y segundo orden de la muestra. En primer lugar, la media, o momento de orden uno respecto del origen es, por definición:

$$\bar{x} = \int_{-\infty}^{+\infty} x f(x) dx \quad (20)$$

de donde, sustituyendo la expresión de  $f(x)$  y realizando el cambio de variable  $z=e^{-a(x-u)}$  se llega (Martínez, 1986) a que:

$$\bar{x} = u - \frac{1}{a} \int_0^{\infty} \log z e^{-z} dz \quad (21)$$

Esa integral, con el signo cambiado, no es más que la constante de Euler  $\gamma = 0,577\ 215\ 7\dots$ , obteniéndose, en definitiva, que:

$$\bar{x} = u + \frac{\gamma}{a} \quad (22)$$


---

La varianza, o momento de orden dos respecto de la media, es:

$$\sigma^2 = -\bar{x}^2 + \int_{-\infty}^{+\infty} x^2 f(x) dx \quad (23)$$

Realizando el mismo cambio de variable anterior e integrando se llega a que:

$$\sigma^2 = \frac{\pi^2}{6} + \gamma^2 \quad (24)$$

Las ec. (22) y (24) constituyen un sistema que nos proporciona los valores de los parámetros de la distribución de Gumbel en la forma:

$$a = \frac{\pi}{\sigma \sqrt{6}} \quad (25)$$

$$u = \bar{x} - \frac{\gamma}{a} = \bar{x} - \frac{\gamma \sigma \sqrt{6}}{\pi} \quad (26)$$

siendo  $\bar{x}$  y  $\sigma$  la media y la desviación típica insesgada, respectivamente, de una muestra de  $n$  valores máximos.

### 2.2.2.b. Método de los mínimos cuadrados.

La estimación basada en el método de mínimos cuadrados parte de la consideración de que la ec. (18) puede expresarse también como:

$$-\log [-\log F(x)] = ax - au \quad (27)$$

y puede representar la ecuación de una recta de pendiente  $a$  y ordenada en el origen  $-au$ , siendo  $x$  la variable independiente e  $y = -\log [-\log F(x)]$  la dependiente. Si ordenamos los datos  $x_i$  por orden creciente  $x_1^o \leq x_2^o \leq \dots \leq x_m^o \leq \dots \leq x_n^o$ , aceptamos (Sneyers, 1975) que la estimación empírica más conveniente para  $F(x_m^o)$  es el valor  $m/(n+1)$ , es decir:

$$y_m = - \log \left( - \log \frac{m}{n + 1} \right) \quad (28)$$

e imponemos que la suma de los cuadrados de los errores, en la forma  $\sum_{m=1}^n (y_m - ax_m^o + au)^2$ , sea mínimo cuadrática, es posible obtener, en definitiva (Martín, 1981):

$$\alpha = \frac{n \sum_{m=1}^n x_m^o y_m - \sum_{m=1}^n x_m \sum_{m=1}^n y_m}{n \sum_{m=1}^n x_m - \left( \sum_{m=1}^n x_m \right)^2} = \frac{\text{Cov} (x^o, y)}{\sigma_x^2} \quad (29)$$

$$u = \bar{x} - \frac{\gamma_n}{\alpha} \quad (30)$$

donde  $\gamma_n$  es el valor medio de los  $n$  valores de  $y$ , habiéndose expresado también el coeficiente  $\alpha$  como el cociente dado por la covarianza de  $x^o$  sobre  $y$  entre la varianza de  $x$ , ambas sesgadas o insesgadas. Cuando  $n$  tiende a infinito,  $\gamma_n$  tiende al valor de la constante de Euler  $\gamma$ , de manera que la expresión para la estimación de  $u$  coincide con la obtenida en la ec. (26) mediante el método de los momentos.

### 2.2.2.c. Método de Chow.

Este método (Chow, 1954, citado por Gumbel, 1958; Elías, 1963), utilizado con frecuencia por el Instituto Nacional de Meteorología, al menos hasta hace poco años (Blanco, 1983; I.N.M., 1988), está relacionado, por una parte, con el método de los momentos (al que reproduce cuando el tamaño de la muestra se hace muy grande) y, por otra, no es más que un procedimiento de mínimos cuadrados, algo más sofisticado que el descrito con anterioridad.

En el método anterior, en el cual se planteaba la regresión de  $y$  sobre  $x$ , se partía de la minimización de la suma de los cuadrados de las distancias horizontales  $\sum_{m=1}^n (y_m - ax_m^o + au)^2$ , pero también podríamos habernos planteado la regresión de  $x$  sobre  $y$ , minimizando la suma de

los cuadrados de las distancias verticales  $\sum_{m=1}^n (x_m^o - \frac{y_m}{a} - u)^2$ . Ambas condiciones no son equivalentes, pues la relación entre  $x$  e  $y$  no es funcional, sino estadística.

El primer procedimiento conducía a las estimaciones vistas en las ec. (29) y (30), a las que denominaremos  $a_1$  y  $u_1$ , mientras que el segundo nos llevaría a otras, a las que llamaremos  $a_2$  y  $u_2$ . La solución de los dos sistemas de ecuaciones involucrados nos conduce a las expresiones (Gumbel, 1958):

$$a_1 = \frac{\text{Cov}(x^o, y)}{\sigma_x^2} ; \quad a_2 = \frac{\sigma_n^2}{\text{Cov}(x^o, y)} \quad (31)$$

$$\bar{x} - u_2 = \frac{\gamma_n}{a_2} ; \quad \bar{x} - u_1 = \frac{\gamma_n}{a_1} \quad (32)$$

donde, como antes  $\gamma_n$  es el valor medio de  $y$ , siendo  $\sigma_n^2$  su varianza (no corregida). Como estimación final se toma  $a = \sqrt{a_1 a_2}$  y  $\bar{x} - u = \sqrt{(\bar{x} - u_1)(\bar{x} - u_2)}$ , es decir, las medias geométricas de las estimaciones, las cuales coinciden aproximadamente con las aritméticas, resultando:

$$a = \frac{\sigma_n}{\sigma_x} \quad (33)$$

$$u = \bar{x} - \frac{\gamma_n}{a} \quad (34)$$

Cuando  $n$  es muy grande  $\gamma_n$  tiende al valor de la constante de Euler  $\gamma$ , mientras que  $\sigma_n$  tiende a  $\pi\sqrt{6}$ , por lo que comparando las expresiones (25) y (26), obtenidas para el método de los momentos, con las (33) y (34) se comprueba que, efectivamente, aquellas eran un caso particular de éstas, en el límite de grandes valores de  $n$ .

2.2.2.d. Método de máxima verosimilitud.

El método de estimación de los parámetros  $a$  y  $u$  más eficaz estadísticamente, es decir, aquel que extrae la mayor cantidad de información de la serie de observaciones, de forma que la varianza de la distribución muestral de los parámetros sea mínima es el que se basa en hacer máxima la función de verosimilitud (*likelihood*) de la distribución, definida (Kotz *et al.*, 1983; Essenwanger, 1986; W.M.O., 1983a) como el producto de las  $n$  funciones de densidad de probabilidad individuales correspondientes a las  $n$  observaciones. Por comodidad, lo que se maximiza es el logaritmo de la función de verosimilitud, es decir, el valor de

$$\mathcal{L}(a,u) = \sum_{i=1}^n \log f(a, u; x_i) = n \log a - \sum_{i=1}^n e^{-a(x_i-u)} - \sum_{i=1}^n a(x_i-u) \quad (35)$$

Las condiciones de máximo serán:

$$\frac{\partial \mathcal{L}}{\partial u} = 0 \quad ; \quad \frac{\partial \mathcal{L}}{\partial a} = 0 \quad (36)$$

La primera de ellas conduce de manera inmediata a que:

$$\hat{u} = - \frac{1}{\hat{a}} \log \frac{\sum_{i=1}^n x_i e^{-\hat{a}x_i}}{n} \quad (37)$$

mientras que la segunda, con la ayuda de esta última expresión, nos proporciona que:

$$\hat{a}^{-1} = \bar{x} - \frac{\sum_{i=1}^n x_i e^{-\hat{a}x_i}}{\sum_{i=1}^n e^{-\hat{a}x_i}} \quad (38)$$

donde, como de costumbre, se ha utilizado el símbolo  $\hat{\alpha}$  para representar las estimaciones de máxima verosimilitud. La resolución de dichas ecuaciones, que se encuentran acopladas, requiere el empleo de algún método numérico, como pudiera ser el siguiente:

- 1º) Se obtiene una primera aproximación del parámetro  $\alpha$ , por el método de los momentos, por ejemplo.
- 2º) Se sustituye el valor obtenido en el segundo miembro de la ec. (38), obteniéndose una nueva estimación.
- 3º) Se reitera el paso anterior cuantas veces deseemos, hasta conseguir la exactitud deseada. En nuestro caso se ha considerado suficiente obtener un error relativo inferior a  $10^{-4}$ . En este punto se recomienda, para facilitar la convergencia, que si  $a^{(1)}$  y  $a^{(2)}$  son las dos primeras estimaciones, la tercera sea  $a^{(3)} = a^{(2)} + (a^{(1)} - a^{(2)}) / 3$ .
- 4º) Una vez conocido  $\hat{\alpha}$ , se obtiene  $\hat{u}$  por la ec. (37).

#### 2.2.2.e. Método de Kimball

Debido a que las ec. (37) y (38) son algo complicadas de resolver y además conducen a estimaciones que sólo son asintóticamente insesgadas, en ocasiones se utiliza otras estimaciones algo menos eficaces pero insesgadas y con un cálculo más cómodo. Entre ellas figura el método propuesto en 1954 por Lieblein (citado por Gumbel, 1958) y la solución de Kimball (1958). Según Sneyers (1975) la segunda de ellas es más eficaz que la primera y presenta además la ventaja de que su eficacia es regularmente creciente con  $n$ . Por otro lado, el método de Lieblein requiere de un proceso de aleatorización de los datos que puede llevar a resultados diferentes, según la ocasión, por lo que nos limitaremos únicamente al método propuesto por Kimball (1958).

---

En la solución de Kimball (1958) se vuelve a considerar los datos  $x_i$  ordenados por orden creciente. De la ec. (37) se tiene que

$$\sum_{m=1}^n e^{-ax_m^o} = n e^{-au} \quad (39)$$

Sustituyendo esta expresión en la ec. (38) y con la ayuda de la expresión (18), tendríamos:

$$n a^{-1} = \sum_{m=1}^m x_m^o [1 + \log F(x_m^o)] \quad (40)$$

Sustituyendo  $\log F(x_m^o)$  por su valor medio se tiene:

$$n a^{-1} = \sum_{m=1}^m x_m^o [1 - (\frac{1}{m} + \frac{1}{m+1} + \dots + \frac{1}{n})] \quad (41)$$

de donde se puede obtener  $a^{-1}$ . Sin embargo, la estimación obtenida no es insesgada y, para evitarlo, se multiplica por un coeficiente corrector  $b_n$ , de manera que:

$$a^{-1} = \frac{b_n}{n} \sum_{m=1}^m x_m^o [1 - (\frac{1}{m} + \frac{1}{m+1} + \dots + \frac{1}{n})] \quad (42)$$

pudiéndose obtener, si  $n > 11$ , el coeficiente  $b_n$  de la expresión:

$$\log (b_n - 1) = - 0,975652 \log (n - 1) + 1,043532 - \frac{1,950309}{\log (n - 1)} + \frac{3,574231}{\log^2 (n - 1)} \quad (43)$$

la cual, para muestras del orden de  $n \approx 50$ , conduce a un valor de  $b_n \approx 1,05$ .

La estimación que se considera para el parámetro  $u$  es, formalmente, la misma que en el método de los momentos, es decir:

$$u = \bar{x} - \frac{\gamma}{\alpha} \quad (44)$$

## 2.3. La distribución SQRT-ET o Raíz.

### 2.3.1. Origen de la distribución Raíz.

Etoh *et al.* (1987) derivaron una nueva distribución biparamétrica de valores extremos para estudiar las lluvias fuertes en Japón, a la que denominaron distribución SQRT-ET-max (*Square-root Exponential Type distribution of the maximum*) y a la que, por abreviar, nos referiremos como la distribución Raíz.

Para su deducción, se supuso que la distribución inicial de las precipitaciones, de la cual se extraen los máximos, era tal que:

- i) la intensidad y la duración de un aguacero son independientes.
- ii) una de ellas se distribuye exponencialmente mientras que la otra obedece a una distribución *Gamma*.
- iii) la cantidad total de lluvia es proporcional al producto de ambas.

Con estas hipótesis, que tienen un adecuado soporte experimental (Etoh y Murota, 1986), es posible obtener que la cantidad de lluvia recogida en un suceso de precipitación debe responder a una f. d. p. dada por la expresión:

$$h(x) = c (\beta x)^{-\nu/2} K_{\nu}(\sqrt{\beta x}) \quad (45)$$

con  $c = 2^{v-1} \beta / \Gamma(1-v)$ , siendo  $v$  y  $\beta$  los parámetros de la distribución y  $K_v$  la función de Bessel modificada de segunda clase. Esta distribución, a la que se le denominó *SQRT-K*, no es más que una generalización de la distribución de Eagleson (en la cual  $v = 0$ ), que, a su vez, es capaz de ajustarse mucho mejor a los datos observados de lluvia que una distribución exponencial uniparamétrica, proporcionando un ajuste tan bueno como una distribución *Gamma*, de dos parámetros. Por todo ello es de esperar que también la distribución *SQRT-K* proporcione una adecuada representación de los datos, como se comprueba con los registros experimentales (Etoh y Murota, 1986).

Tomando para la expresión de la función de Bessel el primer término de su desarrollo asintótico, válido para grandes valores de la variable, se tiene:

$$K_v(\sqrt{\beta x}) \sim \sqrt{\frac{\pi}{2\sqrt{\beta x}}} e^{-\sqrt{\beta x}} \quad (46)$$

Despreciando la variación de la potencia del argumento  $\sqrt{\beta x}$  frente a la de la función exponencial y normalizando a la unidad, nos queda que:

$$h(x) = \frac{\beta}{2} e^{-\sqrt{\beta x}} \quad (47)$$

siendo la f. d. a. correspondiente:

$$H(x) = 1 - (1 + \sqrt{\beta x}) e^{-\sqrt{\beta x}} \quad (48)$$

Admitiendo ahora, como en la deducción formal de la distribución de Gumbel del apartado 2.2.1. que la ocurrencia de las mayores lluvias responde a un proceso de Poisson, hipótesis tradicionalmente aceptada para fenómenos poco frecuentes y sustituyendo la expresión de  $H(x)$  en la ec. (9) obtendríamos una f.d.a. para los máximos en la forma:

$$F(x) = \begin{cases} 0 & x < 0 \\ \exp[-\lambda (1 + \sqrt{\beta x}) \exp(-\sqrt{\beta x})] & x \geq 0 \end{cases} \quad (49)$$

siendo  $\lambda$  un parámetro de forma y  $\beta$  un parámetro de escala. Se trata pues de una distribución biparamétrica, al igual que la de Gumbel. Sin embargo, su cola es mucho mayor, dado que en el argumento negativo que aparece en la segunda exponencial figura la raíz cuadrada de la variable aleatoria, mientras que en la de Gumbel la relación era lineal. En consecuencia, la distribución Raíz asignará una mayor probabilidad de ocurrencia a las lluvias más fuertes. La presencia de la raíz cuadrada obliga a condicionar la probabilidad para  $x \geq 0$ , de forma que la correspondiente f. d. p. será:

$$f(x) = \frac{\lambda \beta e^{-\sqrt{\beta x}} F(x)}{2 (1 - e^{-\lambda})} \quad (50)$$

La presencia del término entre paréntesis en el denominador es debida al proceso de condicionamiento, siendo su efecto despreciable si  $\lambda$  es suficientemente grande, de forma que  $e^{-\lambda} \ll 1$ , en cuyo caso la ec. (49) no necesitaría ser modificada para obtener la f. d. a. .

### ***2.3.2. Estimación de parámetros de la distribución Raíz. Método de máxima verosimilitud.***

Aunque la estimación de los parámetros  $\lambda$  y  $\beta$  podría efectuarse por diversos métodos, como en el caso de la distribución de Gumbel, el único propuesto por los creadores de la distribución Raíz es el basado en las condiciones de máxima verosimilitud. Es de suponer que esto es debido, en principio, a las características de la función de distribución, que harían que los cálculos involucrados en los métodos de los momentos o de los mínimos cuadrados, por ejemplo, no fuesen inmediatos y requiriesen procesos numéricos de cálculo. Por otra parte, el método de máxima verosimilitud demostró su bondad cuando lo aplicaron al estudio de lluvias extremas en Japón, obteniendo una gran estabilidad en la estimación de los parámetros (consecuencia lógica del argumento negativo en forma de raíz cuadrada que aparece en la distribución), una mayor verosimilitud en comparación con la distribución de Gumbel (en una proporción aproximada de cinco casos frente a uno) y otras con un mayor número de parámetros (como la distribución *Log-normal* de tres parámetros) así como una fiabilidad aceptable para los

máximos históricos. Dado que, además, está reconocido como el método de estimación más eficiente, nos limitaremos a él.

Siguiendo un procedimiento similar al expuesto en el apartado correspondiente al caso de la distribución de Gumbel, los parámetros de la distribución Raíz pueden estimarse como aquellos que maximizan el valor de la función de verosimilitud logarítmica  $\mathcal{L}$ :

$$\begin{aligned} \mathcal{L}(\lambda, \beta) &= \sum_{i=1}^n \log f(\lambda, \beta; x_i) = \\ &= \sum_{i=1}^n \log \left\{ \frac{\lambda \beta e^{-\sqrt{\beta x_i}} \exp[-\lambda(1 + \sqrt{\beta x_i}) \exp(-\sqrt{\beta x_i})]}{2(1 - e^{-\lambda})} \right\} \end{aligned} \quad (51)$$

Para imponer la condición de máximo se deriva parcialmente respecto del parámetro  $\beta$  y se iguala a cero, obteniéndose que la estimación máximo-verosímil para el parámetro  $\lambda$  viene ligada a la estimación del parámetro  $\beta$  mediante la relación:

$$\hat{\lambda} = \frac{\sum_{i=1}^n \sqrt{\hat{\beta} x_i} - 2n}{\sum_{i=1}^n \hat{\beta} x_i e^{-\sqrt{\hat{\beta} x_i}}} \quad (52)$$

Sustituyendo esta expresión para  $\lambda$  en la ecuación de  $\mathcal{L}$  dada por la (51) habremos reducido el número de parámetros a uno sólo, es decir  $\beta$ , siendo fácil determinar el valor de dicho parámetro que maximiza el valor de  $\mathcal{L}$ . En estas condiciones, reducida la participación a sólo un parámetro, la programación del cálculo automatizado mediante un proceso de inspección de precisión creciente es cómoda, con una rapidez de ejecución aceptable. En nuestro caso, es suficiente llevar a cabo la estimación de  $\beta$  con un error absoluto que sea inferior a  $10^{-4}$ . Una vez conocido el valor de  $\hat{\beta}$  es inmediato obtener el de  $\hat{\lambda}$  mediante la ec. (52). De esta forma, el hecho de haber recurrido a procesos numéricos de cálculo sí que estaría justificado por la máxima eficacia estadística que se consigue mediante el método de máxima verosimilitud.

## 2.4. Conclusiones parciales.

Resumiendo, hemos visto cómo la distribución de Gumbel podía derivarse de múltiples maneras. Se trata de la distribución asintótica de extremos a la que, bajo ciertas condiciones, tienden las distribuciones iniciales de tipo exponencial, tan frecuentes en climatología, por lo que necesariamente ha de desempeñar un papel clave. Ha sido ampliamente utilizada, habiéndose desarrollado numerosos métodos de estimación de sus parámetros, los cuales hemos ido recopilando, entre los que cabría establecer dos grupos. El primero de ellos, que denominaremos Grupo 1, estaría compuesto por aquellos métodos que suponen una solución de mínimos cuadrados: el de mínimos cuadrados propiamente dicho, el de Chow y el de los momentos (en la medida en que es el límite al que tiende el de Chow cuando el tamaño de la muestra es muy grande). El Grupo 2 lo integrarían aquellos otros que recurren en mayor o menor medida a las condiciones de máxima verosimilitud, como la solución de Kimball y el propio de máxima verosimilitud.

Sin embargo, cabe la posibilidad de que las hipótesis sobre las que se apoya esa distribución, aunque inicialmente plausibles, no lleguen a verificarse, en ocasiones, de manera suficientemente aproximada en un medio tan complejo como la atmósfera. La solución, en esas ocasiones, tal vez pase por esquivar los procesos deductivos en los que se homogeneizan las características de la distribución inicial y, en cambio, se mantenga alguna peculiaridad esencial de la misma, como es el caso de la existencia de un argumento no lineal en una distribución como la Raíz, deducida específicamente para el tratamiento de las lluvias, aun al precio, razonablemente admisible, de una mayor complejidad algorítmica.

## CAPÍTULO 3.

### RESULTADOS Y DISCUSIÓN.

---

La aplicación de los métodos de estimación de los parámetros de las distribuciones estadísticas vistas en el Capítulo 2 requiere la satisfacción previa, por parte de las series de datos a manejar, de determinados requisitos, tales como la fiabilidad y la continuidad de las observaciones, así como una longitud suficiente de los períodos de observación, con los cuales se pretende garantizar la calidad y la representatividad de los datos. También es preciso asegurar el carácter aleatorio de las series, pues éste constituye una hipótesis fundamental en el análisis estadístico de las mismas y, por tanto, una condición necesaria antes de proceder a cualquier otro análisis complementario.

En general, la forma real de la ley de distribución de probabilidades de las observaciones no es conocida, de manera que es preciso elegir aquella que mejor concuerde con las series de observaciones, tras lo cual adquirirá mayor sentido la estimación paramétrica. Para proceder a esa elección habrá que realizar una prueba de adherencia, como la basada en el estadístico  $\chi^2$ . A continuación será conveniente asegurarse de la buena adecuación de la ley de distribución someténdola a una nueva crítica estadística, con algún complemento de información, como pueda ser la capacidad de explicar de manera razonable la ocurrencia de los valores máximos históricos. En el hipotético caso de que más de una ley de distribución de probabilidades resultase satisfactoria, convendría elegir en todo caso aquella que resultase más simple o que estuviese dotada de una mayor estabilidad frente a las muestras de datos.

---

### **3.1. Descripción de los datos utilizados.**

Los datos a los que se ha recurrido han sido extraídos de las series de precipitaciones máximas anuales en 24 horas correspondientes a la red principal de observatorios que el Instituto Nacional de Meteorología, a través de su Centro Meteorológico Territorial de Murcia, mantiene en la provincia de Murcia. En dichas series, cada año se encuentra representado por el valor máximo de precipitación registrada en el observatorio durante un día, iniciando éste a las 0700 TMG (día pluviométrico).

La utilización de series de máximos anuales diarios se fundamenta, como adelantamos al establecer al marco del trabajo, en la propia disponibilidad de los datos y en la existencia de buenas bases teóricas para la extrapolación más allá del intervalo de la observación, aun cuando hicimos constar la limitación que constituye el no tener en cuenta la posibilidad de que el segundo máximo anual pudiera superar a máximos de otros años. El hipotético uso de series parciales obligaría a introducir un nuevo parámetro para dar cuenta del umbral seleccionado y, además, complicaría la teoría estadística a utilizar, debido a la dependencia de los sucesos, que podrían incluso sucederse diariamente uno a otro.

### **3.2. Control de calidad de los datos.**

Para establecer un primer control de calidad de las series nos hemos limitado a la red principal de observatorios, atendida exclusivamente por personal profesional, lo que supone una garantía en la fiabilidad de los datos, reduciendo al mínimo los errores de tipo instrumental, los derivados de la transmisión de datos y los debidos a fallos humanos.

---

La longitud de la serie va a constituir el siguiente criterio para la selección de los datos. De acuerdo con las recomendaciones comúnmente establecidas para el caso de datos de precipitación, el tamaño de la muestra necesario para obtener distribuciones de frecuencia estables viene a situarse en torno a los 40-50 años. Si imponemos la ausencia de cambios de emplazamiento en los últimos decenios, así como un tamaño mínimo de 40 datos, nos vemos obligados a restringirnos a las series correspondientes a los aeródromos de Alcantarilla, ubicado en el interior de la provincia, con 53 datos (en el período 1941-1994), y de San Javier, a orillas del Mar Menor, con 48 datos (en el período 1945-1994), en los que sólo se registran lagunas de 1 y 2 años, respectivamente.

### **3.3. Aleatoriedad de las series.**

Antes de proceder a cualquier otro análisis estadístico, es preciso garantizar el carácter aleatorio de las series. Para el caso de observaciones meteorológicas, y siguiendo a Sneyers (1975, p. 5) “... se puede considerar el carácter aleatorio simple de una serie como suficientemente bien establecido si la aplicación de un test de correlación serial y la de un test de tendencia conducen ... a la aceptación de la hipótesis nula [serie aleatoria]”. Aun cuando es lógico que determinadas series de tiempo de carácter meteorológico presenten persistencia (Essenwanger, 1976), en principio, no hay razones para que ésta deba admitirse en el caso de las lluvias extremas anuales.

El test de correlación serial más apropiado y general es el de Wald-Woldfowitz (cuya simplificación conduciría al test de las secuencias), que se aplica en su forma unilateral (sólo se produce rechazo en la eventualidad de correlación serial positiva, es decir, persistencia). Como test de tendencia podríamos utilizar el que está basado en el cálculo del coeficiente  $r_s$  de Spearman o bien en el del coeficiente  $t$  de Kendall (test de Mann), ambos en su forma bilateral.

---

La combinación de los resultados de las dos pruebas (correlación serial y tendencia) y de los correspondientes niveles críticos de significación puede hacerse utilizando el test de significación múltiple de Fisher.

Siguiendo los procedimientos descritos por Sneyers (1975) para la aplicación de estas pruebas se han obtenido los distintos niveles de significación que figuran en la tabla 1. De ella se deduce, fijando el nivel de significación en el valor  $\alpha_0 = 5 \%$ , que no es rechazable la hipótesis del carácter aleatorio de la serie correspondiente a San Javier, que proporciona valores del nivel de significación múltiple de Fisher del orden del 60 %, independientemente de que combinemos el test de correlación serial del Wald-Woldfowitz con los de tendencia de Spearman o de Mann. Por el contrario, los valores correspondientes a Alcantarilla se reducen hasta valores cercanos al 2 %, por lo que según el criterio indicado, hemos de rechazar la hipótesis de aleatoriedad y esto es debido, fundamentalmente, al bajísimo nivel crítico de significación obtenido en el test de correlación serial, que se sitúa en apenas el 1 %, aproximadamente, y que parece indicar la existencia de persistencia en la serie. En consecuencia, por la exigencia previa de aleatoriedad, la única serie que puede ser seleccionada para el estudio de la aplicabilidad de las distribuciones estadísticas es la correspondiente al aeródromo de San Javier.

TEST DE ALEATORIEDAD		OBSERVATORIO	
		Alcantarilla	San Javier
Correlación serial	(Wald-Woldfowitz)	1,1	47,3
Tendencia	Spearman	29,7	52,4
	Mann	20,3	53,4
Test de significación múltiple (Fisher)	Wald-W. $\oplus$ Spearman	2,1	59,4
	Wald-W. $\oplus$ Mann	1,6	60,1

**Tabla 1. Niveles críticos de significación (%) obtenidos en las pruebas aplicadas a las series de datos de los observatorios.**

### 3.4. Bondad del ajuste general.

Los valores estimados para los parámetros de las distribuciones de Gumbel y Raíz, según los distintos métodos expuestos en el capítulo 2 son los que figuran en la tabla 2. En el caso de la distribución Raíz, y de un modo directo, el valor del parámetro  $\lambda = 13,3$  puede interpretarse como la frecuencia media anual de aguaceros independientes. Por otro lado, el valor del parámetro  $\beta$  nos puede dar información acerca de la distribución inicial de tales aguaceros, a través de la ec. (47). En el caso de la distribución de Gumbel, admitiendo que la distribución inicial viniese dada por la ec. (15), lo cual sólo era estrictamente necesario en la deducción basada en el proceso de Poisson, el número medio  $\lambda$  de sucesos anuales puede obtenerse, para cada uno de los métodos de estimación, mediante la relación  $u = \frac{\log \lambda}{a}$ . Esta expresión conduce a valores para  $\lambda$  comprendidos entre 2,2 (para el método de mínimos cuadrados) y 5,5 (en el caso del método de máxima verosimilitud).

<b>DISTRIBUCIÓN DE GUMBEL</b>		
<b>Método</b>	<b>a (10<sup>-2</sup> mm<sup>-1</sup>)</b>	<b>u (mm)</b>
Momentos	2,57	43,2
Chow	2,32	42,0
Mínimos cuadrados	2,06	39,0
Máxima verosimilitud	3,57	47,9
Kimball	3,33	48,3
<b>DISTRIBUCIÓN RAÍZ</b>		
<b>Método</b>	<b><math>\beta</math> (10<sup>-1</sup> mm<sup>-1</sup>)</b>	<b><math>\lambda</math></b>
Máxima verosimilitud	4,04	13,3

Tabla 2. Valores de las estimaciones de los parámetros  $a$  y  $u$  de la distribución de Gumbel y de los parámetros  $\beta$  y  $\lambda$  de la distribución Raíz.

Aunque estos últimos valores ya parecen demasiado bajos, para juzgar la bondad o idoneidad de las funciones de distribución que genera cada una de las estimaciones, hay que recurrir a pruebas de significación que nos informen acerca de la adecuación existente entre las distribuciones teóricas y las observaciones.

El primero de los índices numéricos a manejar puede ser el estadístico  $\chi^2$ , el cual nos proporciona una medida de la discrepancia existente entre las frecuencias teóricas y las empíricas, supuesto un tamaño  $n$  de la muestra suficientemente grande, a un nivel de significación dado, que situaremos en  $\alpha_0 = 5\%$ , en ensayo unilateral. Denominaremos  $\alpha$  al nivel crítico de significación asociado a cada valor de  $\chi^2$ , de forma que  $(1-\alpha)$  nos dará una medida del grado de incompatibilidad de la situación observada con la hipótesis de adherencia a cada una de las distribuciones ensayadas.

Idealmente, convendría distribuir los datos en un gran número de intervalos de clase. Sin embargo, puesto que el tamaño de la muestra está limitado a  $n = 48$ , en la prueba de  $\chi^2$  será preciso adoptar una solución de compromiso entre dicho número y la frecuencia absoluta esperada para cada uno de ellos. En general, no se recomienda que dicha frecuencia sea inferior a 10. La limitada cantidad de datos disponible no permitirá estudiar las distribuciones muy matizadamente, sino que únicamente podrá entrarse a analizar los aspectos más generales, de modo que las pruebas diseñadas tal vez no puedan ser muy potentes. En estas condiciones, para que la potencia del test a emplear sea la máxima posible, conviene que todos los intervalos de clase sean equiprobables (Houghton, 1985). Por todo ello, será razonable fijar el número de intervalos de clase en 4, es decir, con un valor para las frecuencias absolutas esperadas de 12, lo que cumplirá con los requerimientos mínimos para ambos condicionantes. De esta manera, los intervalos de clase estarán formados por cuartiles.

Por otra parte, al ser el número de intervalos tan reducido, surge otro problema, pues no debemos olvidar que estamos intentando manejar discretamente una distribución que, en

---

realidad, tiene naturaleza continua. En este caso, en el cual los grados de libertad del estadístico  $\chi^2$  están limitados a la unidad, se recomienda aplicar alguna corrección, pues el valor de  $\chi^2$  estará, en general, sobreestimado, siendo la corrección más habitual la de Yates, calculándose con ella el valor de  $\chi^2$  como:

$$\chi^2 = \sum_{c=1}^4 \frac{(|E_c - O_c| - 1/2)^2}{E_c} \quad (53)$$

siendo  $E_c$  y  $O_c$  las frecuencias esperadas y observadas, respectivamente, en el intervalo de clase  $c$ , obedeciendo la introducción del factor  $1/2$  a la corrección por continuidad. Aunque algunos autores recomiendan la aplicación de esta corrección de manera rutinaria, otros la limitan al caso de un grado de libertad. En nuestro caso, no habrá lugar para la polémica, pues habíamos reducido el número de grados de libertad a la unidad. A pesar de todo, a igual que con otras distribuciones de probabilidad, pero especialmente en las de decrecimiento exponencial como es nuestro caso, la exactitud de la prueba basada en el estadístico  $\chi^2$  podría estar limitada, permaneciendo un grado de incertidumbre sobre la bondad del ajuste (Houghton, 1985). La consideración de estas reservas podría hacernos rebajar el nivel de significación hasta el valor  $\alpha_0 = 1\%$  antes de rechazar la hipótesis de ajuste, aún a costa de que el test se hiciese más conservador.

En la tabla 3 se muestra, en las primeras columnas, los valores de los extremos inferiores de los intervalos de clase segundo, tercero (o mediana) y cuarto, obtenidos para cada método de estimación, necesarios para asignar las frecuencias empíricas a cada uno de los intervalos. Obsérvese que las estimaciones de máxima verosimilitud de las dos distribuciones conducen a valores similares para esos límites de los intervalos, lo cual también explica la semejanza de resultados en el cálculo del estadístico  $\chi^2$  y del nivel crítico de significación, reflejados en las últimas columnas. Sin embargo, las mayores diferencias relativas entre ambas distribuciones se producirán después del límite inferior del cuarto intervalo de clase, como se pondrá de manifiesto más adelante.

En las figuras 1 y 2 se muestra la diferencia entre esas frecuencias empíricas y la frecuencia teórica, en valores absolutos, de las distribuciones de Gumbel y Raíz asociadas a cada uno de los métodos de estimación. En la cola izquierda de las distribuciones de Gumbel correspondientes a los métodos de los momentos, de Chow y de mínimos cuadrados, que habíamos quedado en denominar como métodos del Grupo 1, se observa (fig. 1) una fuerte discrepancia. Se infravaloran las dos colas (cuartiles primero y cuarto), de manera más significativa la cola final. Lógicamente, se produce una sobrevaloración en los cuartiles centrales, donde las frecuencias observadas son muy superiores a las esperadas. Los valores del estadístico  $\chi^2$  obtenidos para estos métodos, una vez efectuada la corrección de Yates, son muy superiores a los críticos, tanto al nivel de significación del 5 % ( $\chi^2 = 3,84$ ) como en el del 1 % ( $\chi^2 = 6,63$ ), correspondiéndoles grados de incompatibilidad que oscilan entre el 99,41 % ( $\alpha = 0,59$  %) para el método de los momentos y el 99,99 % ( $\alpha = 0,01$ %) para el método de los mínimos cuadrados, aproximadamente. Dicho de otro modo, si los datos provinieran de una distribución de Gumbel, ajustada mediante el método de los mínimos cuadrados, por ejemplo, tan sólo en una de cada  $10^4$  pruebas que realizáramos encontraríamos un valor tan alto del estadístico  $\chi^2$ . Estos resultados obligan al rechazo de la hipótesis de adherencia de los datos a una distribución de Gumbel cuando se recurre a alguno de estos métodos de estimación.

MÉTODO	L. i. int. 2º	L. i. int. 3º	L. i. int. 4º	$\chi^2$	$\alpha$ ( $10^{-2}$ )
Momentos	30,5	57,4	91,7	7,58	0,590
Mínimos cuadrados	23,2	56,8	99,5	14,75	0,012
Chow	27,9	57,8	95,8	10,92	0,095
Máxima verosimilitud (G)	38,7	58,1	82,8	3,42	6,44
Kimball	38,5	59,3	85,7	4,92	2,66
Máxima verosimilitud (R)	36,4	54,4	81,5	3,75	5,28

**Tabla 3. Valores de los límites inferiores (mm) de los intervalos de clase 2º al 4º, del estadístico  $\chi^2$  y del nivel crítico de significación  $\alpha$ , para los distintos métodos de estimación de la distribución de Gumbel (G) y Raíz (R).**

El análisis de los resultados obtenidos mediante los métodos de máxima verosimilitud y de Kimball, que denominamos métodos del Grupo 2, o bien cuando recurrimos a la estimación máximo-verosímil de la distribución Raíz, ofrece un panorama radicalmente distinto, como se observa en la figura 2. En estos casos, las discrepancias son pequeñas, con valores mucho más bajos del estadístico  $\chi^2$ , evitando el rechazo de la hipótesis de adherencia al nivel de significación del 1 %, e incluso del 5 % en el caso de los dos métodos de máxima verosimilitud. Recuérdese que estos tres métodos de estimación tienen en común el hecho de emplear, al menos, alguna condición de máxima verosimilitud, como vimos en el capítulo 2, por lo que se

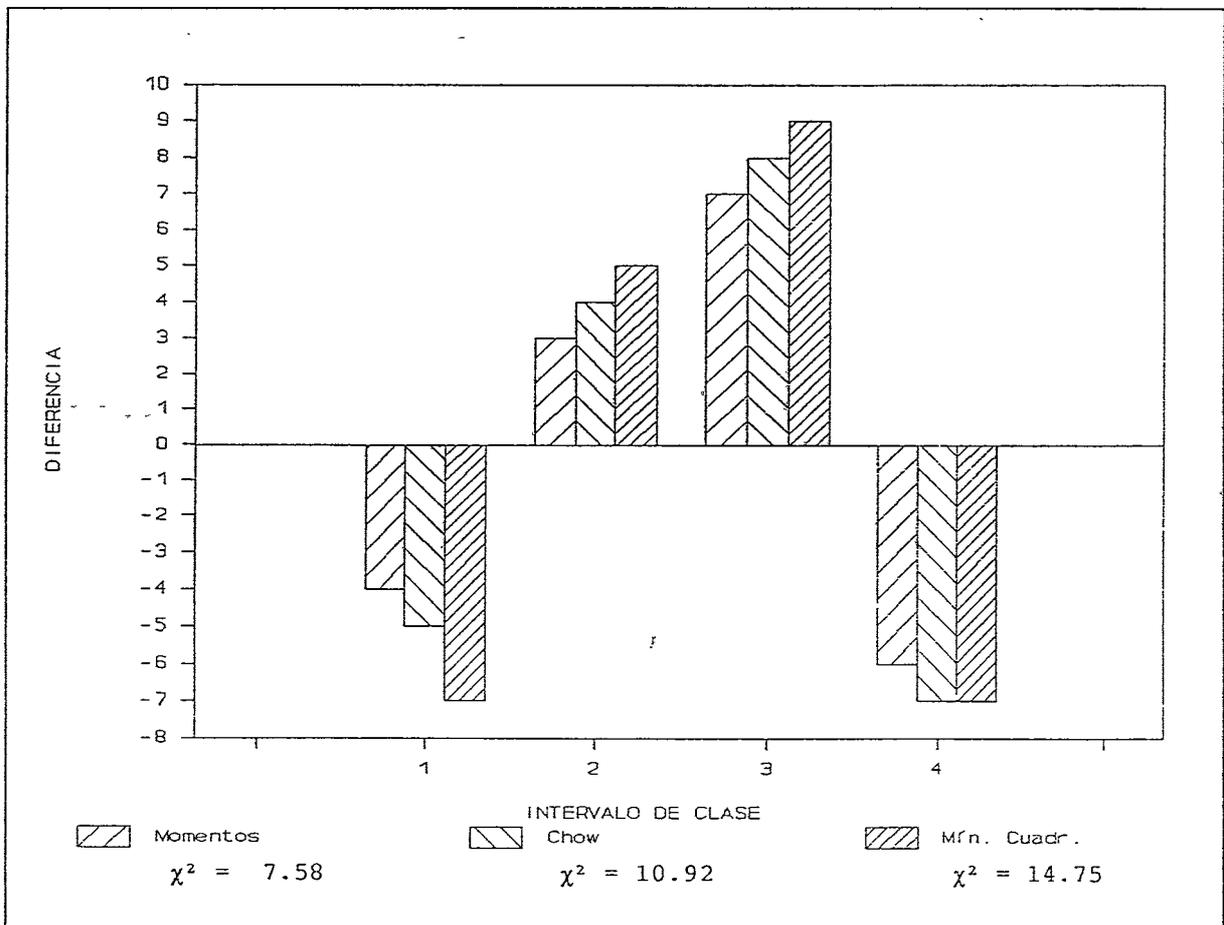
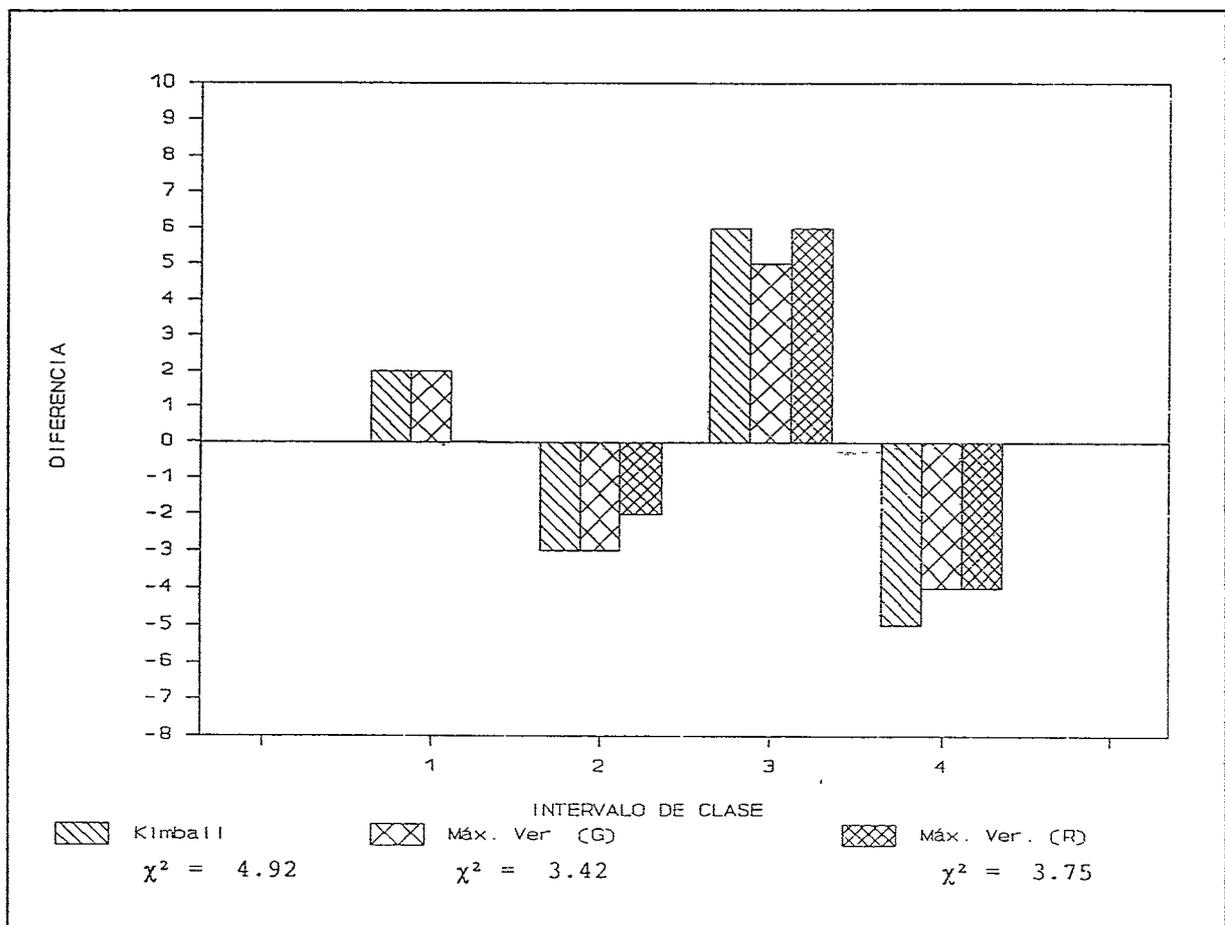


Figura 1. Diferencia entre las frecuencias absolutas empíricas y la teórica, según los intervalos de clase, correspondientes a las distribuciones de Gumbel asociadas a los métodos de estimación del Grupo 1.

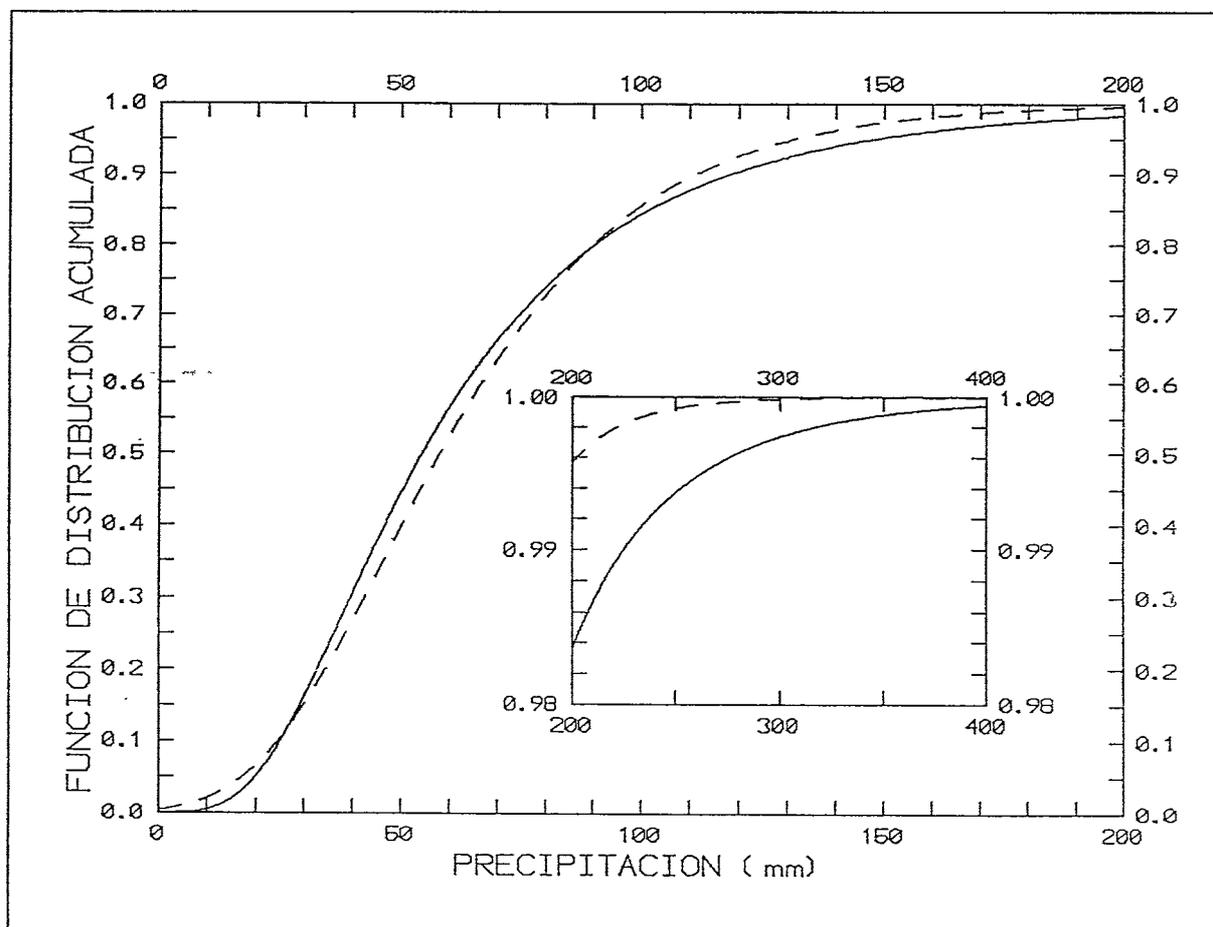
deduce que la imposición de dicha condición está estrechamente relacionada con la bondad del ajuste general.

Los máximos de las funciones de verosimilitud (logarítmicas) para las distribuciones de Gumbel y Raíz alcanzan los valores  $\mathcal{L}(\hat{a}, \hat{u}) = -238,4$  y  $\mathcal{L}(\hat{\beta}, \hat{\lambda}) = -236,0$ , respectivamente, por lo que en nuestro caso puede afirmarse que la distribución Raíz ha presentado una mayor verosimilitud que la distribución de Gumbel.



**Figura 2.** Diferencia entre las frecuencias absolutas empíricas y la teórica, según los intervalos de clase, correspondientes a las distribuciones de Gumbel (asociadas a los métodos de estimación del Grupo 2) y a la distribución Raíz estimada por el método de máxima verosimilitud.

La figura 3 nos ilustra gráficamente la mayor cola final que posee la distribución Raíz, o sea, los valores menores de la f. d. a., consecuencia del argumento en forma de raíz cuadrada que figura en su segunda exponencial, frente al argumento lineal de la Gumbel. En ella se ha representado ambas distribuciones, con estimación máximo-verosímil de los parámetros. Los puntos de corte están dados por los valores de la precipitación  $x = 26,6$  mm y  $x = 88,6$  mm. Por encima de este último valor la f. d. a. correspondiente a la distribución Raíz es menor o, lo que es lo mismo, otorga una mayor probabilidad de ocurrencia a los valores más altos de la precipitación, como vamos a ver, con mayor detalle, a continuación.



**Figura 3. Funciones de distribución acumulada de Gumbel (línea discontinua) y Raíz (línea continua), estimadas por el método de máxima verosimilitud.**

### 3.5. Fiabilidad de los máximos históricos.

El intervalo final de las distribuciones de máximos posee un interés preferente, al afectar a la predicción de las lluvias más fuertes, permitiendo la extrapolación temporal más allá de la longitud de los registros históricos. La capacidad de explicar los máximos históricos puede ser evaluada en función de la probabilidad teórica que proporciona la distribución, recurriendo al período medio de retorno  $T(x)$  para un valor  $x$  de la precipitación, que se define en términos de la f.d.a. (Sneyers, 1975; W.M.O., 1983b) como:

$$T(x) = \frac{1}{1 - F(x)} \quad (54)$$

Dado un valor de  $x$ , la probabilidad  $Q(x)$  de que durante  $n$  años no se produzca un valor que lo supere vendrá dada, en términos de su período de retorno  $T(x)$ , por:

$$Q(x) = [F(x)]^n = \left[ 1 - \frac{1}{T(x)} \right]^n \quad (55)$$

y, por tanto, la probabilidad  $P(x)$  de que se alcance o supere dicho valor al menos una vez en  $n$  años será (W.M.O., 1983a):

$$P(x) = 1 - Q(x) = 1 - \left[ 1 - \frac{1}{T(x)} \right]^n \quad (56)$$

La probabilidad de ocurrencia  $P$  del máximo histórico de precipitación nos da una idea de la magnitud de la extrapolación efectuada cuando al valor de dicho máximo le hacemos corresponder su período de retorno  $T$ . Si  $T \gg n$ , lo cual se cumple en nuestro rango de valores, se tiene:

$$\left[ 1 - \frac{1}{T} \right]^n \approx 1 - \frac{n}{T} \quad (57)$$

con lo cual la probabilidad  $P$  puede escribirse también como:

$$P \approx \frac{n}{T} \quad (58)$$

que no es más que el cociente entre el tamaño de la muestra  $n$  y el período de retorno  $T$ . La magnitud de la extrapolación  $E$  puede definirse entonces (Garrido y Caselles, 1995) como:

$$E = \frac{1}{P} \approx \frac{T}{n} \quad (59)$$

La tabla 4 ofrece los valores obtenidos por los diversos métodos de estimación para el período de retorno  $T$  del máximo histórico, registrado en el año 1987 con un valor de 330 mm, y la extrapolación  $E$  que se efectúa.

Limitándonos inicialmente a la distribución de Gumbel, se observa que, en general, los períodos de retorno teóricos son muy grandes, llegando a ser del orden de  $10^4$  años en los métodos de estimación de la distribución de Gumbel correspondientes al Grupo 2 (máxima verosimilitud y Kimball), lo cual es irreal si se tiene en cuenta que el tamaño de la serie original

MÉTODO	$T$ (años)	$P$ (%)	$E$	$S$ (%)
Momentos	1590	3,0	34	1,5
Mínimos cuadrados	400	5,9	9	2,3
Chow	790	11,3	17	1,3
Máxima Verosimilitud (G)	23.800	0,4	496	12,4
Kimball	11.800	0,2	246	21,1
Máxima Verosimilitud (R)	620	7,4	13	60,4

**Tabla 4.** Valores del período de retorno  $T$  del máximo histórico, de la probabilidad de ocurrencia  $P$ , de la extrapolación  $E$  y de la estabilidad  $S$ , para los distintos métodos de estimación de la distribución de Gumbel (G) y Raíz (R).

(48 años) es varios cientos de veces inferior y, por tanto, las probabilidades  $P$  sólo serán algo superiores a  $10^{-3}$ . Si imponemos, como criterio objetivo, que la probabilidad  $P$  deba ser al menos del 5% y, por consiguiente, se exige que las extrapolaciones deban ser tales que el período de retorno estimado para el máximo no supere en más de 20 veces el tamaño de la muestra ( $E \leq 20$ ), tendremos que rechazar los métodos del Grupo 2 y, también, el método de los momentos, aunque este último, al menos, quedaría en la franja de probabilidad  $P$  comprendida entre el 1 y el 5 %.

En la figura 4 se muestra simultáneamente los valores de  $P$  y los niveles críticos de significación  $\alpha$  obtenidos del test  $\chi^2$ , con una recta de ajuste de interés meramente cualitativo. Para los métodos de estimación de la distribución de Gumbel se aprecia, en general, que cuando aumenta la probabilidad  $P$  de ocurrencia del máximo disminuye el valor de la significación del ajuste, quedando todos los métodos en la región de rechazo, ya sea por lo irreal del período de retorno obtenido para el máximo ( $P < 5\%$ ) o porque el ajuste general proporcionado es malo ( $\alpha < 5\%$ ). Es más, no hay ningún método de los ensayados que supere ambas pruebas, incluso rebajando las exigencias al 1%.

Para la distribución Raíz, por el contrario, se pone de manifiesto, a través de la tabla 3, que el nivel crítico de significación  $\alpha$  es del orden de los obtenidos mediante los métodos de estimación de la distribución de Gumbel del Grupo 2, como era deseable. Por otra parte, véase la tabla 4, la probabilidad de ocurrencia  $P$  es similar a la que se consigue con los métodos del Grupo 1, las cuales eran razonables, resultando así que el empleo de la distribución Raíz no es rechazable (figura 4). El período de retorno obtenido en este caso, 620 años, para el máximo histórico de 330 mm, aunque es alto, no es disparatado, pudiendo representar la excepcionalidad del suceso.

Invirtiendo los términos de la discusión, si, como es habitual en muchas aplicaciones, se diseñase una estructura capaz de soportar la precipitación diaria correspondiente a un período de retorno de 500 años, obtendríamos que el valor de diseño según la distribución Raíz sería de

---

317 mm, sólo inferior en un 4 % al máximo histórico. En el caso de la distribución de Gumbel, estimados sus parámetros también por el método de máxima verosimilitud, se hubiese obtenido un valor de 222 mm, que sólo representa las dos terceras partes de dicho máximo, lo que habría podido poner en grave peligro la resistencia de la estructura.

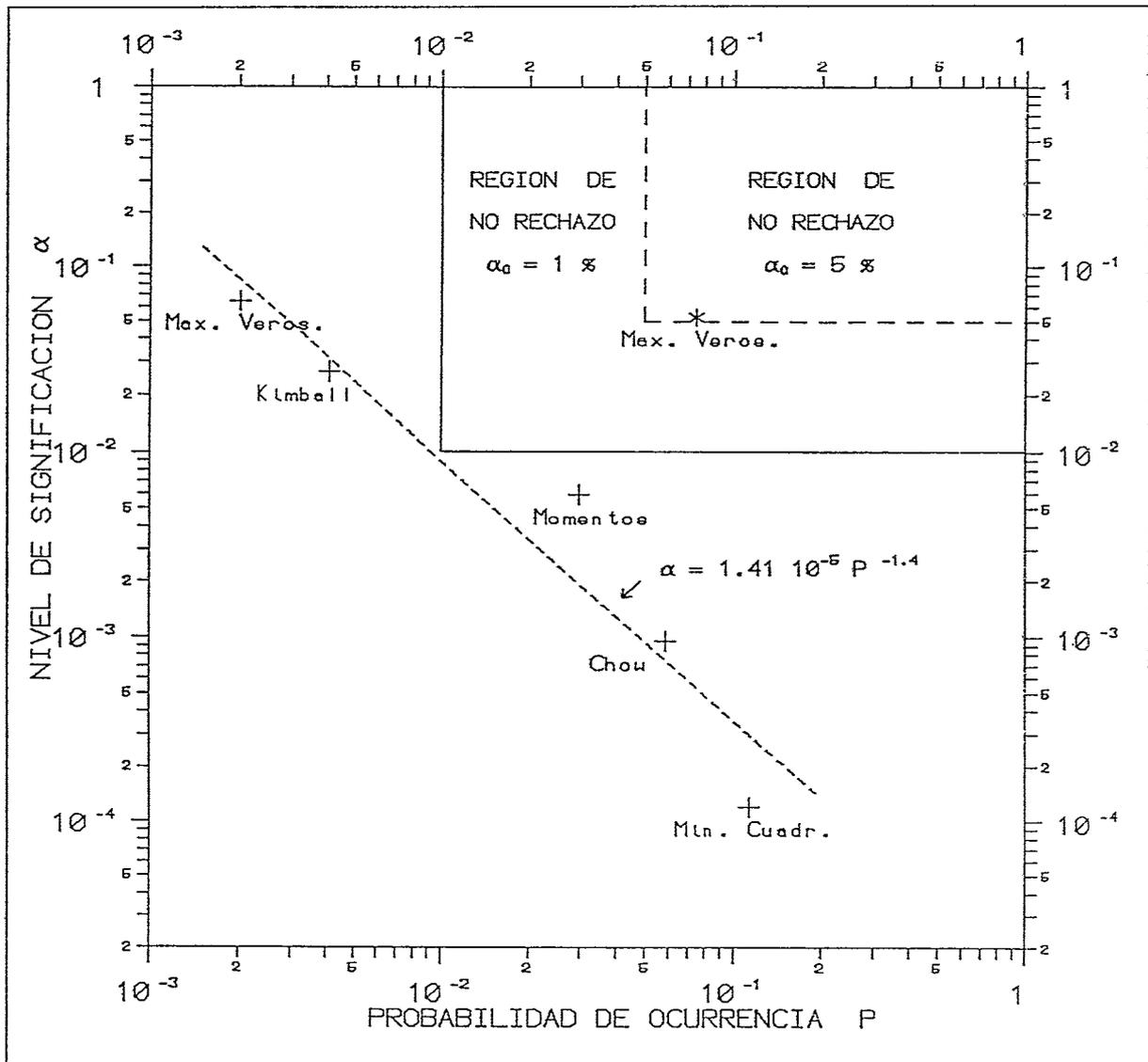


Figura 4. Nivel crítico de significación  $\alpha$  del test  $\chi^2$  frente a la probabilidad de ocurrencia  $P$  del máximo histórico, para los diversos métodos de estimación de la distribución de Gumbel (+) y de la distribución Raíz (\*).

Estos resultados son cualitativa, y casi cuantitativamente, idénticos a los obtenidos en otro trabajo previo, con una serie de datos menos extensa (Garrido y Caselles, 1995). Dado que el fracaso de la distribución de Gumbel no ha podido ser achacable al método de estimación empleado, nos encontramos con que es la propia aplicación de la distribución la que queda cuestionada y, con ella, la validez de las hipótesis de partida. El fallo no parece deberse a una casualidad y su explicación habría que buscarla en la posibilidad de que el tipo de distribución inicial no sea necesariamente exponencial, o bien que las observaciones diarias no sean independientes o que no se hayan realizado bajo idénticas condiciones, o bien que el tamaño de la muestra anual de la que se extrae cada máximo tal vez no sea suficientemente grande. Ese tamaño no puede ser ampliado por limitaciones naturales ( $N = 365$ , y no en todos los días del año se registran precipitaciones). Lo que sí podría ampliarse es la base empírica utilizada, extendiendo el presente estudio a otras series de observaciones más largas o bien aumentando la cobertura territorial, o la resolución espacial con otro tipo de estaciones meteorológicas, lo cual puede ser el objeto de un futuro trabajo.

La existencia de discrepancias entre la distribución de Gumbel y las observaciones disponibles, podría deberse a alguna o varias de las causas enumeradas anteriormente, pero, estrictamente, “no deben considerarse como un fallo de la teoría” de valores extremos, aplicable bajo determinadas circunstancias, de la que se deriva dicha distribución (Gumbel, 1958, p. 347). Del mismo modo que la distribución normal desempeña un papel estelar en la estadística de valores medios, es posible que la distribución de Gumbel lo juegue en la de valores extremos y que no sea capaz de dar buenos resultados en todo tipo de circunstancias. La distribución Raíz, la cual ha resultado mucho más satisfactoria en este caso, también es de tipo exponencial, luego debería conducir asintóticamente a la distribución de Gumbel. Esto podría suceder para grandes valores del parámetro  $\lambda$ , pero no para valores medianos, como ha sido nuestro caso, del mismo modo que la distribución *t de Student* conduciría a la normal para grandes valores del número de grados de libertad.

---

### 3.6. Estabilidad de la distribución.

Resulta interesante establecer una medida de la estabilidad de las distribuciones o, equivalentemente, de los parámetros que la definen, de manera que, cuanto mayor sea esa estabilidad, sea menor la dependencia de las estimaciones respecto de la muestra considerada. La estabilidad no es que sea recomendable por sí misma, sino que está supeditada a otros índices de bondad, siendo su importancia secundaria frente a la de otros índices de bondad. Tan sólo a igualdad de dichos índices sería preferible una distribución lo más estable posible. Téngase en cuenta, por ejemplo, que una distribución de probabilidad constante sería completamente estable, pero tendría un interés nulo.

Para cuantificar la dependencia con la muestra considerada vamos a estimar, en primer lugar, los parámetros a partir de la muestra completa y, en segundo lugar, prescindiendo de su máximo histórico, con lo que obtendremos dos evaluaciones distintas para la frecuencia de dicho máximo. En estas condiciones, la estabilidad  $S$  podrá establecerse mediante el cociente de los períodos de retorno asociados a dichas frecuencias (Etoh *et al.*, 1987), es decir:

$$S = \frac{T_n}{T_{n-1}} \quad (60)$$

siendo  $T_n$  y  $T_{n-1}$  los períodos medios de retorno del máximo histórico cuando se considera la serie completa de  $n$  datos o incompleta de  $n-1$  datos, respectivamente. En general  $T_n \leq T_{n-1}$  y el índice  $S$  así definido variará entre 0 y 1. En algún caso particular puede suceder que  $T_{n-1} \leq T_n$  pudiéndose redefinir, en esa ocasión,  $S$  como su inverso.

Según se deduce de la última columna de la tabla 4, en el caso de la distribución de Gumbel los valores obtenidos para la estabilidad de los distintos métodos de estimación oscilan entre el 1 % para el método de los mínimos cuadrados y el 21 % para el de máxima

verosimilitud. Por el contrario, para la distribución Raíz la estabilidad alcanza un valor del 60 %, mostrándose como la distribución más estable, con gran diferencia.

Los valores de la estabilidad  $S$  se han representado frente a los de las probabilidades  $P$  en la figura 5, observándose que, en general, al aumentar  $P$  disminuye  $S$  cuando se trata de la

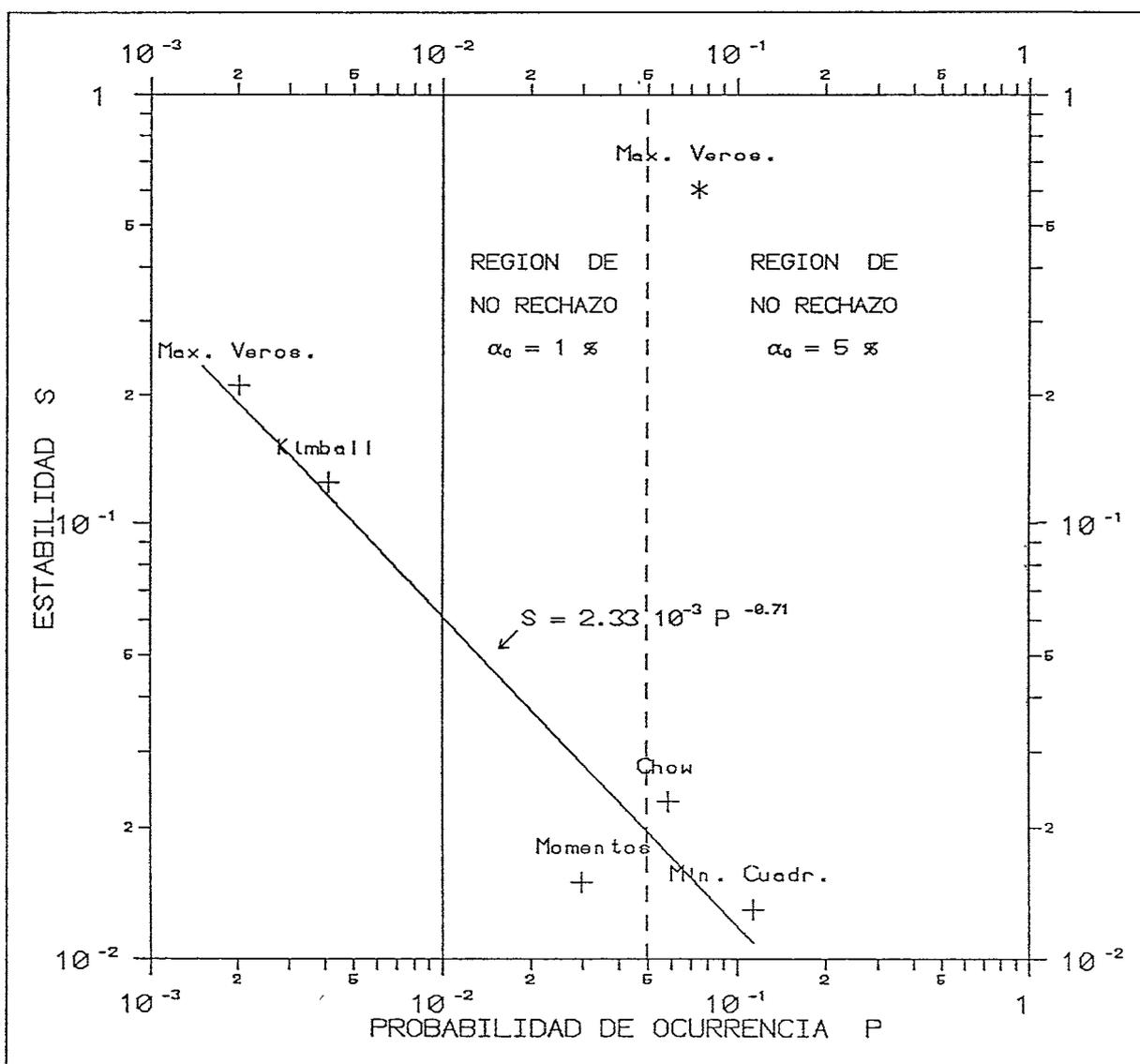


Figura 5. Estabilidad  $S$  frente a la probabilidad de ocurrencia  $P$  del máximo histórico, para los diversos métodos de estimación de la distribución de Gumbel (+) y de la distribución Raíz (\*).

distribución de Gumbel, de acuerdo con la recta cualitativa de ajuste. Los métodos del Grupo 1, que no realizan extrapolaciones muy grandes, son muy inestables, al reducirse  $S$  hasta valores del 2 %, es decir, el período de retorno del máximo histórico varía en un factor del orden de 50, según que se considere o no en la muestra al propio máximo. Por otro lado, los métodos del Grupo 2 resultan más estables que los anteriores, con una variación del período de retorno del máximo aproximadamente 10 veces inferior a la de los métodos del Grupo 1, aunque, como vimos, no son admisibles por la magnitud de los períodos de retorno que implicaban, es decir, de las extrapolaciones efectuadas. En el caso de la distribución Raíz, que presentaba un buen ajuste y una extrapolación razonable, nos encontramos además con una gran estabilidad, que alcanza el 60 %, de manera que la variación del período de retorno del máximo no llega a suponer una duplicación del mismo.

### **3.7. Conclusiones parciales.**

En resumen, la aplicación de la distribución de Gumbel a la serie de observaciones del aeródromo de San Javier y el análisis de los resultados obtenidos ha conducido a conclusiones insatisfactorias, independientemente de los métodos de estimación paramétrica. Mientras que para unos métodos el ajuste general proporcionado al conjunto de los datos es rechazable, para otros lo absurdo del período de retorno calculado para el máximo histórico deja sin sentido la aplicación de dicha distribución.

Sin embargo, el uso de la distribución Raíz, también biparamétrica, proporciona un ajuste general suficientemente adecuado, mientras que el período de retorno del máximo histórico cae dentro de los límites admisibles. Además, la estimación paramétrica máximo-verosímil para esta ley de distribución resulta muy estable.

### CONCLUSIONES.

---

Como resumen de este trabajo, se obtienen las siguientes conclusiones:

1. Las regiones expuestas al riesgo de precipitaciones intensas deben contar con evaluaciones realistas de vulnerabilidad. Aunque usualmente se viene utilizando la distribución estadística de Gumbel para valorar las probabilidades de ocurrencia, sus resultados no son suficientemente satisfactorios de manera general.
  2. Con el fin de determinar esas probabilidades, se han recopilado los cinco métodos de estimación de los parámetros de la distribución de Gumbel más universalmente aceptados: de los momentos, de mínimos cuadrados, de Chow, de máxima verosimilitud y de Kimball. Los tres primeros los agrupamos en lo que denominamos Grupo 1 y los dos restantes en el Grupo 2. También se recoge el método de estimación máximo-verosímil para una distribución alternativa, biparamétrica como la anterior, denominada SQRT-ET o Raíz, la cual fue específicamente derivada para el estudio del fenómeno de la lluvia y que posee, como peculiaridad matemática esencial frente a la anterior, la presencia de un argumento en forma de raíz cuadrada.
  3. Se estudia la aplicabilidad de estas distribuciones a las series anuales de precipitaciones máximas diarias de los observatorios meteorológicos de la provincia de Murcia. Imponiendo criterios de control de calidad de los datos (atención de personal profesional, longitud suficiente de las series y constancia de emplazamiento), así como la superación
-

de pruebas de aleatoriedad, los datos correspondientes al aeródromo de San Javier, ubicado junto al mar, han sido seleccionados para estudiar su distribución estadística.

4. Para la muestra seleccionada, y en el caso de la distribución de Gumbel, los tres métodos incluidos en el Grupo 1, deben ser rechazados dado el pésimo ajuste general que proporcionan, evaluado mediante el test  $\chi^2$ , que arroja niveles críticos de significación muy pobres, comprendidos entre el 0,01 y el 0,6 %, conduciendo a grados de incompatibilidad entre la distribución teórica y la empírica que llegan hasta el 99.99 % en el caso del método de los mínimos cuadrados. Los dos métodos restantes, pertenecientes al Grupo 2, proporcionan ajustes no rechazables al nivel de significación del 1 %, pero sólo podríamos mantener el de máxima verosimilitud al nivel de significación usual del 5 %. Estos dos últimos métodos generan períodos de retorno para el máximo histórico, con un valor de 330 mm, superiores a  $10^4$  años, los cuales son completamente irrealistas si los comparamos con la longitud de la serie (48 años), por lo que también resultan insatisfactorios.
5. Al no ser achacable el fracaso de la distribución de Gumbel al método de estimación paramétrica empleado, es la propia validez general de la distribución para valorar el riesgo de lluvias fuertes la que queda cuestionada. Adicionalmente, se ha encontrado que la estabilidad de la distribución es mucho mayor cuando se emplean los métodos del Grupo 2 que cuando se recurre a los del Grupo 1. En estos últimos, dependiendo de que se incluya o no en la muestra el propio máximo histórico, el período de retorno puede variar en un factor del orden de 50.
6. Cuando se ha empleado la distribución Raíz, el ajuste general que se obtiene es suficientemente bueno, con una verosimilitud mayor que en el caso de la distribución de Gumbel. El período de retorno proporcionado para el máximo histórico resulta admisible, de manera que el valor de diseño para un período de retorno de 500 años

coincide prácticamente con el de dicho máximo. Por otro lado, la estabilidad paramétrica de la distribución Raíz es muy superior a la de Gumbel, de modo que el factor de variación del período de retorno no llega a alcanzar el valor 2.

7. En consecuencia, en nuestro caso de estudio, la distribución Raíz se ha mostrado más apropiada que la tradicional de Gumbel para valorar el riesgo de lluvias fuertes. La ampliación de la base observacional, aumentando la resolución espacial o extendiendo el trabajo a un ámbito geográfico mayor, podrá ayudar a un mejor conocimiento de la vulnerabilidad ante los desastres naturales más preocupantes en la vertiente mediterránea española.

## BIBLIOGRAFÍA.

---

- 
- (1) Almarza, C. y López, J. A. 1995. Las temperaturas máximas absolutas en zonas costeras. El caso de La Coruña. *Situaciones de riesgo climático en España*, 153-161. Inst. Pirenaico de Ecología- C.S.I.C., Huesca.
- (2) Blanco, A. 1983. *Distribución de valores extremos*. Notas de Climatología, 2. Inst. Nac. de Meteor., Madrid, 9 pp.
- (3) Carrera, A. R. y Martínez, M., 1991.: Aproximaciones óptimas de las variables que describen sucesos extremos en meteorología. *Rev. de Geofísica*, **47**, 85-92.
- (4) Castillo, E. 1988. Estadística de Valores Extremos. Distribuciones Asintóticas. *Estad. Española*, **116**, 5-34.
- (5) Elías, F., 1963. *Precipitaciones máximas en España. Régimen de intensidades y frecuencias*. Ministerio de Agric., Serv. de Conservación de Suelos. Bol. Técn. nº 3, Madrid. 267 pp.
- (6) Etoh, T. y Murota, A. 1986. Probabilistic model of rainfall of a single storm. *Journal of Hydroscience and Hydraulic Engineering*, **4**, 65-67.
- (7) Etoh, T. *et al.* 1987. SQRT-Exponential Type distribution of maximum. *Hidrologic frequency modeling*, 253-264. Reidel Publ. Comp.
- (8) Essenwanger, O. M. 1976. Applied statistics in atmospheric science. Part A: Frequencies and curve fitting. *Developments in Atmospheric Science*, **4A**. Elsevier, Amsterdam. 412 pp.
- (9) Essenwanger, O. M. 1986. General Climatology, 1B. Elements of Statistical Analysis. *World Survey of Climatology*, **1B**. H. E. Landsberg (ed.). Elsevier Sc. Publ. B. V., Amsterdam, 424 pp.
-

- 
- (10) Font, I. 1983. *Climatología de España y Portugal*. Inst. Nac. de Meteor., Madrid. 296 pp.
- (11) Gumbel, E.J. 1958. *Statistics of extremes*. Columbia University Press, New York. 375 pp.
- (12) Garrido, R. 1992. Limitaciones de la distribución de Gumbel en la valoración del riesgo de lluvias fuertes. *Ponencias y Comunicaciones del 1<sup>er</sup> Congr. Iberoamericano sobre Técnicas Aplicadas a la Gestión de Emergencias para la Reducción de Desastres Naturales*, 217-226. Univ. Politécn. de Valencia.
- (13) Garrido, R. y Caselles, V. 1995. Estudio comparativo de la valoración del riesgo de lluvias fuertes mediante las distribuciones de Gumbel y Raíz. Caso de un observatorio del litoral mediterráneo. *Avances en Geofísica y Geodesia*, **1**, 239-249.
- (14) Houghton, D. D. (ed.). 1985. *Handbook of applied meteorology*. Wiley-Interscience, New York. 1461 pp.
- (15) I. G. N. 1992. *Atlas nacional de España: Climatología (Secc. 2, Grupo 9)*. Inst. Geográf. Nac., Centro Nac. de Información Geográfica, Madrid. 4 pp. + 24 lám.
- (16) I. N. M. 1988. *Menú para tratamiento estadístico de series*. Inst. Nac. de Meteor., Secc. de Inform. Climat., Madrid. 15 pp + 2 anexos.
- (17) I. N. M. 1992. *Estilo. Manual de términos meteorológicos*. Inst. Nac. de Meteor., Madrid. 29 pp.
- (18) Jenkinson, A. F. 1955. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quart. J. Royal Meteor. Soc.*, **87**, 158-171.
-

- (19) Jenkinson, A. F. 1969. *Statistic of extremes. Estimation of Maximum Floods*. World Meteorological Organization. Tecn. Note 98, 183-227.
- (20) Jenkinson, A. F. 1975. Extreme value analysis in meteorology. *4th Conf. Probability Stat. Atmos. Sc.*, 83-89.
- (21) Kimball, B. F. 1958. Sufficient Estimation Functions and Maximum Likelihood Estimations. *Statistics of extremes*, 229-232. E. J. Gumbel. Columbia Univ. Press.
- (22) Kotz (ed.) *et al.* 1983. *Encyclopedia of Statistical Sciences*. J. Wiley and Sons, Inc. New York.
- (23) Lahore, J. P. 1992. El decenio internacional para la Reducción de Desastres Naturales. *Ponencias y Comunicaciones del 1<sup>er</sup> Congr. Iberoamericano sobre Técnicas Aplicadas a la Gestión de Emergencias para la Reducción de Desastres Naturales*, 1-6. Univ. Politécn. de Valencia.
- (24) López, F. 1993. Desertificación: factores y procesos. *La teledetección en el seguimiento de los fenómenos naturales. Climatología y Desertificación*, 183-203. S. Gandía y J. Meliá (coord.). Dpto. de Termodinámica, Univ. de Valencia.
- (25) Martín, V. 1981. Ajustes de distribuciones útiles en Ingeniería Civil. *Estadística y Simulación aplicadas a la Ingeniería Civil*, 81-100. Col. de Ing. de Caminos, Canales y Puertos. Madrid.
- (26) Martínez, I. 1986. *Estadística aplicada a la hidrometeorología*. Inst. Nac. de Meteor., Public. D-47, Madrid. 212 pp.
-

---

(27) Nobilis, F. *et al.* 1991. Statistical Considerations Concerning Probable Maximum Precipitation (PMP) in the Alpine Country of Austria. *Theor. Appl. Climatol.*, **44**, 89-94.

(28) Plate, E. J. 1979. Extreme value models. *Inputs for Risk Analysis in Water Systems*, 3-26. McBean, E. A. *et al.* (ed.). Univ. of Waterloo. Ontario.

(29) R. A. E. 1992. *Diccionario de la Lengua Española. 21ª ed.* Real Acad. Española, Madrid. 1517 pp.

(30) Redaño, A. *et al.* 1986. Climatología de las intensidades extremas de lluvia en Barcelona. *Rev. de Geofísica*, **42**, 193-198.

(31) Sevruk, B. y Geiger, H. 1981. *Selection of distribution types for extremes of precipitation.* World Meteorological Organization, No. 560. Geneva. 64 pp.

(32) Sneyers, R. 1975. *Sobre el análisis estadístico de las series de observaciones.* World Meteorological Organization, Nota Técn. nº 143, Geneva. 186 pp.

(33) Upadhyay, D. S. y Surinder, K. 1990. Finite mixture of extreme value distribution for rainfall frequency analysis. *Mausam*, **41**, 627-630.

(34) W.M.O. 1983a. *Guide to hydrological practices*, Vol. 2 (*Analysis, forecasting and other applications*). World Meteorological Organization, No. 168, Geneva.

(35) W.M.O. 1983b. *Guide to climatological practices.* World Meteorological Organization, No. 100, Geneva. 8 capt.