Universidad de Córdoba

DEPARTAMENTO
DE
INGENIERÍA RURAL

ANÁLISIS REGIONAL DE FRECUENCIAS DE VALORES EXTREMOS

José Luis Ayuso

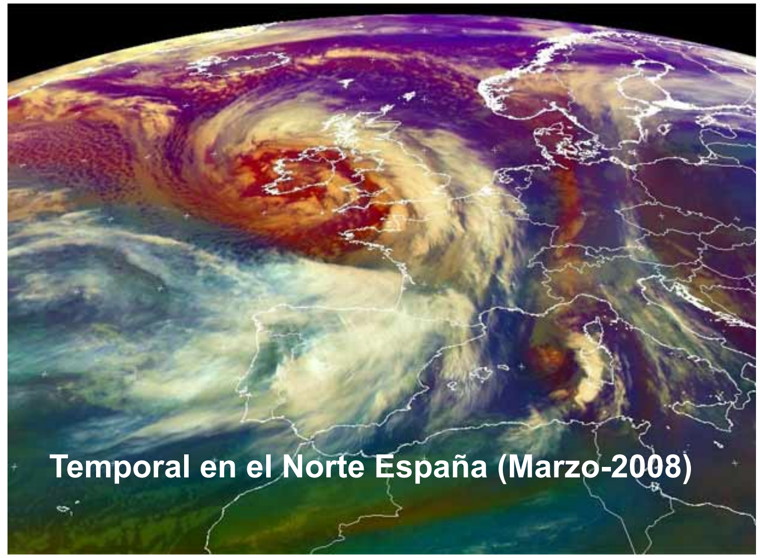
Santiago de Chile

Enero, 2012

ANÁLISIS REGIONAL DE FRECUENCIAS DE VALORES EXTREMOS

1. INTRODUCCIÓN

Los episodios hidrológicos de carácter catastrófico como las tormentas de elevada intensidad y las avenidas ocurren periódicamente produciendo cuantiosas pérdidas por daños a las propiedades e incluso pérdidas de vidas humanas, constituyendo un grave problema económico y social.



Ante estos sucesos el ingeniero ha de estimar la magnitud y frecuencia de tales eventos

Análisis local de frecuencias

Análisis regional de frecuencias

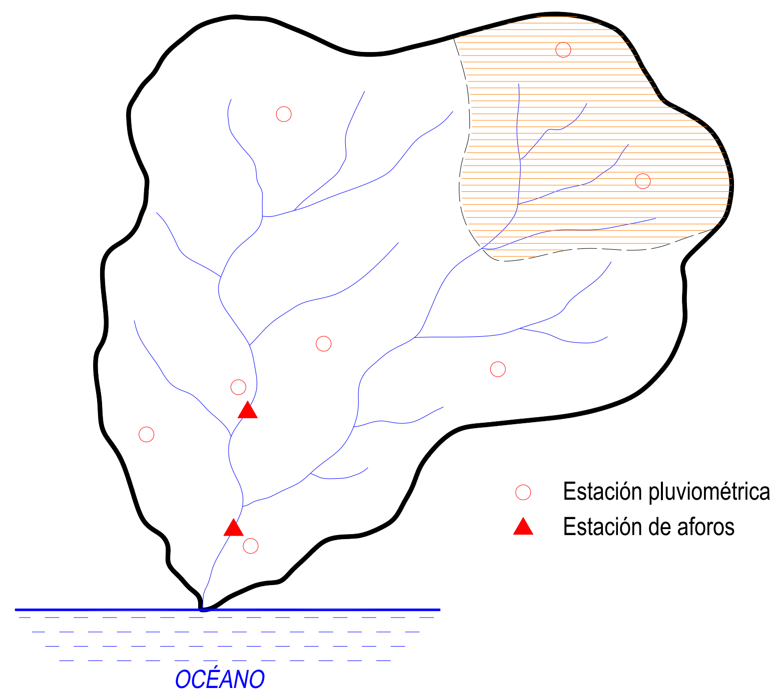
Cuestiones a responder *¿Cual es la probabilidad de que un evento extremo de determinada magnitud ocurra en cualquier año, en los próximos 50 años, o en los próximos 100 años?*

En el diseño de estructuras hidráulicas el ingeniero ha de estimar el caudal de proyecto para un determinado periodo de retorno

Datos Hidrológicos

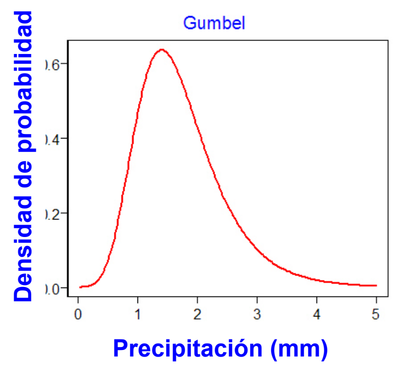
- **Caudales**
  - Medios diarios
  - Punta diarios
- **Precipitación**
  - Diaria
  - Registros pluviográficos con resolución de 5 o 10 minutos

Series de valores extremos



Objetivo

El principal objetivo del análisis de frecuencias es *relacionar la magnitud de eventos extremos con su frecuencia de ocurrencia a través del uso de distribuciones de probabilidad.*



Magnitud de un evento extremo (Cuantil) → Frecuencia de ocurrencia (Periodo Retorno)

Los datos observados durante un extenso periodo de tiempo en un sistema hidrológico se analizan mediante el análisis de frecuencias.

2. PERIODO DE RETORNO Y PROBABILIDAD

Periodo de retorno *T*, o intervalo de recurrencia, se define al *tiempo medio entre dos ocurrencias consecutivas de un fenómeno*.

Las grandes avenidas tienen grandes periodos de retorno y viceversa

$$T = \frac{1}{P} \qquad (1)$$

En Hidrología, los acontecimientos que pueden producir daños suelen ser mayores a una cierta cantidad, por ejemplo: precipitación superior a un determinado valor. La probabilidad de ocurrencia será

$$P = P(X > x) = 1 - P(X \leq x) = 1 - F(x) \qquad (2)$$

siendo:

F(x) = función de distribución acumulada del proceso.

De (1) y (2) el periodo de retorno puede definirse como $$T = \frac{1}{1 - F(x)} \qquad (3)$$

(4)

Problemas que se plantean en el análisis de frecuencias:

1. Conociendo la función de distribución del proceso que se analiza, deducir la probabilidad de ocurrencia de una magnitud dada

2. El problema inverso: determinar la magnitud del acontecimiento tal que su periodo de retorno es *T*.

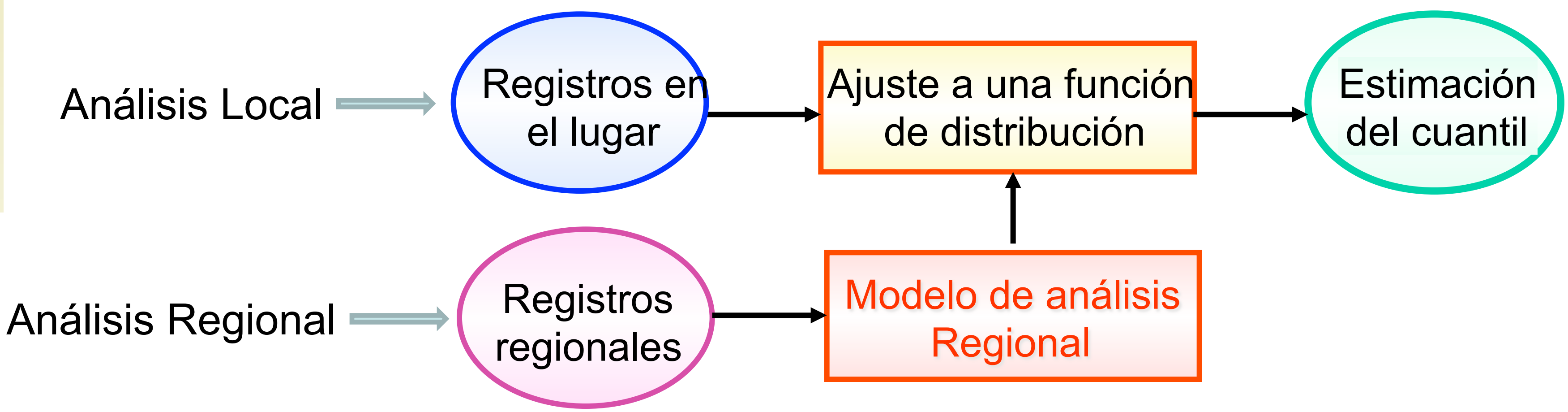
3. ANÁLISIS REGIONAL DE FRECUENCIAS DE VALORES EXTREMOS

En términos estadísticos el problema planteado en el análisis de frecuencias es estimar la cola de una distribución de probabilidad $F(x)$ desconocida, basada en un conjunto limitado de datos x_1, x_2, .. x_n.

Los pocos datos extremos disponibles hacen difícil tal estimación

La mayoría de las veces se requieren estimaciones que sobrepasan al mayor valor observado, lo que hace necesario la extrapolación más allá del intervalo observado

Los datos ajustados a una distribución no necesariamente se ajustarán bien en los extremos de la distribución



¿**Cuándo se aplica el Análisis Regional de Frecuencias**?

1. Cuando las series de valores extremos (máximos anuales) son demasiadas cortas para hacer una estimación fiable de los eventos extremos

2. Cuando no hay registros en el lugar de interés (caso general en países en desarrollo o subdesarrollados)

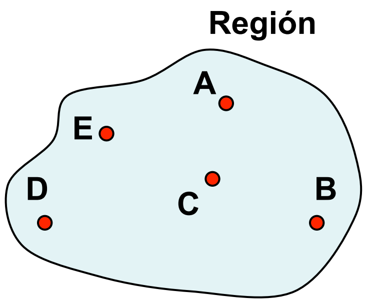
Se combinan los registros de datos (lluvia, caudal, sequía, etc..) de diferentes lugares en una región que pueda asumirse que tiene similares características (todos las series de datos de la región proceden de una misma distribución parental)

Se estima una Distribución Regional de Frecuencias de valores extremos (precipitación, avenidas, sequias, caudales bajos, etc.) para toda la región que proporciona informa-ción en lugares con datos escasos o carentes de ellos

Se reduce la incertidumbre en las estimaciones de los cuantiles

Este método asume que *M* estaciones con *N* años de registro, equivalen a una Estación-año, que proporciona análoga información que una estación con *MxN* años

Método del índice de avenida



Se establece la variable *Y*, del conjunto de datos de la región

$$Y = \left\{ \frac{x_{Ai}}{\bar{x}_A}, \frac{x_{Bi}}{\bar{x}_B}, \ldots\ldots \frac{x_{Ei}}{\bar{x}_E} \right\}$$

dividiendo los datos x_i de cada estación por la media de dicha estación.

Este método asume que los datos adimensionalizados, en cada estación, siguen la misma ley de distribución en toda la región

2. Se ajusta una ley de distribución al conjunto de datos *Y*, y se obtienen los parámetros de dicha distribución.

3. El cuantil en cada estación, *j*, se obtiene como

$$x_{T_j} = \bar{x}_j \cdot Y_T$$

Siendo Y_T el cuantil obtenido de la distribución regional.

4. MÉTODO DE LOS MOMENTOS LINEALES

Los momentos-*L* son estadísticos de muestras de datos y de distribuciones de probabilidad.

Son análogos a los momentos convencionales − proporcionan medidas de localización, dispersión, sesgo, curtosis y otros aspectos de la forma de las distribuciones de probabilidad o de las muestras de datos − pero se *calculan por combinaciones lineales de los elementos ordenados de una muestra* (de aquí el término *L*)

Ventajas de los momentos *L* sobre los momentos convencionales

Son capaces de caracterizar una amplia gama de distribuciones

Son más robustos a la presencia de valores anómalos (*outlieres)* en la muestra

Las estimaciones de los momentos *L* están menos sujetas al sesgo que los momentos convencionales

Estimación de los momentos-L a partir de datos de la muestra

La estimación de los momentos *L* se basa en la muestra de tamaño n, de los datos disponibles

Para ello, se ordenan los datos según magnitud creciente: $X_{1:n} < X_{2:n} < \ldots < X_{n:n}$

El estimador del momento de probabilidad ponderada es:

$$b_r = n^{-1}\binom{n-1}{r}^{-1}\sum_{j=r+1}^{n}\binom{j-1}{r}x_{j:n}$$

$$b_0 = n^{-1}\sum_{j=1}^{n} x_{j:n}$$

$$b_1 = n^{-1}\sum_{j=2}^{n}\frac{(j-1)}{(n-1)}x_{j:n}$$

$$b_2 = n^{-1}\sum_{j=3}^{n}\frac{(j-1)(j-2)}{(n-1)(n-2)}x_{j:n}$$

$$b_r = n^{-1}\sum_{j=r+1}^{n}\frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)}x_{j:n}$$

Los momentos *L* de la muestra se definen como

$$\boxed{\begin{aligned} \ell_1 &= b_0 \\ \ell_2 &= 2b_1 - b_0 \\ \ell_3 &= 6b_2 - 6b_1 + b_0 \\ \ell_4 &= 20b_3 - 30b_2 + 12b_1 - b_0 \end{aligned}}$$

Los cocientes de los momentos *L* de la muestra se definen como

$$C_v L = t = \frac{\ell_2}{\ell_1}$$

$$\boxed{t_r = \frac{\ell_r}{\ell_2}}$$

Terminología de los Momentos-L

Momentos lineales de la distribución	Momentos lineales de la muestra
λ_1 *es la media-L de la distribución*	ℓ_1
λ_2 *es el parámetro de escala-L*	ℓ_2
τ *es el C_v-L*	$t = \ell_2/\ell_1$
τ_3 *es el coeficiente de sesgo-L*	$t_3 = \ell_3/\ell_2$
τ_4 *es el cociente de curtosis-L*	$t_4 = \ell_4/\ell_2$

Etapas fundamentales en el ARF

1. Identificación de Regiones Homogéneas
2. Selección de la Distribución Regional de Frecuencias apropiada
3. Estimación de Cuantiles en los lugares de interés
   - Lugares aforados
   - Lugares no aforados

Cunnane (1988) y GREHYS (1996) presentaron detalladamente diversas metodologías de estimación regional para el ARF, incluyendo las tres etapas anteriores

La técnica de los momentos-L se usa en las tres etapas del ARF

1 Identificación de Regiones Homogéneas

Es la etapa más difícil y la que requiere mayor dosis de juicio subjetivo

Concepto de homogeneidad

Una región homogénea no presupone que sea una región geográfica, puesto que la proximidad geográfica no es garantía de homogeneidad.

Lugares geográficamente próximos pueden tener características muy diferentes de la variable que se analiza, sobre todo si la variabilidad espacial de las características fisiográficas e hidrológicas es grande.

Una región homogénea agrupa lugares (estaciones) con similar comportamiento estadístico (habitualmente cuantificado por su C_v)

Técnicas de identificación de regiones homogéneas

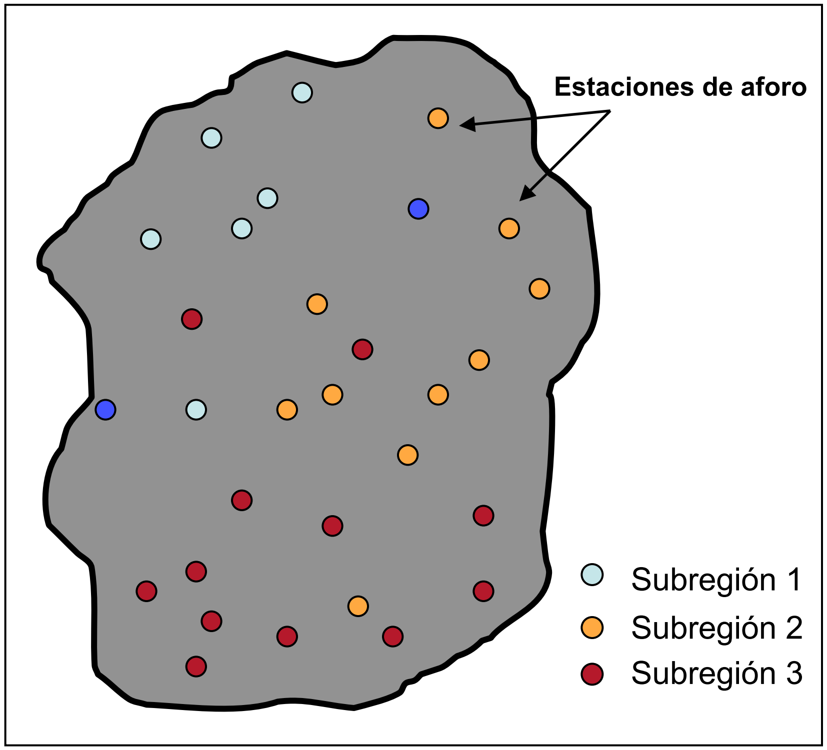
Existen diversos modos de identificar regiones homogéneas. Entre las técnicas más usuales están:

- Análisis Cluster (o análisis de conglomerados) (Jingyi y Hall, 2004)
- Análisis de Componentes Principales (Garcia-Marín y col., 2011)
- Método de los resíduos (Wiltshire, 1985; Nathan y McMahon, 1990; Jingyi y Hall, 2004)
- Redes Neuronales Artificiales (Jingyi y Hall, 2004)
- Lógica Borrosa (Fuzzy logic) (Jingyi y Hall, 2004)

Todas las técnicas requieren

- Estadísticos de cada lugar → Deducidos de las series de valores máximos anuales objeto de estudio
- Características del lugar → P.e., Precipitación media anual, media de la precipitación máxima anual en 24 h, área de la cuenca, longitud de la corriente principal, pendiente media del río, altitud del lugar, altitud media de la cuenca, y otras características geográficas o fisiográficas

Una vez identificadas tentativamente las diversas regiones homogéneas, han de calcularse en cada subregión las medidas de la discordancia y la homogeneidad basadas en los estadísticos de los momentos-L



Si algún lugar de una región es discordante con la región en su conjunto, se eliminará de la región, pudiéndose considerar la posibilidad de desplazarlo a otra región.

Medida de la discordancia

Dado un grupo de lugares o estaciones hay que identificar aquellos que sean fuertemente discordantes con el grupo como un todo.

La discordancia se mide con los momentos-L de los datos de los lugares.

Valores críticos para el estadístico de la discordancia

Nº de lugares en la región	Valor crítico	Nº de lugares en la región	Valor crítico
5	1,333	11	2,632
6	1,648	12	2,757
7	1,917	13	2,869
8	2,140	14	2,971
9	2,329	≥ 15	3
10	2,491		

Un lugar se considera discordante si el valor de D_i es mayor que el valor crítico dado en la Tabla 1

Para el lugar *i*, D_i es

$$\boxed{D_i = \frac{1}{3} N (u_i - \bar{u})^T A^{-1} (u_i - \bar{u})}$$

Siendo el vector $\mathbf{u}_i$ el traspuesto del vector que contiene los valores $\boldsymbol{t}$, $\boldsymbol{t}_3$ y $\boldsymbol{t}_4$ del lugar *i*

$$u_i = \left[t^{(i)} t_3^{(i)} t_4^{(i)} \right]^T$$

$\bar{u}$ la media del grupo

$$\bar{u} = \frac{\sum_{i=1}^{N} u_i}{N}$$

A la matriz suma de los cuadrados y productos transversales

$$A = \sum_{i=1}^{N} (u_i - \bar{u})(u_i - \bar{u})^T$$

Medida de la heterogeneidad

Se estima el grado de heterogeneidad en un grupo de lugares para evaluar si los lugares pueden, razonablemente, ser tratados como una región homogénea.

Hay que analizar si la dispersión entre lugares de los cocientes de los momentos-L de la muestra para el grupo de lugares es mayor de lo que podría esperarse de una región homogénea.

La medida de la heterogeneidad se realiza mediante un estadístico H

La región se declara: Heterogénea si H es suficientemente grande

Aceptablemente homogénea si $H < 1$

Posiblemente heterogénea si $1 \leq H \leq 2$

Definidamente heterogénea si $H > 2$

② Selección de la Distribución Regional de Frecuencias apropiada

Una vez definida la región homogénea los pasos a seguir para seleccionar la distribución regional son:

Agrupar los datos de las series de máximos anuales de las estaciones incluidas en la región previamente normalizados por el valor medio de cada estación (Método de la Avenida Índice)

Ajustar una función de distribución de valores extremos al conjunto de datos máximos anuales normalizados de la región

Las distribuciones de probabilidad más usuales para caracterizar la relación entre las magnitudes de los eventos y sus frecuencias son:.

- **Generalizada de Valores Extremos (GEV)** — Recomendada en los estudios de avenidas en Gran Bretaña
  - Gumbel (EV1)
  - Frechet
  - Weibull
- **Pearson tipo III y Log Pearson tipo III** — Utilizada de forma normativa en los EEUU
- **Lognormal (2 y 3 parámetros)**
- **Logística Generalizada**
- **Generalizada de Pareto**
- **Wakeby con límite inferior**
- **SQRT-ETmax** — De gran auge en España

En lugares de interés no aforados, la curva regional puede redimensionarse con una estimación del parámetro de normalización (típicamente el valor medio de la variable estudiada) del lugar obtenido de características de la región

Hay que establecer una relación regional de regresión entre **los valores medios de la variable analizada en las estaciones de la región** y características de cuenca e hidrológicas como variables explicativas

Se ha procedió a realizar un análisis regional de frecuencias, intraestación (in-site) con las nueve series de valores máximos anuales de la estación de Málaga (Aeropuerto)

Metodología Propuesta por Hosking y Wallis (1993, 1997)

1 Obtención de los momentos-*L* de las 9 series de valores máximos anuales

Programa RAZMOMEN

Archivo de entrada

Datos 10 minutos Malaga
4.2
6.5
9.0
31.4
6.1
28.0
12.6
6.8
23.7
7.0
7.5
7.3
7.0
23.6
18.2
8.7
16.0
6.0
5.7
4.2
15.1
8.9
11.2
13.0
7.4
15.7
6.3
9.0
12.5
8.4

Archivo de salida

Datos 10 minutos Málaga
Datos registrados ordenados por magnitud creciente

4.2	4.2	5.7	6.0	6.1	6.3	6.5	6.8	7.0	7.0
7.3	7.4	7.5	8.4	8.7	8.9	9.0	9.0	11.2	12.5
12.6	13.0	15.1	15.7	16.0	18.2	23.6	23.7	28.0	31.4

Cocientes de los Momentos-L

Numero de cocientes de momentos-L calculados: 5

Orden del Momento-L	Magnitud	
1	**11.5667**	l_1
2	**3.7103**	l_2
3	**.3722**	τ_3
4	**.1704**	τ_4
5	**.0328**	τ_5

De manera análoga se calcularon los momentos-L para cada una de las series, y los correspondientes cocientes de los momentos-L, $\tau_2=l_2/l_1$ (C_v-*L*), τ_3 (C_s-*L*) y τ_4 (C_k-*L*), así como sus promedios regionales.

SERIE	C_v-L	C_s-L	C_k-L
	τ_2	τ_3	τ_4
10’	0,3208	0,3722	0,1704
20’	0,3288	0,3482	0,1783
30’	0,3311	0,3602	0,2049
1h	0.3230	0,3281	0,1519
2h	0,3376	0,3487	0,1813
3h	0,3186	0,2957	0,1234
6h	0,3036	0,2408	0,0661
12h	0,2942	0,2529	0,0884
24h	0,2783	0,2259	0,0378
Medias	**0,3115**	**0,3081**	**0,1336**

2 Identificación si la región es homogénea

Programa XTEST

Archivo de entrada

```
9          Datos Malaga
10m    30     11.57  0.3208  0.3722  0.1704  0.0328
20m    30     16.78  0.3288  0.3482  0.1783  0.0887
30m    30     19.95  0.3311  0.3602  0.2049  0.1356
1h     30     25.04  0.3219  0.3279  0.1544  0.0748
2h     30     33.58  0.3376  0.3487  0.1813  0.1266
3h     30     40.53  0.3186  0.2957  0.1234  0.0939
6h     30     50.95  0.3063  0.2451  0.0594  0.0313
12h    30     58.52  0.2942  0.2529  0.0884  0.0487
24h    30     72.24  0.2783  0.2259  0.0378  0.0357
```

Constituido por los valores de los cocientes de los de los momentos-L obtenidos en el paso anterior

Archivo de Salida PROGRAMA XTEST

1

	Datos Malaga				9 LUGARES	
10m	30	11.57	.3208	.3722	.1704	.0328
20m	30	16.78	.3288	.3482	.1783	.0887
30m	30	19.95	.3311	.3602	.2049	.1356
1h	30	25.04	.3219	.3279	.1544	.0748
2h	30	33.58	.3376	.3487	.1813	.1266
3h	30	40.53	.3186	.2957	.1234	.0939
6h	30	50.95	.3063	.2451	.0594	.0313
12h	30	58.52	.2942	.2529	.0884	.0487
24h	30	72.24	.2783	.2259	.0378	.0357

SITE	N	NAME	L-CV	L-SKEW	L-KURT	D(I)
1	30	10m	.3208	.3722	.1704	**2.29**
2	30	20m	.3288	.3482	.1783	**.22**
3	30	30m	.3311	.3602	.2049	**1.15**
4	30	1h	.3219	.3279	.1544	**.05**
5	30	2h	.3376	.3487	.1813	**.69**
6	30	3h	.3186	.2957	.1234	**.29**
7	30	6h	.3063	.2451	.0594	**1.67**
8	30	12h	.2942	.2529	.0884	**1.13**
9	30	24h	.2783	.2259	.0378	**1.50**
WEIGHTED MEANS			.3153	.3085	.1331	

2

***** **HETEROGENEITY MEASURES** *****
(NUMBER OF SIMULATIONS = 500)

OBSERVED S.D. OF GROUP L-CV	= .0182
SIM. MEAN OF S.D. OF GROUP L-CV	= .0321
SIM. S.D. OF S.D. OF GROUP L-CV	= .0080
STANDARDIZED TEST VALUE **H(1)**	**= -1.76**
OBSERVED AVE. OF L-CV / L-SKEW DISTANCE	= .0504
SIM. MEAN OF AVE. L-CV / L-SKEW DISTANCE	= .0745
SIM. S.D. OF AVE. L-CV / L-SKEW DISTANCE	= .0170
STANDARDIZED TEST VALUE **H(2)**	**= -1.42**
OBSERVED AVE. OF L-SKEW/L-KURT DISTANCE	= .0694
SIM. MEAN OF AVE. L-SKEW/L-KURT DISTANCE	= .0952
SIM. S.D. OF AVE. L-SKEW/L-KURT DISTANCE	= .0209
STANDARDIZED TEST VALUE **H(3)**	**= -1.23**

3

***** **GOODNESS-OF-FIT MEASURES** *****
(NUMBER OF SIMULATIONS = 500)

GEN. LOGISTIC	L-KURTOSIS = .246	Z VALUE= 4.24
GEN. EXTREME VALUE	L-KURTOSIS = .220	Z VALUE= 3.26
GEN. NORMAL	L-KURTOSIS = .198	Z VALUE= 2.42
PEARSON TYPE III	L-KURTOSIS = .159	Z VALUE= .97 *
GEN. PARETO	L-KURTOSIS = .148	Z VALUE= .55 *

4

PARAMETER ESTIMATES FOR DISTRIBUTIONS ACCEPTED AT THE 90% LEVEL

PEARSON TYPE III	1.000	.620	1.852		
GEN. PARETO	.351	.685	.057		
WAKEBY	.367	-.258	1.956	.818	-.135

QUANTILE ESTIMATES

	.500	**.800**	**.900**	**.960**	**.980**	**.990**
PEARSON TYPE III	.821	1.394	1.815	2.365	2.777	3.188
GEN. PARETO	.817	1.406	1.831	2.368	2.755	3.128
WAKEBY	.810	1.423	1.854	2.369	2.719	3.038

PROCESADOS TODOS LOS DATOS

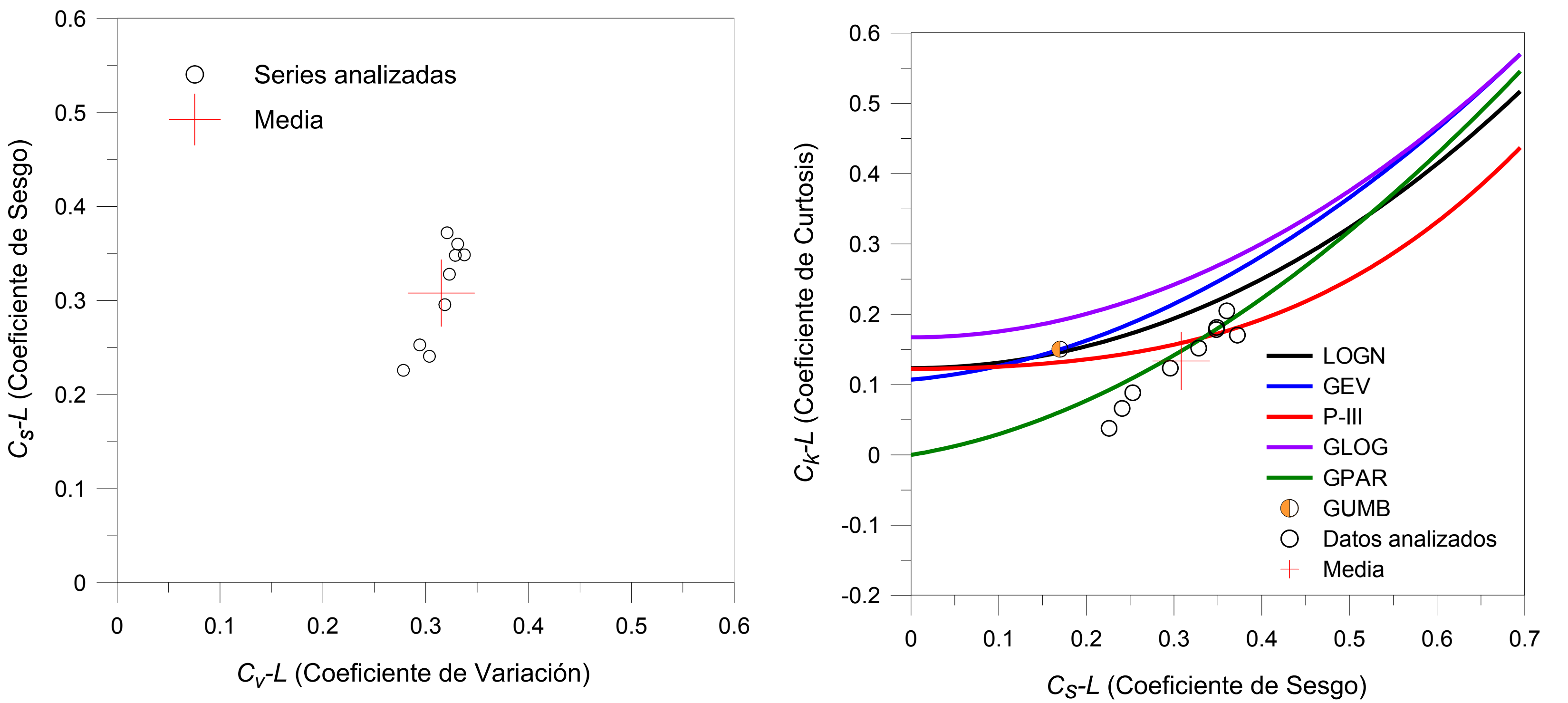
Para una región con 9 lugares el valor crítico D_i de la discordancia es **2.329** ⟹ No hay ningún lugar discordante

Puesto que H < 1, la región es homogénea

3 Selección de la Distribución Regional de Frecuencias apropiada

Comprobada la homogeneidad de la región ha de seleccionarse una función de distribución de frecuencias para ajustarla a los datos normalizados de la región

Un primer test de bondad de ajuste de la distribución puede realizarse con los diagramas de los cocientes de los momentos-L



Puede observarse que la distribución que más se aproxima al valor medio de C_s-L y C_k-L de la región, y de los valores de las series analizadas es la Distribución Generalizada de Pareto (GPAR) seguida de la Pearson III

También la salida del Programa **XTEST** especifica la distribución de frecuencias más apropiada en función del índice **Z_{dist}** siendo el valor límite de Z_{dist} ⟹ $|Z_{dist}| < 1.64$

4 Ajuste de la distribución seleccionada a los datos de la región

Archivo de entrada

```
Análisis Regional datos Aeropuerto de Málaga
 9
10 minutos
 30
   4.2    6.5    9.0  31.4   6.1  28.0   12.6   6.8  23.7   7.0
   7.5    7.3    7.0  23.6  18.2   8.7   16.0   6.0   5.7   4.2
  15.1    8.9   11.2  13.0   7.4  15.7    6.3   9.0  12.5   8.4
20 Minutos
 30
   6.4    8.8   12.8  45.5   6.7  43.4   21.7   9.5  35.2  12.4
  13.0   10.6    9.3  35.4  21.8  15.1   22.6   8.7   7.0   6.4
  21.6    2.4   14.5  20.6   8.7  19.8    8.6  14.0  15.4  15.4
30 Minutos
 30
   8.6    9.9   16.1  61.7   7.4  48.7   25.5  10.7  40.4  16.1
  18.3   11.5   11.1  38.4  27.6  17.5   26.0  10.3   8.7   7.7
  22.7   14.9   15.9  26.8   9.3  21.4   10.7  21.3  16.5  16.7
1 Hora
  30
  11.3   18.5   20.0  67.9   8.2  50.9   38.9  11.6  61.0  16.3
  23.0   12.8   13.7  42.6  38.6  23.8   31.9  11.9  15.5  10.9
  25.0   21.6   18.8  35.7  13.0  27.4   12.1  29.1  17.9  21.3
2 Horas
 30
  14.3   34.4   20.8  73.0  13.3  54.5   59.1  20.2 112.5  16.7
  33.2   22.8   15.1  44.0  41.1  30.7   61.2  12.8  22.3  16.2
  28.3   27.6   23.2  52.9  14.6  34.1   12.8  45.7  24.4  25.7
3 Horas
 30
  18.1   48.0   24.4  77.6  21.0  60.1   60.8  21.9 129.3  17.7
  36.4   26.1   18.7  65.1  51.5  31.4   64.2  13.4  24.3  17.6
  30.2   33.3   29.8  55.7  19.0  37.5   27.9  59.5  30.5  64.8
6 Horas
 30
  22.0   86.5   34.8  85.7  21.0  62.8   77.8  30.5 134.3  19.1
  42.0   33.7   25.8  88.3  58.9  48.7   82.2  24.4  28.1  22.4
  34.8   33.3   38.8  65.1  23.4  49.5   45.7  92.5  42.3  74.2
12 Horas
 30
  27.1  127.5   53.0  94.3  27.3  62.8   88.5  34.6 138.5  21.9
  46.2   39.2   32.4  96.0  59.2  53.5   85.0  27.9  35.2  27.4
  46.2   35.5   42.6  65.6  26.3  63.9   46.3 104.0  62.8  85.0
24 Horas
 30
  34.2  162.5   75.4 126.6  29.7  82.9   90.3  40.8 143.6  39.2
  60.0   45.0   34.5 100.7  63.9  69.3   85.1  46.0  46.1  35.4
  48.4   43.7   56.1 110.5  37.5 113.3   46.3 109.2  95.5  95.6
```

Programa XFIT

Ajusta la distribución seleccionada a los datos de la región

Archivo de salida programa XFIT

ANALISIS REGIONAL

PARAMETROS DE REPRESENTACION DE POSICION DE LOS MOMENTOS-L -0.3500 0.0000

LUGAR 1 10 minutos	N= 30	COCIENTES DE MOMENTOS-L	11.57	3.7023	0.3643	0.1836	0.0540
LUGAR 2 20 Minutos	N= 30	COCIENTES DE MOMENTOS-L	16.78	5.4999	0.3429	0.1883	0.0942
LUGAR 3 30 Minutos	N= 30	COCIENTES DE MOMENTOS-L	19.95	6.5835	0.3538	0.2104	0.1288
LUGAR 4 1 Hora	N= 30	COCIENTES DE MOMENTOS-L	25.04	8.0428	0.3244	0.1685	0.0831
LUGAR 5 2 Horas	N= 30	COCIENTES DE MOMENTOS-L	33.58	11.2948	0.3437	0.1902	0.1214
LUGAR 6 3 Horas	N= 30	COCIENTES DE MOMENTOS-L	40.53	12.8857	0.2951	0.1421	0.0950
LUGAR 7 6 Horas	N= 30	COCIENTES DE MOMENTOS-L	50.95	15.5980	0.2490	0.0889	0.0475
LUGAR 8 12 Horas	N= 30	COCIENTES DE MOMENTOS-L	58.52	17.2273	0.2556	0.1140	0.0611
LUGAR 9 24 Horas	N= 30	COCIENTES DE MOMENTOS-L	72.24	20.1541	0.2308	0.0735	0.0489

Análisis Rgional datos Aeropuerto de Malaga

COCIENTES PROMEDIOS REGIONALES DE LOS MOMETOS-L 1.0000 0.3148 0.3066 0.1511 0.0816

PARAMETROS REGIONALES DE LA DISTRIBUCION PARETO GENERALIZADA
0.3512 0.6887 0.0613 0.0000 0.0000

LUGAR				CUANTILES						
NUMERO	2.	3.	4.	5.	10.	15.	20.	25.	50.	100.
REGION	**0.818**	**1.083**	**1.266**	**1.407**	**1.830**	**2.066**	**2.236**	**2.363**	**2.746**	**3.114**
1	9.47	12.53	14.65	16.27	21.17	23.91	25.86	27.33	31.77	36.02
2	13.73	18.18	21.25	23.60	30.70	34.67	37.52	39.64	46.07	52.24
3	16.33	21.61	25.26	28.06	36.51	41.23	44.60	47.13	54.78	62.11
4	20.49	27.13	31.71	35.22	45.82	51.75	55.99	59.16	68.77	77.97
5	27.49	36.38	42.53	47.24	61.46	69.41	75.10	79.35	92.23	104.58
6	33.17	43.91	51.32	57.00	74.16	83.76	90.62	95.75	111.30	126.20
7	41.70	55.13	64.44	71.67	93.24	105.17	113.79	120.39	139.93	158.66
8	47.90	63.40	74.11	82.32	107.09	120.96	130.87	138.27	160.72	182.24
9	59.13	78.27	91.48	101.61	132.20	149.31	161.55	170.69	198.40	224.96

Ajuste de la distribución Generalizada de Pareto y Pearson III a los datos regionales

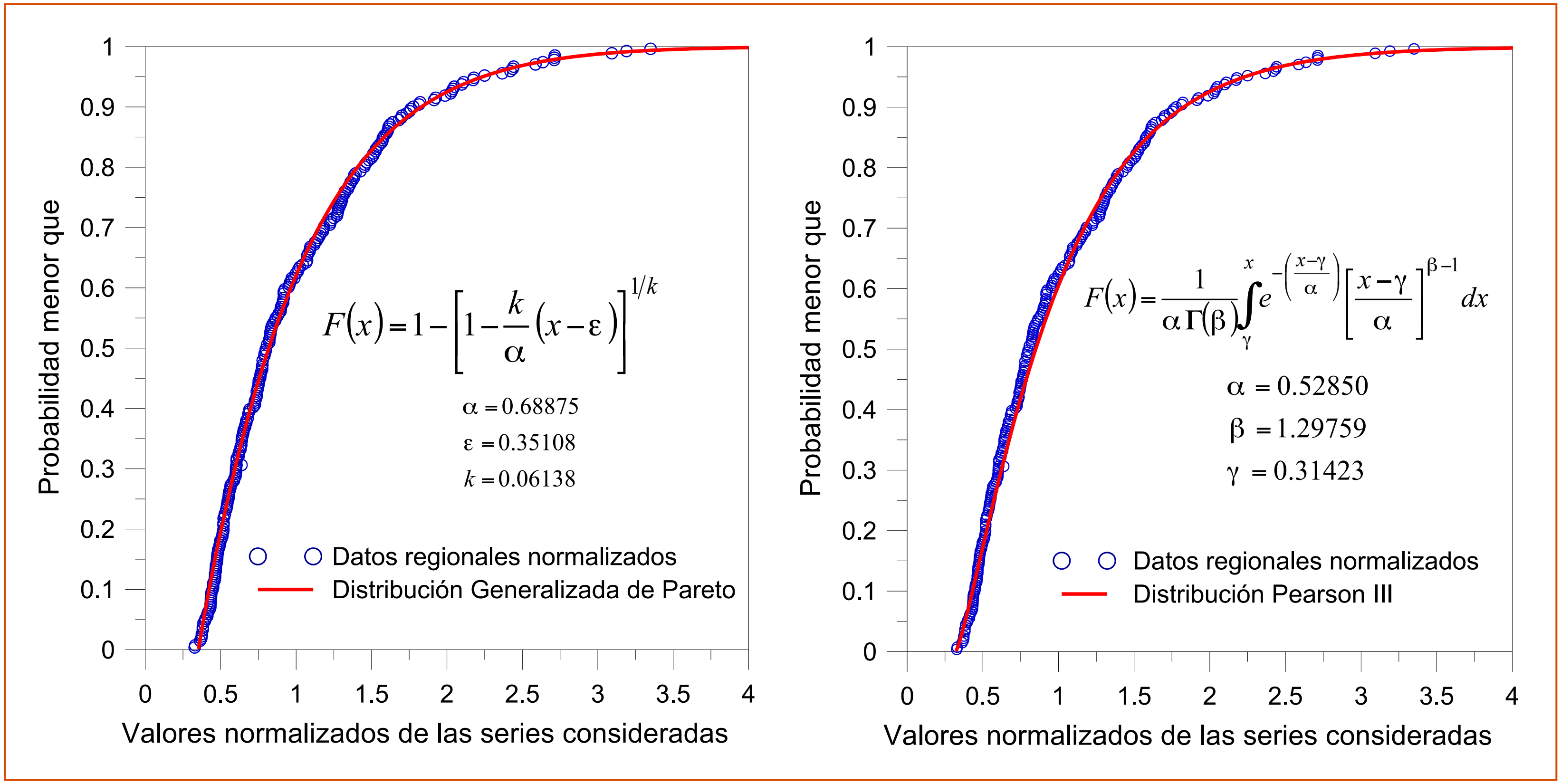


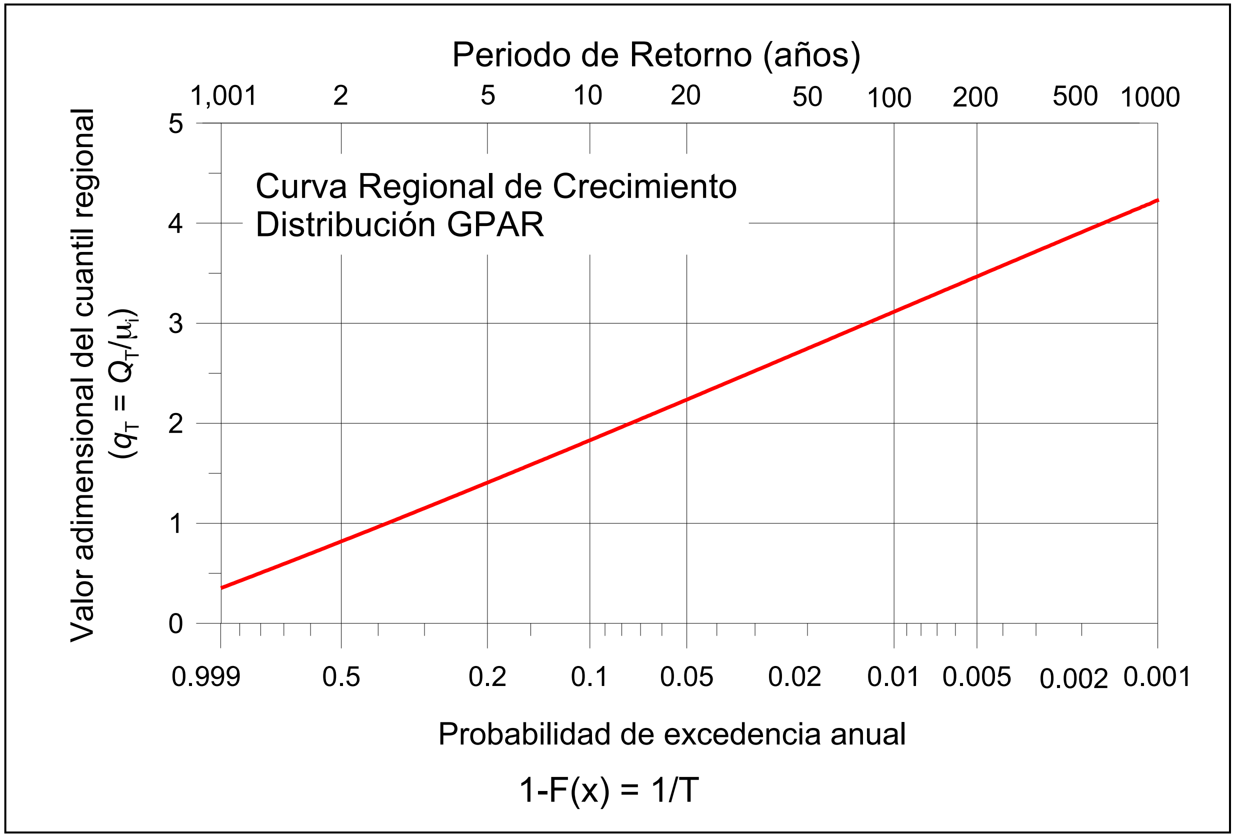
Tabla 1 Cuantiles obtenidos con la función de distribución GPAR

Serie	Periodo de Retorno									
	2	3	4	5	10	15	20	25	50	100
Región	**0.818**	**1.083**	**1.266**	**1.407**	**1.830**	**2.066**	**2.236**	**2.363**	**2.746**	**3.114**
10 min	9.47	12.53	14.65	16.27	21.17	23.90	25.86	27.33	31.77	36.02
20 min	13.73	18.18	21.25	23.60	30.70	34.67	37.51	39.64	46.07	52.29
30 min	16.33	21.61	25.26	28.06	36.50	41.22	44.59	47.13	54.78	62.11
1 h	20.49	27.13	31.71	35.22	45.82	51.74	55.98	59.16	68.77	77.97
2 h	27.49	36.39	42.53	47.24	61.46	69.40	75.08	79.35	92.23	104.58
3 h	33.17	43.91	51.32	57.00	74.16	83.75	90.61	95.75	111.30	126.20
6 h	41.70	55.20	64.53	71.67	93.24	105.29	113.92	120.39	139.93	158.60
12 h	47.90	63.41	74.11	82.32	107.09	120.94	130.84	138.27	160.72	182.24
24 h	59.13	78.27	91.49	101.61	132.20	149.29	161.51	170.69	198.40	224.96

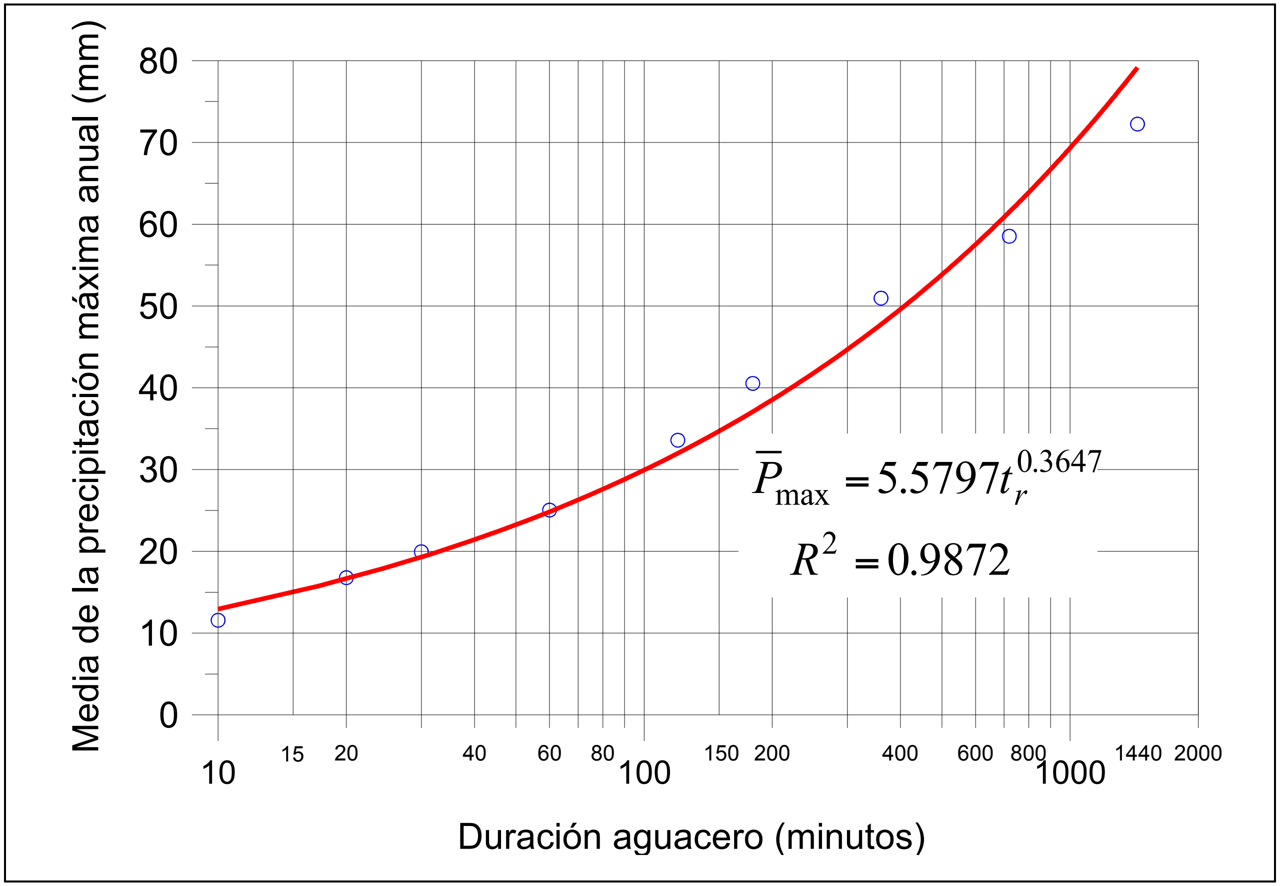
Tabla 2 Cuantiles obtenidos con la función de distribución Pearson-III

Serie	Periodo de Retorno									
	2	3	4	5	10	15	20	25	50	100
Región	**0.823**	**1.080**	**1.257**	**1.394**	**1.813**	**2.053**	**2.227**	**2.360**	**2.770**	**3.178**
10 min	9.52	12.49	14.54	16.12	20.97	23.75	25.76	27.30	32.04	36.76
20 min	13.80	18.11	21.09	23.38	30.42	34.44	37.37	39.59	46.47	53.31
30 min	16.41	21.53	25.07	27.80	36.17	40.95	44.43	47.07	55.25	63.39
1 h	20.60	27.03	31.48	34.90	45.40	51.41	55.78	59.10	69.36	79.57
2 h	27.63	36.26	42.22	46.81	60.90	68.95	74.81	79.26	93.03	106.72
3 h	33.34	43.75	50.94	56.49	73.49	83.20	90.27	95.64	112.26	128.79
6 h	41.92	55.01	64.05	71.02	92.39	104.61	113.50	120.25	141.14	161.92
12 h	48.14	63.18	73.57	81.57	106.12	120.15	130.36	138.12	162.11	185.98
24 h	59.43	77.99	90.81	100.69	131.00	148.32	160.92	170.50	200.11	229.58

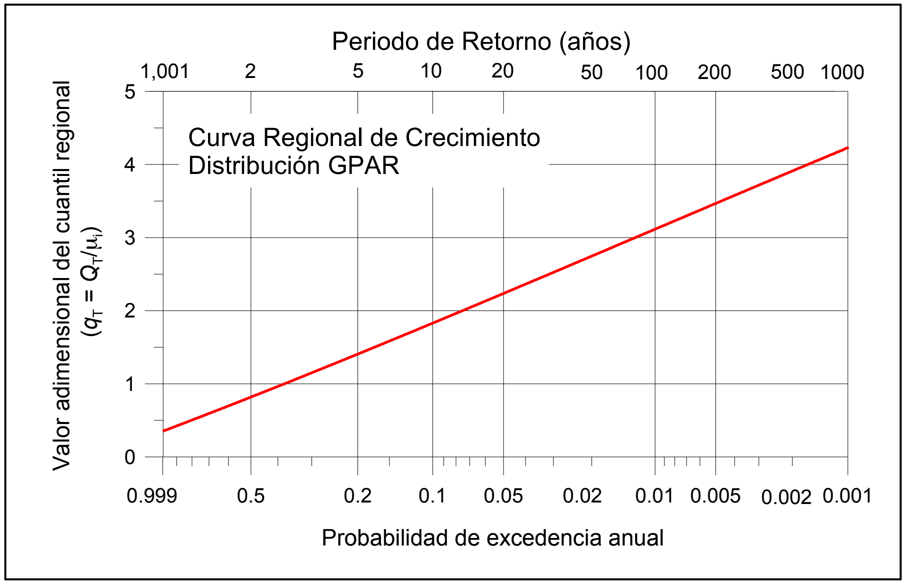
Una vez realizado el ajuste a la distribución apropiada, se construye la curva regional de crecimiento

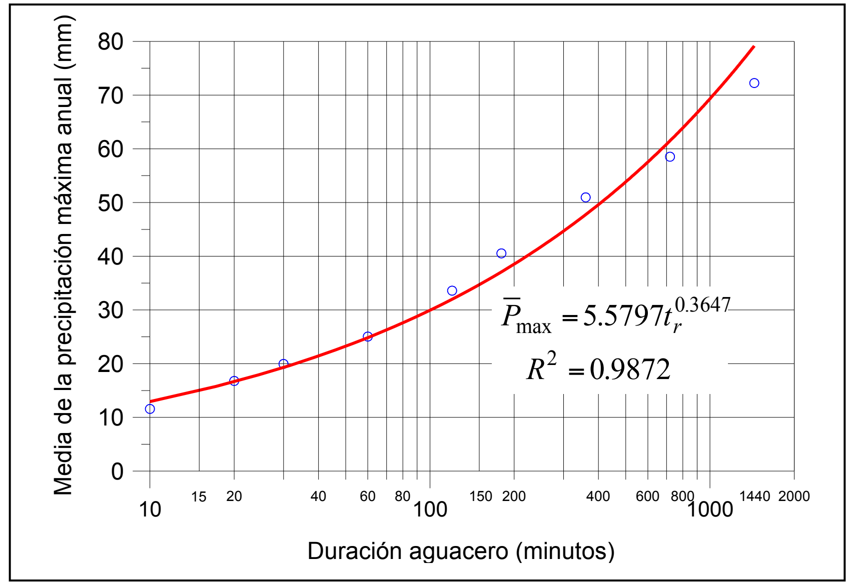


Relación regional entre los valores medios de las series de precipitaciones máximos anuales y la duración



Uso de las relaciones regionales obtenidas





Ejemplo de uso de las relaciones

Se quiere estimar la precipitación de 1h de duración y 100 años de periodo de retorno en la localidad de Málaga

Solución

De la Curva Regional de Crecimiento, para T = 100 años $\Rightarrow$ el valor adimensional del cuantil regional es $q_T = 3{,}11$

En consecuencia, el valor del cuantil para la duración de 1 h será $Q_T = q_T\ \mu_i$

Se necesita conocer la magnitud de μ_i (valor medio de la precipitación máxima anual de 1 h de duración) que según la relación de la segunda gráfica es $\mu_i = 24{,}8$

Por consiguiente, la precipitación de 1 h y 100 años será:

$$\boxed{Q_T = P_{100}^{1h} = q_T \times \mu_i = 3{,}11 \times 24{,}8 = 77{,}1\,mm}$$

Valor sensiblemente igual al dado en las Tabla de la distribución Generalizada de Pareto

DESARROLLO DE UNA RELACION Altura-Duración-Frecuencia PARA MÁLAGA

La Curva Regional de Crecimiento corresponde a la Función de Distribución Generalizada de Pareto ajustada a los valores adimensionales de las series de máximos anuales consideradas (región).

$$F(x) = 1 - \left[1 - \frac{k}{\alpha}(x - \varepsilon)\right]^{1/k} \qquad (1)$$

Siendo x el cuantil regional adimensional, q_T

La relación entre el periodo de retorno T, y F(x) es

$$T = \frac{1}{1 - F(x)} \qquad (2)$$

Combinando ambas ecuaciones (1) y (2) $\Longrightarrow$

$$\frac{1}{T} = \left[1 - \frac{k}{\alpha}(x - \varepsilon)\right]^{1/k}$$

$$\boxed{q_T = x = \left[1 - \left(\frac{1}{T}\right)^k\right]\frac{\alpha}{k} + \varepsilon} \qquad (3)$$

De la regresión entre los valores medios de las precipitaciones máximas y la duración, se tiene

$$\boxed{\mu_i = \overline{P}_{\max} = 5{,}5797\, t_r^{0{,}3647}} \qquad (4)$$

Combinando las ecuaciones (3) y (4) se obtiene el cuantil de la precipitación para un periodo de retorno T y duración t_r, $Q_T = P_T^{t_r} = q_t\, \mu_i$ resultando

$$\boxed{P_T^{t_r} = \left\{\left[1 - \left(\frac{1}{T}\right)^k\right]\frac{\alpha}{k} + \varepsilon\right\}\left(5.5797\, t_r^{0.3647}\right)} \qquad (5)$$

Sustituyendo en (5) los valores de los parámetros, α, *k* y ε, se obtiene la relación Altura-Duración-Frecuencia de la precipitación en Málaga Capital

$$P_T^{t_r} = \left\lfloor 62.6103\left(1-(1/T)^{0.06138}\right)+1.9589\right\rfloor t_r^{0.3647} \qquad (6)$$

Siendo $P_T^{t_r}$ = altura de precipitación de T años de periodo de retorno y duración t_r
T = periodo de retorno (años)
t_r = duración del aguacero (minutos)

SISTEMA DE ESTIMACIÓN DE EVENTOS EXTREMOS DE LLUVIA

FONDEF

Universidad de Talca EIAS Universidad de Córdoba