



UNIVERSITAT DE VALÈNCIA

Facultat de Física
Departament de Física de la Terra i
Termodinàmica

En colaboración con

AZTI-Tecnalia
Unidad de Investigación Marina




Nuevos modelos de probabilidad de la precipitación: Aplicación a España, y en particular, al País Vasco

Tesis presentada por
ROBERT MONJO I AGUT
(*Roberto Moncho Agud*)
Para optar al título de Doctor

Dr. VICENTE CASELLES MIRALLES, Catedrático de Física Aplicada de la Facultat de Física de la Universitat de València y Dr. GUILLEM CHUST PETERS, Investigador Senior de la Unidad de Investigación Marina de AZTI-Tecnalia,

CERTIFICAN: Que la presente Memoria *“Nuevos modelos de probabilidad de la precipitación: Aplicación a España, y en particular, al País Vasco”*, ha sido realizada bajo nuestra dirección en el Departament de Física de la Terra i Termodinàmica de la Facultat de Física, en colaboración con la Unidad de Investigación Marina de AZTI-Tecnalia, por Robert Monjo i Agut, y constituye su Tesis Doctoral dentro del Programa de Doctorado código 345D titulado, “Física del Medio Ambiente” del citado departamento.

Y para que así conste, en cumplimiento de la legislación vigente firmamos el presente certificado en Burjassot, a 21 de noviembre de dos mil once.



Fdo.: Vicente Caselles Miralles



Fdo.: Guillem Chust Peters



VNIVERSITAT ID VALÈNCIA

Acta de lectura de tesi doctoral

(Art. 21 del RD 1393/2007)

ESTUDIANT/A

Nom i cognoms: Roberto Moncho Agud

Data de naixement: 22/12/1984

D.N.I.: 20039410-Q

Núm. de registre: 5590

Títol del treball: Nuevos modelos de probabilidad de la precipitación: Aplicación a España, y en particular, al País Vasco.

Departament/Institut: Fís. De la terra i Termodinàmica

Director/a: Vicente Caselles Miralles

Guillem Chust Peters

TRIBUNAL

President/a: FERNANDO BELDA ESPLUGUES

1r. Vocal: JAVIER MARTÍN VIDE

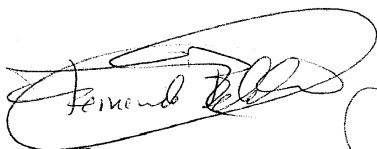
2n. Vocal: JON SAENZ AGUIRRE

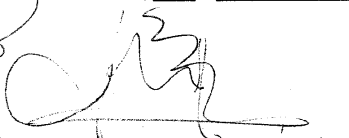
3r. Vocal: JORGE TAMAYO CARMONA


Secretari/a: CÉSAR COLL COMPANYY

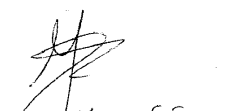
QUALIFICACIÓ: EXCELLENT CUM LAUDE

València, 31 d ENERO de 2012


Signat: FERNANDO BELDA ESPLUGUES


Signat: JAVIER MARTÍN VIDE


Signat: JORGE TAMAYO CARMONA


Signat: JON SAENZ


Signat: CÉSAR COLL COMPANYY

Art. 21.7 - Les qualificacions que es poden atorgar són: "NO APTA", "APROVAT", "NOTABLE", "EXCELLENT", i "EXCELLENT CUM LAUDE". (En cap cas s'hi farà servir la fórmula "per unanimitat").

COMISSIÓ DE DOCTORAT

A mi familia

A Darío

El presente estudio ha sido posible gracias a la beca Iñaki Goenaga de la Fundación Centros Tecnológicos, con la que he podido formarme en la unidad de Investigación Marina de AZTI-Tecnalia. Este trabajo se enmarca en un proyecto financiado por el Programa de subvenciones Ertortek del Gobierno Vasco (proyecto K-Egokitzen) y por el Departamento de Medio Ambiente, Planificación Territorial, Agricultura y Pesca (Proyecto Ref.: 0.39/SGTB/2007/4.1). Además, la elaboración de este trabajo no hubiera sido posible sin la ayuda y colaboración de ciertas personas e instituciones, a las cuales deseo expresar aquí mi más sincero agradecimiento:

En primer lugar a mis directores, el Dr. Vicente Caselles Miralles y el Dr. Guillem Chust Peters, quienes me ofrecieron la posibilidad de iniciarme en el mundo de la investigación y despertaron en mí el interés por los temas relacionados con este trabajo. A ellos tengo que agradecer su tiempo y disponibilidad para atender mis inquietudes, así como la confianza que depositaron en mí desde un primer momento. En especial hay que agradecer también al Dr. Fernando Belda, Delegado de la Agencia Estatal de Meteorología (AEMET) en Murcia, por su apoyo y empuje inicial durante las prácticas que desembocaron en muchos de los aspectos científicos del actual trabajo.

A Margarita Martín, Delegada de la AEMET en el País Vasco y a José Ángel Núñez, jefe de Climatología de la AEMET en Valencia, por la aportación de todos los datos necesarios para la elaboración de este estudio.

A Maddalen Mendizábal por su apoyo incondicional y por plantear la cuestión que dio pie a gran parte de esta tesis: ¿Cómo se puede corregir la precipitación diaria de los modelos climáticos? Gracias por tus inquietudes científicas.

A mis compañeros de despacho de aquí y allá: Juan Manuel Sánchez, Raquel Niclòs, Joan Miquel Gálvez, Maitane Grande, Richard Curtin, Nerea Goikoetxea, Stefani Novoa, Francisco Javier Pórtolles, Juan Quintana y Emma Gaitán, por toda la ayuda brindada que me ha servido para resolver muchas dudas.

A todos mis compañeros del Departament de Física de la Terra i Termodinàmica; en especial a Sandra, Almudena, Amparo, Alex y Cristina, así como a todos mis compañeros de la Unidad de Investigación Marina de AZTI-Tecnalia, por hacer más agradable y ameno el trabajo del día a día.

Especialmente, a mis padres, por su apoyo en todos los estudios cursados y en todos los trabajos realizados, porque sin ellos sería imposible haber llegado hasta aquí. Gracias por vuestra ayuda incondicional.

A Darío, porque contigo todo es más fácil. Gracias por todos los momentos que hemos vivido juntos. Gracias por entenderme y apoyarme durante todo este tiempo, a pesar de la gran carga de trabajo, y el poco tiempo libre. Gracias por hacer tan llevadero todo ese esfuerzo. Gracias por todo.

RESUMEN

El agua en España es un bien escaso, pero en algunas ocasiones puede derivar en un riesgo de inundaciones que repercute no sólo en el medio urbano y en la actividad humana, sino también en el medio natural. Ese es el caso del litoral mediterráneo, pero también cada vez más de otras regiones como la vertiente cantábrica. En el contexto del cambio climático, el estudio de las precipitaciones en el País Vasco es de gran trascendencia para la planificación del territorio y la conservación de los ecosistemas.

Uno de los campos no resueltos totalmente por la climatología es la modelización probabilística de las precipitaciones intensas a escala diaria y subdiaria. Por ello, resulta importante desarrollar una metodología que permita corregir los errores sistemáticos de la precipitación proyectada por los modelos climáticos. En este trabajo, se han construido dos modelos de probabilidad basados, respectivamente, en una modificación de la distribución de Gumbel y otra de la distribución Log-Logística.

Los nuevos modelos propuestos han sido sometidos a dos pruebas. En primer lugar se ha estudiado la bondad de ajuste para 108 observatorios de España. En este caso se analizaron las relaciones climáticas de los parámetros de ajuste y se vio que los modelos de probabilidad se ajustan mejor para los climas lluviosos que para los secos. En la segunda prueba, la misma metodología ha sido empleada para corregir la distribución de probabilidad de diferentes modelos regionales del clima, en base a 156 observatorios del País Vasco. En concreto, se transformó la distribución de probabilidad diaria de 7 modelos regionales del clima (Aladin-ARPEGE, CLM-HadCM3Q0, HIRHAM-HadCM3Q0, HIRHAM-BCM, RECMO-ECHAM5-rt3, REMO-ECHAM-rt3 y PROMES-HadCM3Q0.).

Los resultados mostraron que el método desarrollado en esta tesis corrige adecuadamente las distribuciones de probabilidad de los modelos, obteniendo un error relativo medio que en general fue inferior al 5% sobre la simulación de la media, y del 1% sobre la simulación del número de días sin precipitaciones.

Con ello, a lo largo del siglo XXI y bajo el escenario A1B, cabe esperar una intensificación de la precipitación diaria y subdiaria. Así por ejemplo, la precipitación con un período de retorno de 100 años aumentará hasta un 30% en 2001-2040 respecto a 1961-2000, en áreas del oeste de Vizcaya. Para el período 2041-2080, el aumento puede incluso superar el 50% en la mayor parte del territorio. Esta intensificación se interpreta como una consecuencia del proceso de mediterraneización que experimentaría el clima oceánico más meridional, si se cumple la expansión de la célula de Hadley prevista por los modelos climáticos globales.

La principal novedad de ese trabajo ha sido el desarrollo de los dos nuevos modelos de probabilidad, así como su aplicación para la corrección del error sistemático de series temporales de precipitación. Sin embargo, estos modelos también pueden emplearse para otros tratamientos de datos, tales como rellenar los huecos de series incompletas o corregir inhomogeneidades. Finalmente, otra de las novedades de este estudio es la profundización matemática del concepto de la intensidad, mediante un índice n que hace referencia al comportamiento de la intensidad máxima de la precipitación en función del tiempo.

ÍNDICE

RESUMEN	9
CAPÍTULO I. INTRODUCCIÓN	13
I.1. MOTIVACIÓN	15
I.2. CONCEPTOS PREVIOS	16
I.2.1. Intensidad de la precipitación	16
a) Definición.....	16
b) Caracterización.....	17
c) Relación entre intensidad, duración y frecuencia	18
I.2.2. Distribuciones de probabilidad clásicas	19
a) Distribución normal y semi-normal	19
b) Distribución log-normal y semi log-normal	20
c) Distribuciones gamma.....	21
d) Distribuciones exponenciales	22
d) Distribuciones potenciales	23
e) Distribuciones logísticas	24
I.2.3. Corrección del error sistemático de la precipitación.....	25
a) Corrección lineal	25
b) Corrección geométrica.....	26
c) Corrección probabilística.....	28
I.3. OBJETIVOS Y NOVEDADES.....	29
CAPÍTULO II. NUEVOS MODELOS DE PROBABILIDAD.....	31
II.1. NUEVA VERSIÓN DE LA DISTRIBUCIÓN DE GUMBEL OPUESTA	33
II.1.1. Probabilidad acumulada.....	33
II.1.2. Densidad de probabilidad.....	33
II.1.3. Función inversa.....	33
II.2. NUEVA VERSIÓN DE LA DISTRIBUCIÓN LOG-LOGÍSTICA GENERALIZADA	34
II.2.1. Probabilidad acumulada.....	34
II.2.2. Función de densidad	34
II.2.3. Función inversa.....	35
II.3. DISTRIBUCIÓN TEMPORAL	36
CAPÍTULO III. PRIMER TEST: AJUSTE A LOS CLIMAS DE ESPAÑA.....	37
III.1. INTRODUCCIÓN	39
III.2. MATERIAL.....	39
III.2.1. Área de estudio	39
III.2.2. Datos de precipitación.....	41
III.3. METODOLOGÍA.....	41
III.3.1. Modelos de probabilidad	41
III.3.2. Inferencia estadística	42
III.4. RESULTADOS.....	42
III.5. DISCUSIÓN	45
III.5.1. Ventajas y limitaciones	45
III.5.2. Relaciones climáticas.....	48
III.6. CONCLUSIONES	50
III.6.1. Sobre la bondad de ajuste.....	50
III.6.2. Sobre las relaciones climáticas.....	50

CAPÍTULO IV. SEGUNDO TEST: CORRECCIÓN DE MODELOS CLIMÁTICOS EN EL PAÍS VASCO	52
IV.1. INTRODUCCIÓN	54
IV.2. MATERIALES	55
IV.2.1. Área de estudio.....	55
IV.2.2. Datos de precipitación observada.....	56
IV.2.3. Modelos regionales del clima.....	57
IV.3. METODOLOGÍA.....	58
IV.3.1. Distribución de probabilidad	58
IV.3.2. Estandarización probabilística	59
IV.3.3. Ampliación de series temporales.....	59
IV.3.4. Transformación probabilística de los modelos climáticos.....	60
IV.3.5. Análisis de los extremos de precipitación	62
IV.4. RESULTADOS.....	62
IV.4.1. Rellenado y transformación probabilística	62
IV.4.2. Validación de la transformación probabilística.....	65
IV.4.3. Análisis del cambio de los extremos de precipitación.....	68
IV.5. DISCUSIÓN	72
IV.5.1. Elección de la metodología	72
IV.5.2. Transformación de la probabilidad.....	73
IV.5.3. Análisis del cambio de la precipitación extrema.....	76
IV.6. CONCLUSIONES	77
IV.6.1. Sobre la metodología	77
IV.6.2. Sobre el cambio de la precipitación intensa	77
CAPÍTULO V. CONCLUSIONES GENERALES.....	79
V.1. AJUSTE A LOS CLIMAS DE ESPAÑA.....	81
V.1.1. Sobre la metodología.....	81
V.1.2. Sobre las relaciones climáticas	81
V.2. CORRECCIÓN DE ESCENARIOS DE PRECIPITACIÓN DEL PAÍS VASCO.....	81
V.2.1. Sobre la metodología.....	81
V.2.2. Sobre el cambio de la precipitación intensa.....	82
V.3. SOBRE EL ÍNDICE N	82
GLOSARIO DE TÉRMINOS Y ACRÓNIMOS.....	83
BIBLIOGRAFÍA	91
ANEXO 1. AN ALTERNATIVE MODEL FOR PRECIPITATION PROBABILITY DISTRIBUTION	99
ANEXO 2. PROBABILISTIC CORRECTION OF RCM PRECIPITATION FOR THE BASQUE COUNTRY (NORTHERN SPAIN)	130

CAPÍTULO I. INTRODUCCIÓN

I.1. Motivación

Una de las principales preocupaciones medioambientales de la sociedad se focaliza en los posibles efectos del cambio global. El agua es el recurso natural más importante para la mayoría de actividades humanas, pero sobretodo para los ecosistemas naturales, con los que convivimos y nos nutrimos. A menudo, el agua es un bien escaso, y claramente limitante en los medios naturales y actividades como la agricultura y la industria. Pero en algunas ocasiones, la abundancia o la mala distribución del agua pueden derivar en un riesgo de inundaciones que repercute en el medio urbano y en la actividad humana (Olcina, 2008). Por todo ello, el estudio de las precipitaciones en el contexto del cambio global es determinante para una planificación territorial coherente con la conservación del medio ambiente y adecuada para las actividades socioeconómicas.

Sin embargo, los modelos climáticos presentan importantes errores sistemáticos en la proyección de la precipitación. La baja resolución espacial de los modelos globales y regionales provoca que las proyecciones subestimen la importante variabilidad espacial dando valores suavizados; es decir, suelen subestimar la precipitación extrema y sobreestimar el número de días con precipitación.

Por lo tanto, antes de usar los productos de precipitación de los modelos, es necesario corregir el error sistemático usando datos de observatorios meteorológicos. No obstante, la precipitación es una variable muy irregular en su distribución espacial y temporal, por lo que dista mucho del comportamiento de otras variables físicas como la temperatura, la presión, etc. Por ejemplo, la distribución de probabilidad de la temperatura se aproxima a una distribución normal o *gaussiana*, con un valor medio más o menos centrado en una curva que aproximadamente es simétrica. Sin embargo, el valor más probable de la precipitación diaria es “cero”, por lo que el pico de probabilidad de su distribución se sitúa entorno a ese valor, provocando que la curva de probabilidad sea totalmente asimétrica: a priori no hay precipitaciones negativas, y la cola de la distribución baja suavemente para las precipitaciones positivas.

Desde el punto de vista físico, es posible tomar la evaporación como un tipo de precipitación negativa. Por ejemplo, en los procesos convectivos de una nube, puede haber simultáneamente precipitación y evaporación (virgas); de ese modo la precipitación que llega al suelo es el balance entre lo que precipita desde la nube y lo que se vuelve a evaporar antes de llegar al suelo. Si extendemos ese fenómeno para la precipitación en general, podemos prolongar matemáticamente el dominio de las funciones de probabilidad para intentar completar las curvas.

Sin embargo, a pesar de eso, las curvas de probabilidad siguen siendo asimétricas: los extremos de evaporación son bastante inferiores a los extremos de precipitación. Por todo ello, es necesario buscar nuevas distribuciones de probabilidad, que puedan modelizar y corregir adecuadamente la precipitación de las proyecciones climáticas.

Una de las posibles aplicaciones de la corrección de las proyecciones climáticas es el análisis del riesgo de inundaciones en España, especialmente en áreas del mediterráneo, pero también en áreas del litoral cantábrico como es el caso del País Vasco. En concreto, el riesgo de inundaciones por precipitación extrema unido al ascenso global del nivel del mar en la zona (Chust *et al.*, 2010) es de gran relevancia en la costa. En particular, las zonas estuáricas del País Vasco son especialmente vulnerables a dichas variaciones dado que en ellas se concentran tanto actividades humanas como comunidades bióticas específicas del intermareal.

I.2. Conceptos previos

I.2.1. Intensidad de la precipitación

a) Definición

La intensidad de la precipitación es la relación entre la acumulación de precipitación, P , y el tiempo con el que se acumula, t . De ese modo podemos distinguir entre intensidad media ($I = P/t$), y la intensidad instantánea ($I_i = dP/dt$), pero ¿cómo definimos si una precipitación es muy intensa? En la actualidad, los medios de comunicación emplean casi indistintamente el término de “lluvia intensa” y el de “lluvia torrencial”. Popularmente, este último concepto está relacionado con la capacidad de la lluvia para formar importantes torrentes de agua que suelen ocasionar problemas para las personas, bien directamente o bien a sus recursos naturales o artificiales. Sin embargo, la definición oficial de “lluvia torrencial”, así como de “lluvia intensa”, depende del país o del ámbito de cada organismo meteorológico o de ciencias afines (como la climatología, la hidrología, etc.). Así por ejemplo, la Agencia Estatal de Meteorología define una lluvia torrencial como aquella que supera los 60 mm en una hora (AEMET, 2011). Por otro lado, el Servei Meteorològic de Catalunya define la precipitación torrencial como aquella que supera los 40 mm en 30 minutos (SMC, 2011). El Centro de Estudios Medioambientales del Mediterráneo emplea el término torrencial para referirse a cantidades de lluvia muy cuantiosas en 12 ó 24h. Entonces, ¿pueden unificarse esas definiciones? ¿Qué relación hay entre 40 mm en 30 minutos y 60 mm en una hora?

El carácter torrencial o intenso de la precipitación debe ser un concepto diferenciado del de la duración, por lo tanto habría que definir “torrencial”, “intensa”, “muy intensa”, etc. de una forma independiente de si la lluvia dura un minuto, 60 minutos o 1 día. Sin embargo, es posible ver que en un día de lluvia abundante, ésta no se comporta de forma totalmente regular durante todo el día, sino que hay una hora en la que llueve más intensamente que en las demás. Ni si quiera dentro de esa hora

de máxima lluvia ésta presenta una intensidad constante durante todos los minutos. Por ejemplo, cabe esperar que si en una hora se han registrado 60 mm, entonces en los 30 minutos de mayor intensidad han caído más de 30 mm; ¿pero cuantos más? ¿unos 40 mm?

b) Caracterización

Por lo mencionado anteriormente, es necesario caracterizar adecuadamente los diferentes tipos de precipitación, según la relación existente entre la intensidad media máxima y cada tiempo considerado. En este sentido, en esta tesis proponemos emplear una sencilla relación:

$$I(t) \approx I(t_o) \left(\frac{t_o}{t} \right)^n \quad (I.1)$$

donde $I(t)$ y $I_o(t_o)$ son las intensidades medias máximas registradas en un tiempo t y t_o , respectivamente y n es un parámetro adimensional que caracteriza el tipo de precipitación (Moncho *et al.*, 2009a). Mediante dicha relación, es posible ver que una misma precipitación puede acumular 40 mm en 30 minutos y 60 mm en una hora; en ese caso, el índice n sería de aproximadamente 0,4. Nótese que el índice n sólo puede tomar valores entre 0 y 1. En el caso límite de que $n = 0$, la intensidad sería totalmente constante en el tiempo, por lo contrario, para $n = 1$, la precipitación sería totalmente instantánea, $I = P_o/t$ (Fig. I.1).

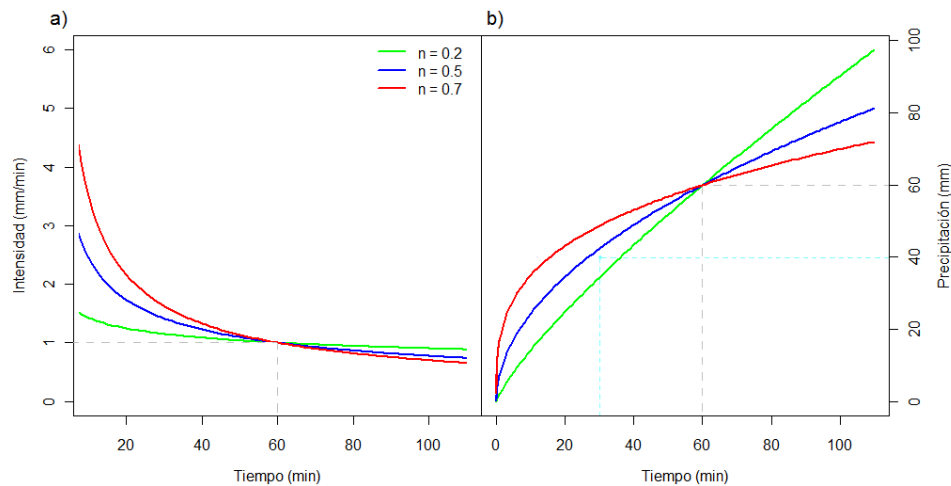


Figura I.1. Ejemplo de tres tipos de lluvia *torrencial*, según el índice n . Con un $n = 0,2$ la intensidad es prácticamente constante, mientras que para $n = 0,7$ y superiores la mayor parte de la precipitación se acumula en pocos minutos: a) Curvas de intensidad media máxima y b) curvas de acumulación máxima.

Por lo tanto, el concepto de intensidad y el carácter torrencial pueden definirse para la precipitación de forma independiente de la duración final. Sin embargo, es necesario introducir un concepto adicional para referirse al carácter excepcional de las precipitaciones más intensas. Se trata de la probabilidad de ocurrencia, o frecuencia de un evento. En este sentido, se habla de relaciones de Intensidad-Duración-Frecuencia (IDF) que llevan implícito tanto el carácter de la precipitación (n) como su frecuencia (Ben-Zvi, 2009).

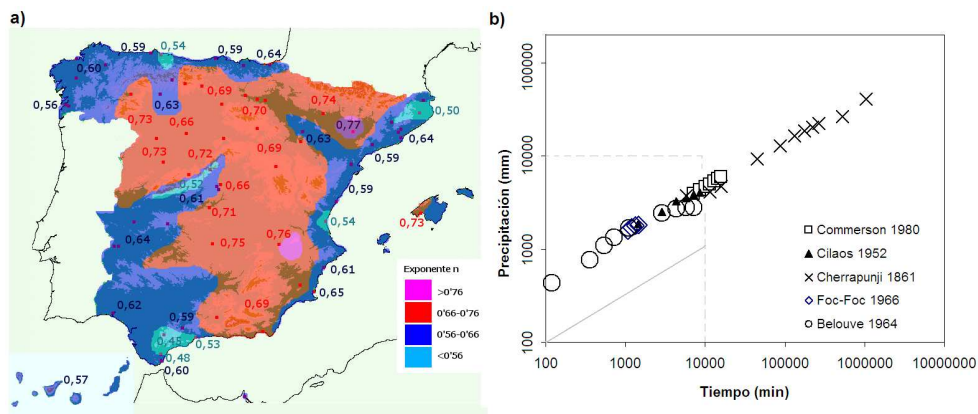


Figura I.2. a) Distribución espacial del índice n asociado a las curvas IDF de la Península Ibérica, con retornos entre 1 y 25 años (Moncho *et al.*, 2009a). b) Algunos de los registros mundiales más extraordinarios (Moncho *et al.*, 2011a): Foc-Foc (La Reunión), entre los días 7-8/01/1966 con $n = 0,49 \pm 0,02$; Commerson (La Reunión) entre los días 16-27/01/1980, con $n = 0,53 \pm 0,02$; Cilaos (La Reunión) durante los días 15-16/03/1952, con $n = 0,45 \pm 0,02$; Cherrapunji (China) en 1860-1861, con $n = 0,49 \pm 0,01$; y Belouve (La Reunión) entre los días 24-29/02/1964 con $n = 0,48 \pm 0,02$.

c) Relación entre intensidad, duración y frecuencia

El concepto de período de retorno p se refiere al intervalo de tiempo que cabe esperar entre dos eventos en los que se supera la cantidad de precipitación $P = I \cdot t$. Es decir, si definimos π como la probabilidad de que en un año se supere un registro de precipitación P , entonces el período de retorno es $p = 1/\pi$.

La probabilidad de ocurrencia de un evento de fuerte precipitación puede describirse según modelos matemáticos que se ajustan a los registros climáticos de diferentes observatorios. Esta información probabilística puede estar implícita en otras relaciones semi-empíricas como las conocidas curvas de Intensidad-Duración-Frecuencia (IDF). Según Témez (1978), estas curvas son el resultado de conectar los puntos representativos de la intensidad de la precipitación media en intervalos de tiempo diferentes, todos correspondientes a la misma frecuencia o período de retorno.

Existe un gran número de aproximaciones semi-empíricas que relacionan la intensidad (I), la duración (t) y el período de retorno (p), a partir de ajustes a potencias tales como:

- La fórmula de Chow (1962) parte de tres parámetros (a , c y n) ajustados para cada período de retorno fijado:

$$I(t) \approx \frac{a}{t^n + c} \quad (1.2)$$

- La función potencial (Aparicio, 1997), basada en la fórmula de Sherman (1931), utiliza cuatro parámetros (k , c , m y n):

$$I(t) \approx k \cdot \frac{P^m}{(t + c)^n} \quad (1.3)$$

Nótese que estas funciones llevan implícita la relación entre la duración y la intensidad media mediante un exponente n . En cuanto a la distribución de probabilidad, en el último caso se puede relacionar con la función potencial de Pareto (Barry, 1983):

$$\pi = \left(\frac{\lambda}{P} \right)^\alpha = \dots = \left(\frac{k \cdot t}{I \cdot t} \right)^{\frac{1}{m}} \xrightarrow{\pi \equiv \frac{1}{p}} I(p) = k \cdot p^m \quad (1.4)$$

donde λ , α , k y m son parámetros ajustables.

1.2.2. Distribuciones de probabilidad clásicas

a) Distribución normal y semi-normal

La probabilidad de ocurrencia, f , de una variable continua, x , puede asemejarse a diferentes funciones matemáticas (Abramowitz y Stegun, 1972). La más común es la distribución normal o *gaussiana*, que en su forma de densidad de probabilidad se escribe como:

$$f = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right) \quad (1.5)$$

donde μ y σ son parámetros de ajuste, conocidos respectivamente como media y desviación estándar. Recordemos que podemos definir la variable normalizada de x , según:

$$\lambda(x) \equiv \frac{x - \mu}{\sigma} \quad (1.6)$$

La probabilidad acumulada $\pi(x \geq x_0)$ de una curva *gaussiana* puede escribirse mediante la función de error (ERF), según:

$$\pi(x \leq x_0) = \int_{-\infty}^{x_0} f dx = \frac{1}{2} \left[1 + \text{ERF}\left(\frac{\lambda(x)}{\sqrt{2}}\right) \right] \quad (1.7)$$

Para variables como la precipitación, que no presentan valores negativos, puede tomarse la versión semi-normal, la cual puede escribirse como:

$$\pi(x \leq x_o) = ERF\left(\frac{\lambda(x)}{\sqrt{2}}\right) \quad (1.8)$$

b) Distribución log-normal y semi log-normal

Otra de las funciones más comúnmente empleadas es la log-normal, que es muy útil para variables con colas de distribución muy elevadas, es decir, con una probabilidad que decrece lentamente a medida que aumenta el valor de la variable. Por lo tanto, la función log-normal es a priori una candidata a describir la probabilidad de la precipitación:

$$\pi(x \leq x_o) = ERF\left(\frac{\ln \lambda(x)}{\kappa\sqrt{2}}\right) \quad (1.9)$$

donde κ es un parámetro ajustable. Nótese que en este caso el dominio de la variable x viene dado por la inecuación $\lambda \geq 1$. En la Figura 1.3 se comparan las diferentes distribuciones normales.

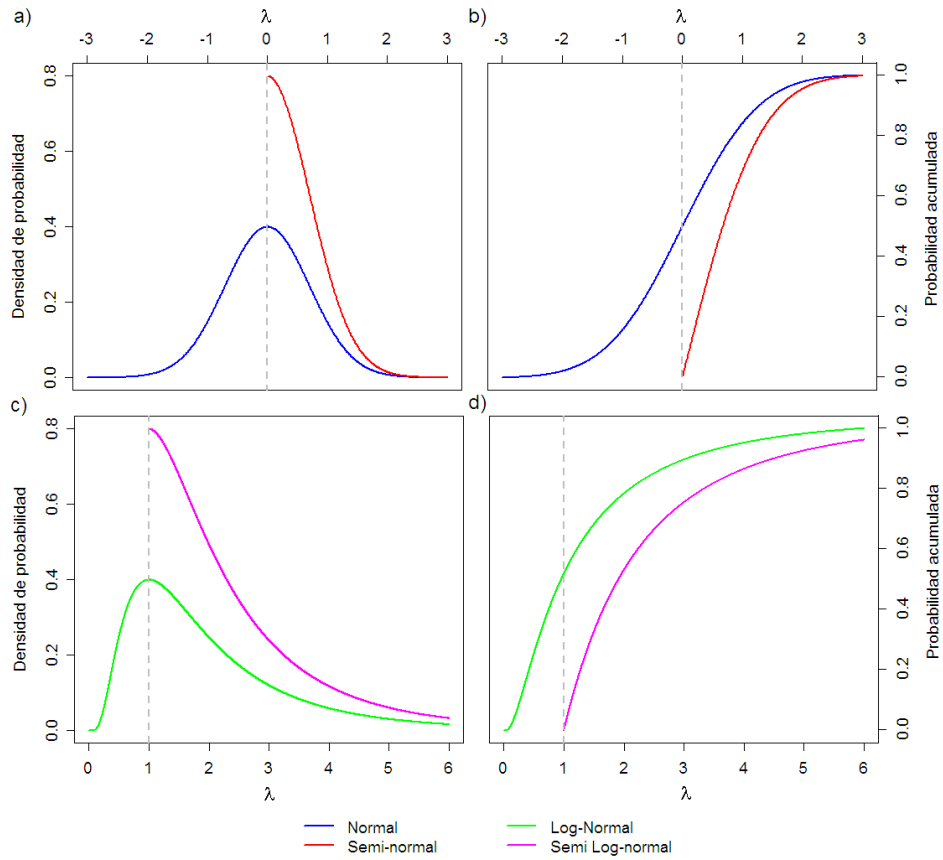


Figura I.3. Comparación de las diferentes distribuciones normales: a) densidad de probabilidad de la distribución normal y semi-normal; b) probabilidad acumulada de la distribución normal y seminormal; c) densidad de probabilidad de la distribución log-normal y semi log-normal; d) probabilidad acumulada de la distribución log-normal y semi log-normal.

c) Distribuciones gamma

Para estudiar la probabilidad de la precipitación, muchos autores emplean distribuciones gamma (Elshamy *et al.*, 2009; Piani *et al.*, 2010), que se construyen mediante la función gamma III (Hogg y Craig, 1978):

$$\pi(x \leq x_o) = \frac{\gamma(k, \lambda(x))}{\Gamma(k)} \quad (I.10)$$

donde γ es la función gamma incompleta, Γ es la gamma completa, y k es un parámetro ajustable (Fig. I.4).

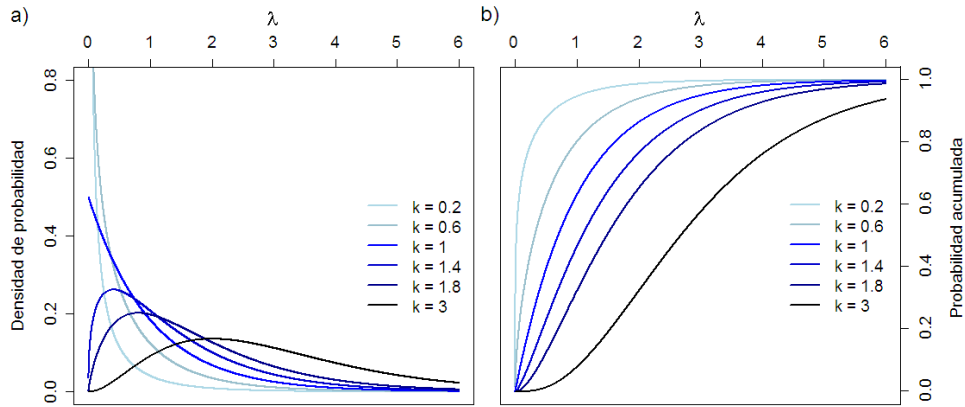


Figura I.4. a) Densidad de probabilidad de la distribución gamma III; b) probabilidad acumulada de la distribución gamma III.

d) Distribuciones exponenciales

Una de las distribuciones de probabilidad más empleadas para la velocidad del viento, x , es la distribución exponencial de Weibull (Johnson *et al.*, 1994):

$$\pi(x \leq x_o) = 1 - \exp(-\lambda^w) \quad (I.11)$$

donde w es un parámetro de ajuste y $\lambda = (x - \mu)/\sigma$, donde μ y σ son también parámetros de ajuste. Nótese que para $w = 1$ se recupera la distribución exponencial clásica. En muchos casos se puede aplicar el modelo de Weibull para ajustar la probabilidad de la precipitación e incluso de caudales de río. Sin embargo, muchos autores prefieren la distribución de Gumbel, que es doblemente exponencial:

$$\pi(x \leq x_o) = \exp[-\exp(-\lambda)] \quad (I.12)$$

o bien la distribución de Gumbel opuesta:

$$\pi(x \leq x_o) = 1 - \exp[-\exp(\lambda)] \quad (I.13)$$

Tanto la distribución de Gumbel como la distribución de Weibull están relacionadas con la distribución Generalizada del Valor Extremo (GEV):

$$\pi(x \leq x_o) = \exp\left\{-\left[1 + \varepsilon \cdot \lambda\right]^{-\frac{1}{\varepsilon}}\right\} \quad (I.14)$$

donde ε es un parámetro ajustable. Nótese que en el límite de ε tendiendo a cero se recupera la distribución de Gumbel. Si se toma valores negativos para x y ε , entonces se recupera una función tipo Weibul. Sin embargo, la distribución GEV opuesta presenta mejores límites:

$$\pi(x \leq x_o) = 1 - \exp\left\{-\left[1 + \varepsilon \lambda\right]^{\frac{1}{\varepsilon}}\right\} \quad (I.15)$$

Obsérvese que la ecuación I.15 pertenece a la familia de funciones de Weibull, ya que podemos definir $w \equiv 1/\varepsilon$, y tomar λ con una forma apropiada:

$$\lambda^* = \frac{x - \mu^*}{\sigma^*} = \frac{x - (\mu - \sigma w)}{\sigma w} = 1 + \frac{\lambda}{w} = 1 + \varepsilon \lambda \quad (I.16)$$

Por otro lado, la función GEV opuesta también se transforma en la distribución de Gumbel opuesta para valores de ε que tienden a cero (Fig. I.5).

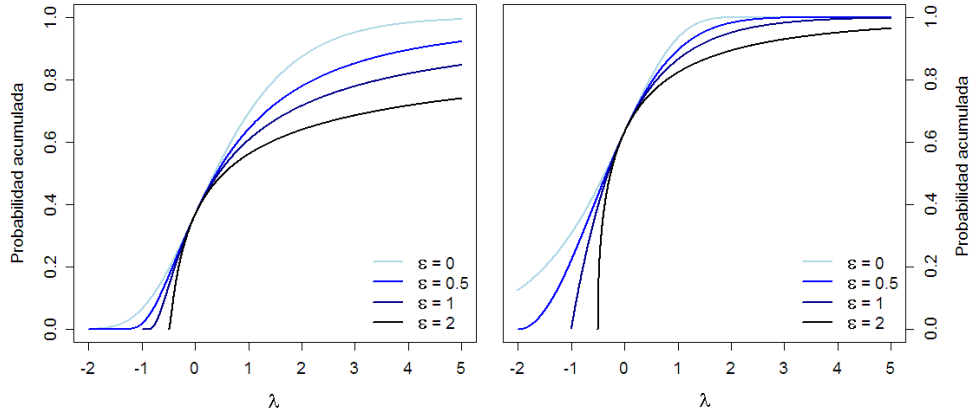


Figura I.5. a) Probabilidad acumulada de la distribución GEV; para $\varepsilon = 0$ se recupera la distribución de Gumbel; b) Probabilidad acumulada de la distribución GEV opuesta; para $\varepsilon = 0$ se recupera la distribución de Gumbel opuesta.

d) Distribuciones potenciales

La función potencial más sencilla que forma una distribución de probabilidad se conoce por el sociólogo Vilfredo Pareto (Barry, 1983), y tiene la siguiente forma:

$$\pi(x \leq x_o) = 1 - \left(\frac{\sigma}{x} \right)^w \quad (I.17)$$

donde σ es un parámetro ajustable. Nótese que es posible definir de nuevo una variable como $\lambda \equiv (x - \mu)/\sigma$, por lo cual la distribución de Pareto queda como:

$$\pi(x \leq x_o) = 1 - \frac{1}{\lambda^w} \quad (I.18)$$

Habitualmente se define la Distribución Generalizada de Pareto (GPD) como:

$$\pi(x \leq x_o) = 1 - \frac{1}{(1 + \varepsilon \lambda)^{\frac{1}{\varepsilon}}} \quad (I.19)$$

Del mismo modo que en la ecuación I.16, en la ecuación I.18 se tiene una distribución de Pareto pero con la peculiaridad que si escogemos el límite cuando ε tiende a cero entonces se recupera la distribución exponencial o Weibull con $w = 1$ (Fig. I.6).

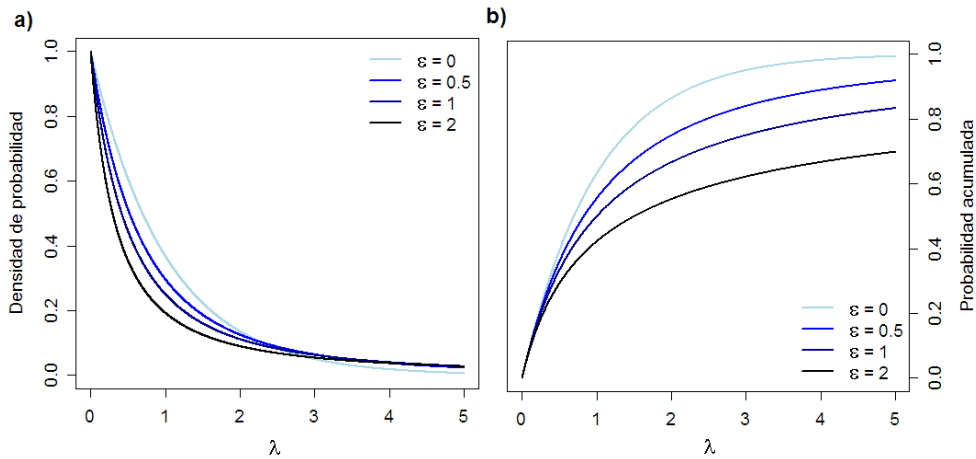


Figura I.6. Ejemplos de la Distribución Generalizada de Pareto, donde para $\varepsilon = 0$ se recupera la distribución exponencial: a) densidad de probabilidad (derivada de la ecuación I.19); b) probabilidad acumulada (ec. I.19).

e) Distribuciones logísticas

Una función similar a la ecuación I.19 es la familia de funciones logísticas:

$$\pi(x \leq x_o) = 1 - \frac{1}{1 + (1 + \varepsilon \lambda)^{\frac{1}{\varepsilon}}} \quad (I.20)$$

Si ε tiende a cero entonces se obtiene la distribución logística:

$$\pi(x \leq x_o) = 1 - \frac{1}{1 + \exp(\lambda)} \quad (I.21)$$

Por otro lado, teniendo en cuenta de nuevo la ecuación I.16, entonces la ecuación I.20 se convierte en la distribución log-logística (Fig. I.7):

$$\pi(x \leq x_o) = 1 - \frac{1}{1 + \lambda^w} \quad (I.22)$$

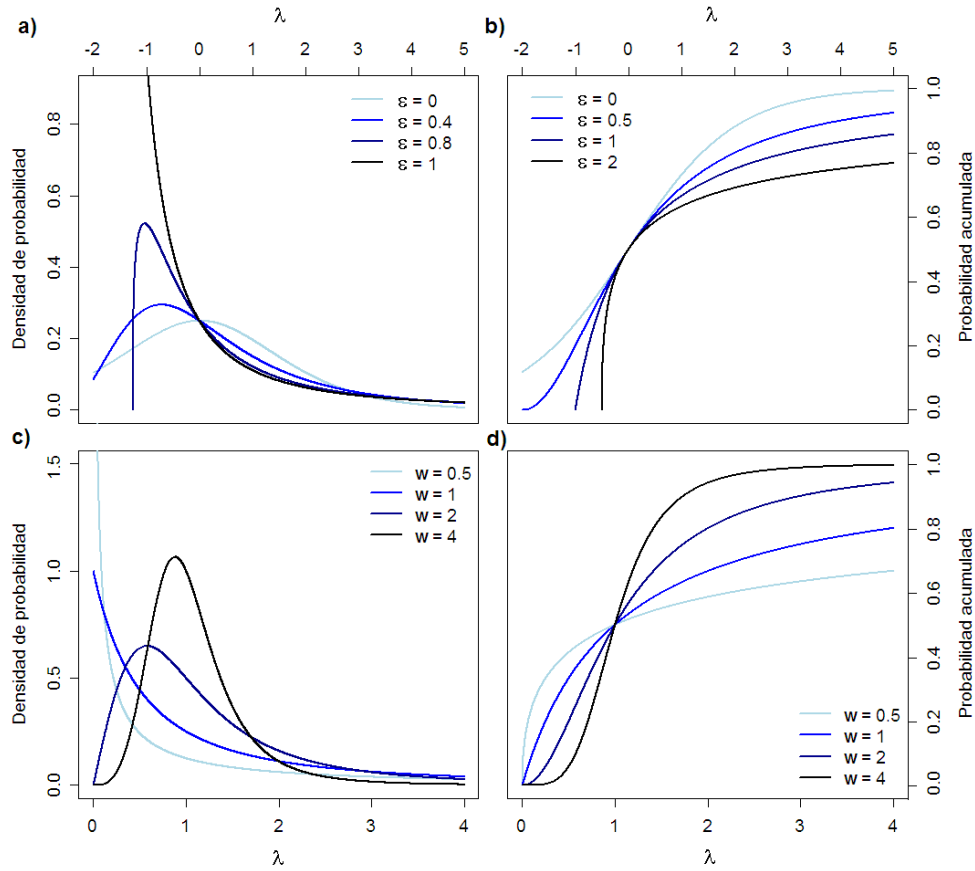


Figura I.7 a) Densidad de probabilidad de la familia de curvas logísticas (derivada de la ecuación I.20). b) Probabilidad acumulada de las curvas logísticas (ec. I.20), donde para $\varepsilon = 0$ se tiene la distribución Logística (ec. I.21). c) Densidad de probabilidad de la distribución Log-Logística (derivada de la ecuación I.22). d) Probabilidad acumulada de la distribución Log-Logística (ec. I.22).

I.2.3. Corrección del error sistemático de la precipitación

a) Corrección lineal

Las salidas numéricas de los modelos climáticos globales y regionales presentan importantes errores sistemáticos que a menudo deben corregirse para distintas aplicaciones prácticas (Fig. I.8). Por ejemplo, la precipitación anual y mensual puede corregirse linealmente mediante la diferencia de la media (sesgo) y el cociente de las desviaciones típicas (Terink *et al.*, 2010):

$$P_{jC} \equiv (P_j - \bar{P}) \frac{\sigma_{obs}}{\sigma} + \bar{P}_{obs} \quad (I.8)$$

donde, P_{jc} es la precipitación del año-j de la serie calibrada, P_j es la precipitación del año-j de la serie sin calibrar, mientras que \bar{P} y σ son la media y desviación estándar, respectivamente, para la serie sin calibrar. Por otro lado, \bar{P}_{obs} y σ_{obs} son la media y desviación típica de referencia.

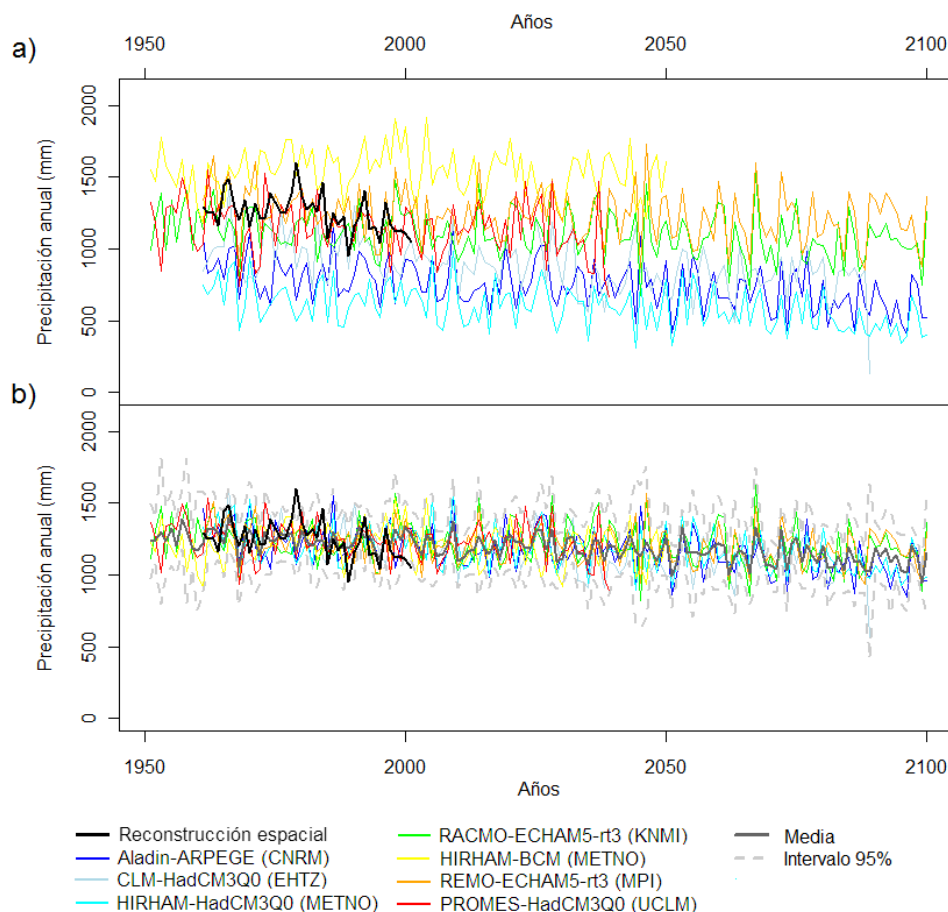


Figura I.8 Ejemplo de corrección lineal de la precipitación anual de 7 modelos regionales del clima (RMC), a partir de una serie de referencia del País Vasco obtenida mediante reconstrucción espacial (Moncho *et al.*, 2009b): a) Series temporales antes de corregir y b) después de corregir (Moncho, 2009).

b) Corrección geométrica

Puesto que la corrección lineal no puede aplicarse a escala diaria, necesariamente deben usarse correcciones no-lineales. La más sencilla de todas es la corrección geométrica. Este método consiste en aplicar un factor y un exponente que vienen dados, respectivamente, por el cociente entre la media de la serie-problema y de la serie-referencia, por un lado, y entre la desviación típica de la serie-problema y de la serie-referencia, por

otro. Estos factores correctores se aplican iterando para un gran número de pasos, en los que se espera que la serie resultante converja en otra con media y desviación típica muy cercanas a las de referencia. Por ejemplo, la precipitación del día- i correspondiente a la iteración $n+1$ puede estimarse según:

$$(P_i)_{n+1} = \left(\frac{(\bar{P}_i)_{obs}}{(\bar{P}_i)_n} \right)^\alpha (P_i)_n \left(\frac{(\sigma_i)_{obs}}{(\sigma_i)_n} \right)^\alpha \quad (I.9)$$

donde α es un parámetro que toma valores entre 0 y 0,5; $(\bar{P}_i)_{obs}$ es la precipitación media climática de referencia, correspondiente al día- i (o bien, puede tomarse la media mensual en mm/día), mientras que $(\sigma_i)_{obs}$ es la desviación típica de referencia climática estimada a escala diaria correspondiente al día- i , o a su mes (en mm/día). Por otro lado, $(P_i)_n$ es la precipitación del día- i en la iteración- n del proceso de corrección, mientras que $(\bar{P}_i)_n$ es la precipitación media climática resultante de dicha iteración para el día problema y $(\sigma_i)_n$ es la desviación típica (climática) a escala diaria asociada al mismo día. Alternativamente, puede tomarse los valores de referencia a escala mensual en vez de la escala diaria.

La corrección geométrica no hace variar el número de días con precipitación cero, por lo que si fuese necesario corregir dicho valor habría que tratar previamente la serie-problema. En el caso de que haya menos días con ceros en la serie-problema, entonces bastaría con extraer una cantidad pequeña a todos los valores hasta que el número de valores mayores que cero se igualase a la serie-referencia; de ese modo se redefinirían como valores nulos todos aquellos que fuesen menores que cero. Por lo contrario, si la serie-problema tiene más valores nulos que la serie-referencia, entonces habría que suavizar la serie: si el valor del día- i es cero, entonces se cambia por la media del valor anterior y posterior. Este proceso se realizaría hasta conseguir suficientes días con valores mayores que cero.

La corrección geométrica y similares pueden emplearse para corregir series muy próximas a las observadas (Moncho, 2009), pero en muchos casos pueden alterar inadecuadamente la distribución de probabilidad de la serie-problema y provocar que la serie resultante no se parezca a la serie-referencia en aspectos tan importantes como los valores extremos, e incluso en otros rangos de la distribución de probabilidad (Fig. I.9).

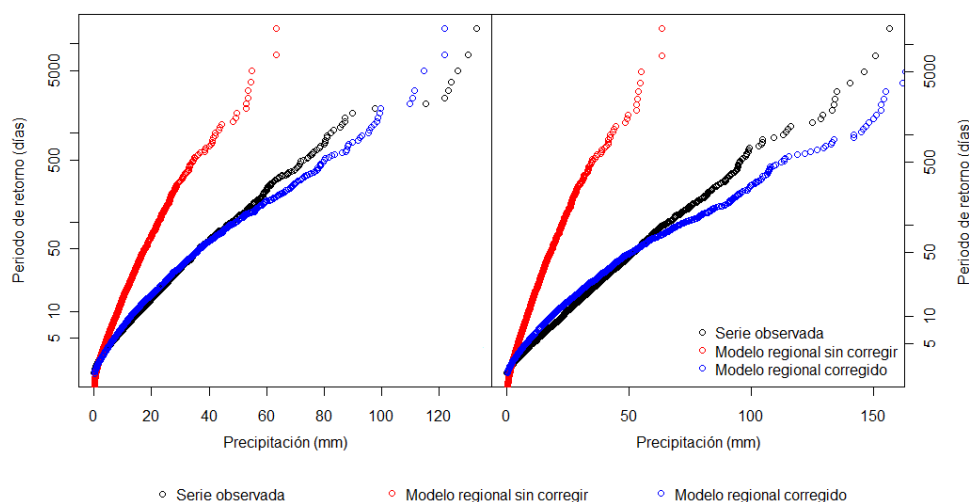


Figura I.9. Ejemplo de transformación geométrica de la precipitación diaria con el modelo CLM-HadCM3Q0-A1B (ETHZ): a la izquierda observatorio de Fuenterrabia (Guipúzcoa), y a la derecha observatorio de Rentería (Guipúzcoa). Fuente: Moncho (2009).

c) Corrección probabilística

Una alternativa a la corrección geométrica es la corrección probabilística o por quantiles (Ines y Hansen, 2006; Benestad, 2010). Si disponemos en dos gráficas diferentes las distribuciones de probabilidad empírica de la serie-problema y la serie-referencia, entonces puede aplicarse una relación de correspondencia entre los valores simulados de la precipitación, asociada a una probabilidad de ocurrencia, y los valores de referencia para la misma probabilidad de ocurrencia.

No obstante, las curvas de las distribuciones de probabilidad empírica (ECDF) pueden presentar irregularidades en las derivadas. Es decir, la probabilidad de ocurrencia de tres valores de precipitación muy cercanos entre sí debería ser ordenada (un valor lejos de la moda no debería tener más frecuencia que sus vecinos inmediatos). Las irregularidades en la continuidad de la curva ECDF pueden deberse a variabilidad natural de corta ocurrencia o bien a errores sistemáticos del instrumento de medición. Si bien, el efecto de la variabilidad de corta ocurrencia se suaviza para series muy largas.

Otro problema de la corrección por cuantiles es que la estimación de los extremos de precipitación presenta un gran intervalo de error. Además, éstos no pueden ser extrapolados a períodos de tiempo más allá de la propia longitud de la serie de referencia (Piani *et al.*, 2010). En este sentido, parece que la corrección por cuantiles no puede ser empleado para corregir la precipitación intensa de los modelos climáticos ya que en general las series de los observatorios son mucho más cortas.

Todos esos problemas pueden solucionarse en parte si se aplica una corrección mediante funciones matemáticas que ajusten de forma muy robusta toda la distribución de probabilidad de la precipitación, de ese modo, la extrapolación de los extremos de precipitación presentaría un error menor, ya que se eliminarían las irregularidades de las ECDF. En esa línea de trabajo, algunos autores han empleado la distribución gamma (Elshamy *et al.*, 2009; Piani *et al.*, 2010); sin embargo, existen otras posibilidades como veremos en los siguientes capítulos.

I.3. Objetivos y novedades

El objetivo principal de esta tesis es probar el adecuado funcionamiento de dos nuevos modelos de distribución de probabilidad para la precipitación diaria y subdiaria. El trabajo se divide en dos aplicaciones principales: en la primera se estudia la bondad de ajuste de los modelos de probabilidad para los diferentes climas de España; en la segunda parte se utilizan dichos modelos para corregir la distribución de probabilidad de diferentes modelos regionales del clima para el País Vasco, en base a un grupo de observatorios de referencia. En este último caso, se pretende construir una herramienta que sirva para el análisis del riesgo de inundación potencial y por lo tanto para cuantificar los posibles impactos derivados de ello.

Para alcanzar los objetivos citados, en este trabajo se presentan algunas novedades destacables en el ámbito de la meteorología y climatología:

a) Novedad principal:

- Desarrollo de dos nuevos modelos de distribución de probabilidad que tratan de ajustar adecuadamente todo el espectro de probabilidades de la precipitación diaria: Versiones ampliadas de la distribución de Gumbel opuesta y de la distribución Log-Logística.

b) Para la corrección de los modelos climáticos regionales:

- Avances en la metodología de transformación probabilística mediante el uso de Series de Período de Retorno (RPS) como forma de estandarización y como base para ampliar series temporales.
- Uso de una nueva reconstrucción espacial de la precipitación como apoyo para completar huecos de observatorios y para validar los modelos regionales del clima.

c) Para el análisis de la precipitación intensa en el País Vasco:

- Nueva síntesis sobre el tratamiento matemático de la intensidad de la precipitación, que permite bajar fácilmente a escala subdiaria. Esto

facilita el análisis del riesgo de inundaciones, mediante la construcción de hietogramas sintéticos.

CAPÍTULO II. NUEVOS MODELOS DE PROBABILIDAD

II.1. Nueva versión de la distribución de Gumbel Opuesta

II.1.1. Probabilidad acumulada

En este trabajo proponemos la siguiente versión de cuatro parámetros de la distribución de Gumbel opuesta:

$$\pi(\lambda_0 \leq \lambda) = 1 - \exp[-\exp(\lambda^w + k)] \quad (II.1)$$

donde w y k son, respectivamente, los parámetros de “forma” y “desplazamiento”. La variable de la curva, λ , es la precipitación relativa, que a su vez depende de un parámetro de localización, P_o , que representa la precipitación más probable y un parámetro de escala, P_1 , a modo de normalización.

$$\lambda \equiv \frac{P - P_o}{P_1} \quad (II.2)$$

Por lo tanto, la probabilidad acumulada de ocurrencia de una precipitación p igual o inferior a P satisface la misma distribución de probabilidad (ecuación II.2), i. e., $\pi(p \leq P) = \pi(\lambda_o \leq \lambda)$. En cuanto al dominio de la variable λ , cabe destacar que es real y positiva.

II.1.2. Densidad de probabilidad

La densidad de probabilidad asociada a la ecuación II.1 se calcula trivialmente mediante su derivada, y queda como:

$$f(\lambda_o \geq \lambda) = w \lambda^{w-1} \exp(\lambda^w + k) \exp[-\exp(\lambda^w + k)] \quad (II.3)$$

Donde k puede tomar cualquier valor real, mientras que w se define como positivo. Nótese que para $w < 1$, la densidad de probabilidad presenta un polo en $\lambda = 0$.

II.1.3. Función inversa

La distribución de probabilidad acumulada puede invertirse fácilmente según:

$$\frac{P - P_o}{P_1} = [\ln(\ln T) - k]^{\frac{1}{w}} \quad (II.4)$$

donde T es el período de retorno definido como la inversa de la probabilidad acumulada $\pi(\lambda_o \geq \lambda)$, donde a su vez $\pi(\lambda_o \geq \lambda) = 1 - \pi(\lambda_o \leq \lambda)$.

II.2. Nueva versión de la distribución Log-Logística Generalizada

II.2.1. Probabilidad acumulada

Alternativamente al modelo del apartado anterior, proponemos otra función de probabilidad acumulada, $\pi(\lambda)$, con similitudes respecto a la distribución Log-logística:

$$\pi(\lambda_0 \leq \lambda) = 1 - \frac{1}{1 + \lambda^{w+\lambda^w e^{-k}}} \quad (II.5)$$

donde w y k son parámetros ajustables positivos, y λ es la precipitación relativa (ecuación II.2), definida también como positiva. Nos referiremos a w como el parámetro de “forma” y a k como el parámetro de “suavidad”.

II.2.2. Función de densidad

De la ecuación II.5, se deduce que la función de densidad de probabilidad asociada $f(\lambda; w, k)$ viene dada por:

$$f(\lambda; w, k) \equiv \frac{d}{d\lambda} \pi(\lambda; w, k) = \frac{\lambda^{w+\lambda^w e^{-k}-1}}{(1 + \lambda^{w+\lambda^w e^{-k}})^2} \left[w + (1 + w \ln \lambda) \lambda^w e^{-k} \right] \quad (II.6)$$

En la figura II.1 se aprecia que el parámetro de “suavidad”, k , solo afecta a los valores de probabilidad medios y altos, mientras que el parámetro de “forma”, w , modula todo el intervalo de probabilidades.

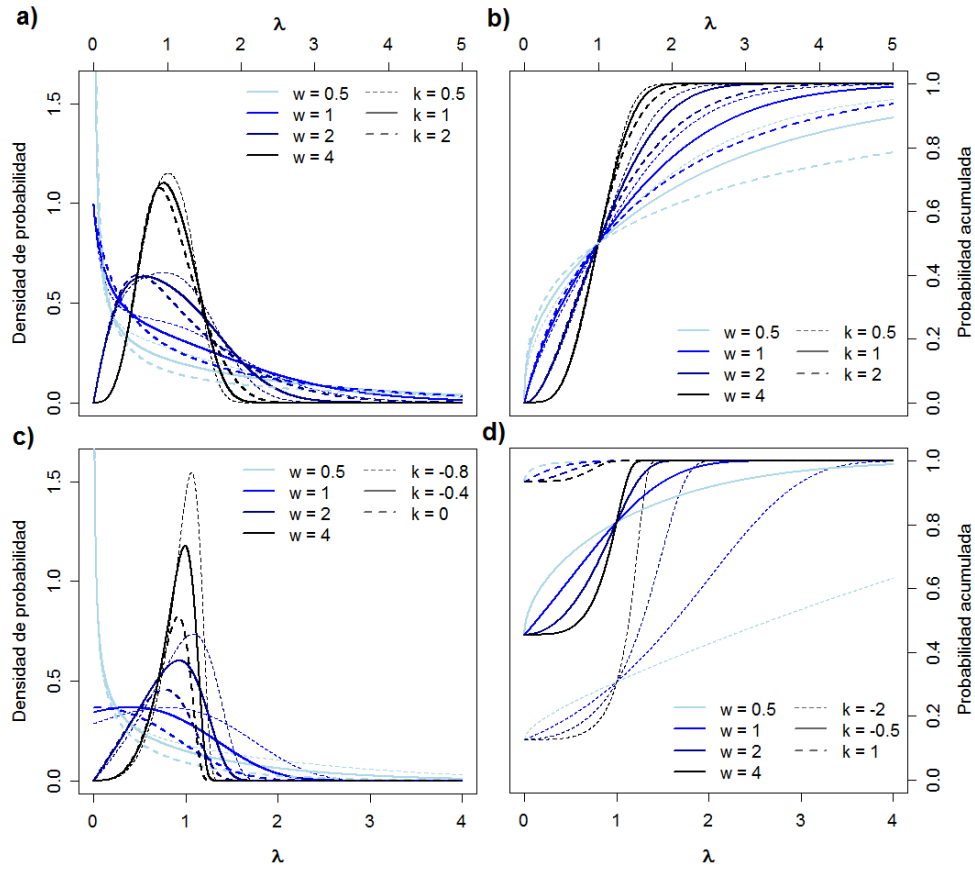


Figura II.1. Comparación entre la nueva versión de la distribución Log-Logística (a y b) y la de Gumbel opuesta (c y d). A la izquierda (a y c), curvas de densidad de probabilidad; a la derecha (b y d), curvas de probabilidad acumulada.

II.2.3. Función inversa

La ecuación II.5 no tiene inversa explícita, por lo tanto es necesario usar algoritmos de programación o bien buscar una función inversa aproximada, de la forma:

$$\frac{P - P_o}{P_1} = (T - 1)^{m(T-1)} \quad (II.7)$$

donde $m(T-1)$ es una función ajustable del período de retorno. Para períodos de retorno elevados, la ecuación anterior puede aproximarse a una distribución de Pareto (Hardy, 2010), de acuerdo con:

$$\frac{P}{P_1} \approx T^{m_o} \quad (II.8)$$

donde P_o y m_o son parámetros ajustables. Para un evento de lluvia con una

duración diferente de un día, podemos considerar la duración como el tiempo esperado para una determinada acumulación de lluvia. Por lo tanto, es posible aproximar la distribución temporal de la máxima acumulación a otra distribución potencial:

$$\frac{P}{P_1} \approx t^{h_o} \quad (II.9)$$

donde t es la duración, mientras que P_1 y h_o son parámetros ajustables. Uniendo las ecuaciones II.8 y II.9, se obtiene una curva de Intensidad-Duración-Frecuencia (IDF; Ben-Zvi, 2009):

$$P(T, t) \approx P_o(T_o, t_o) \left(\frac{T}{T_o} \right)^{m_o} \left(\frac{t}{t_o} \right)^{h_o} \quad (II.10)$$

donde $P(T, t)$ es la acumulación de lluvia esperada para un período de retorno T y una duración t , $P_o(T_o, t_o)$ es la lluvia esperada para un retorno T_o y una duración t_o , y por último, m_o y h_o son parámetros ajustables.

II.3. Distribución temporal

Tanto para las nuevas versiones de la distribución de Gumbel opuesta y Log-logística como para otras posibles distribuciones, podemos describir una distribución temporal en base al índice n . En concreto, a partir de las ecuaciones II.10 y I.1, podemos generalizar una ecuación para la acumulación máxima en función del período de retorno T y la duración t :

$$\frac{P}{P_o} \approx f(T, T_o) \cdot \left(\frac{t}{t_o} \right)^{1-n} \quad (II.11)$$

donde P y P_o es la precipitación esperada para una duración de t y t_o , y para un período de retorno T y T_o , respectivamente, mientras que n es el índice de la precipitación. Nótese que $f(T, T_o)$ es una función del período de retorno que depende del tipo de distribución considerada. Así por ejemplo, para la distribución generalizada de Gumbel opuesta (ec. II.1), esa función viene dada por la ec. II.4:

$$f(T, T_o) \approx \left[\frac{\ln(\ln T) - k}{\ln(\ln T_o) - k} \right]^{\frac{1}{w}} \quad (II.12)$$

donde k y w son dos parámetros ajustables.

**CAPÍTULO III. PRIMER TEST: AJUSTE A LOS
CLIMAS DE ESPAÑA**

III.1. Introducción

En la mayoría de los observatorios meteorológicos, la precipitación más probable en un día es cero, por lo que la curva de probabilidad climática disminuye progresivamente desde la lluvia más débil hasta la más intensa. La precipitación no se comporta como una distribución normal, en parte porque no tiene valores negativos. Por otro lado, aunque se intente completar el dominio de la función de probabilidad con la evaporación como un modo de precipitación negativa, la distribución seguiría siendo muy asimétrica. Esto se debe a que la evaporación rara vez llega a los 20 mm en un día, en la mayoría de los observatorios (McGuinness y Leslie, 1972; Granger, 1989; Jeevananda, 1995), mientras que precipitaciones superiores a 50 mm en un día se registran con mayor frecuencia en algunos países (Cerveny *et al.*, 2007). En particular, las precipitaciones extremas son relativamente frecuentes en climas como el mediterráneo, ya que en ocasiones en un día puede llover más de la mitad de la precipitación anual (Martín-Vide, 2004). Esto causa dificultades para los ajustes matemáticos de la probabilidad de la precipitación diaria.

Posiblemente por esa razón, las distribuciones usadas habitualmente (Generalización de Valores Extremos, Distribución Generalizada de Pareto, Gamma, Gumbel, Weibull) no se ajustan a todo el intervalo de precipitación diaria, desde la más baja (incluyendo el cero) hasta la más alta (Bridges y Haan, 1972; Öztürk, 1981; Etoh *et al.*, 1986; Begueria, 2005). Todos estos modelos usan dos o tres parámetros, por lo que parece necesario usar otras distribuciones de probabilidad, con más parámetros (Park y Jung, 2002; Hanson y Vogel, 2008).

El objetivo de este capítulo es estudiar la bondad de ajuste de los dos nuevos modelos de distribución de probabilidad de la precipitación (incluyendo los días sin lluvia o inapreciable), que han sido descritos en el apartado II. El estudio también se centra en comprobar el funcionamiento de los dos modelos para los principales climas de España: mediterráneo, oceánico, subtropical y de montaña.

III.2. Material

III.2.1. Área de estudio

El área de estudio escogida para este capítulo fue España, incluyendo los dos archipiélagos (Fig. III.1), como área piloto para el primer análisis de la metodología desarrollada. El clima predominante en España es el mediterráneo, si bien también podemos encontrar el clima oceánico (en el norte de la Península Ibérica), clima de montaña (en Sierra Nevada y los Pirineos), y el clima subtropical seco (en las Islas Canarias). La diversidad climática de España se debe en gran parte a su ubicación geográfica y a las características del territorio, con una orografía que a

menudo está fuertemente accidentada. Respecto a la ubicación geográfica, cabe destacar que la Península Ibérica se encuentra dentro de la influencia de las ondas de chorro polar, o las ondas de Rossby, y en ocasiones se ve afectada también por el chorro subtropical (Alves y Verdière 1999, Peliz *et al.* 2002). Dichas masas de aire se humidifican en contacto con el Atlántico o el Mediterráneo, contrastando con otras masas de aire más secas, cálidas o frías, procedentes respectivamente del Sáhara o de Siberia.

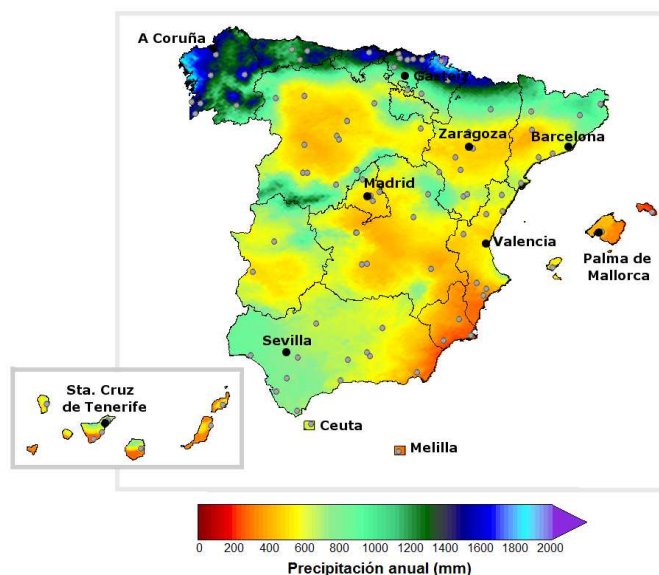


FIG. III.1. Localización de las estaciones meteorológicas empleadas en este estudio y precipitación anual para el período 1950-2000 interpolada con el método Thin Plate Spline multidimensional (latitud, longitud, altitud, pendiente y orientación geográfica).

Por lo tanto, las condiciones climáticas de la precipitación son muy heterogéneas: la acumulación anual de precipitaciones oscila entre los 150 mm en el sureste de Andalucía, con clima árido, y los 2.500 mm en el noreste del País Vasco, con clima oceánico (Capel-Molina, 2000). Este último clima se caracteriza por precipitaciones regulares durante todo el año (en algunos casos con más del 50% de días con lluvia). En cambio, la precipitación del clima mediterráneo se caracteriza por un marcado mínimo estival y lluvias intensas concentradas en unos pocos días (Martín-Vide, 2004; Lana *et al.*, 2004), sobre todo en otoño y primavera. A lo largo del año, el número de días con precipitación por lo general oscila entre el 10 y el 30% de media, aunque puede alcanzar valores por debajo del 10% en julio y en algunos casos más del 50% en abril (Capel-Molina, 2000). Por lo tanto, las diferentes estaciones del año, en cierta medida representan los diferentes climas de España (por ejemplo, el verano del clima mediterráneo es similar a los climas más secos y la primavera a los climas más

húmedos). Por estas razones, España es un área piloto adecuada para probar la metodología desarrollada en este trabajo.

III.2.2. Datos de precipitación

Los datos de precipitación utilizados para este estudio se corresponden con 108 observatorios con una resolución diaria, proporcionados por la Agencia Estatal de Meteorología (AEMET) de España (Fig. 1). Las series temporales seleccionadas satisfacen una longitud de al menos 20 años y el 90% de días con registros. Entre esos observatorios, 62 tienen una longitud de más de 50 años y el 98% de días con registros. Los datos presentan un gran número de días con precipitación inapreciable (es decir, entre 0 y 0,1 mm), en concreto se encuentran entre el 0,6 y el 16% de días con lluvia inapreciable. Estos valores se han cambiado a 0, por lo que el estudio tuvo en cuenta que los ceros representan realmente la clase de $[0, 0,1)$ mm. Además, en las series analizadas se encontraron algunos valores diarios erróneos (acumulado durante varios días), siendo menos del 0,1% sobre el total; sin embargo, no representan valores extremos, por lo tanto no afectan a las estadísticas. En estos casos, hemos dividido la cantidad de precipitación diaria acumulada entre el número de días correspondiente.

III.3. Metodología

III.3.1. Modelos de probabilidad

Para el estudio de la frecuencia de la precipitación diaria en un determinado grupo de estaciones o puntos de malla de un modelo, se puede tomar la precipitación media o la máxima del grupo. Para este estudio se analiza la probabilidad de cada observatorio por separado. La probabilidad acumulada $\pi(p \geq P)$ se define como la probabilidad de que una estación o grupo de estaciones registre una precipitación igual o mayor que P . Lógicamente, la probabilidad de que se supere cada umbral de precipitación, P , se puede estimar empíricamente con la suma normalizada del número de veces que sucede. El período de retorno T o el tiempo esperado (en días) entre dos eventos de precipitación con p igual o superior a P viene dado por la inversa de dicha probabilidad acumulada.

Para ajustar la probabilidad de la precipitación diaria (de todo el año o de un mes del año en particular), en primer lugar se probaron los modelos de uso común: Pareto y Distribución Generalizada de Pareto (GPD), Gamma, Gumbel, Weibull, Exponencial, Log-normal II y III, y Valor Extremo Generalizado (GEV) y su opuesto (capítulo I). En segundo lugar, se probaron los nuevos modelos: versión modificada de la distribución Gumbel

opuesta (ec. II.1) y versión modificada de la distribución Log-Logística (ec. II.5).

III.3.2. Inferencia estadística

Para ajustar los parámetros de todas las funciones estadísticas hemos usado la inferencia por máxima verosimilitud, y en particular con el Perfil de Log-Verosimilitud (PLV) (Raue *et al.*, 2009). Sin embargo, la cantidad minimizada dentro de la PLV no fue exactamente el cuadrado de la diferencia entre la probabilidad empírica (π) y la probabilidad predicha por el modelo ajustado (π'). Sino que dichas probabilidades fueron normalizadas por π y $1 - \pi$ con tal de minimizar el error en las colas de la distribución:

$$\text{Min} \left(\sum_i (\pi'_i - \pi_i)^2 \right) \xrightarrow{\left(\frac{\pi'_i - \pi_i}{\pi_i} \right) \left(\frac{(1 - \pi'_i) - (1 - \pi_i)}{1 - \pi_i} \right)} \text{Min} \left(\sum_i \frac{(\pi'_i - \pi_i)^2}{\pi_i (1 - \pi_i)} \right) \quad (\text{III.1})$$

Con el fin de comparar la bondad de ajuste de los modelos se usó la medida del Error Absoluto Normalizado Medio (MNAE) y el Criterio de Información de Akaike (AIC). El MNAE es un índice de error adecuado para la precipitación porque compara los errores relativos desde la precipitación más escasa hasta la más abundante (ver MNAE en glosario). Por otro lado, de acuerdo con el Criterio de Akaike, el modelo teórico con un menor AIC es el que mejor se ajusta a los datos estudiados (Akaike, 1975; Burnham y Anderson, 2002). Además, se comparó si las distribuciones de probabilidad predichas son significativamente diferentes de las distribuciones observadas. Para ello, se usaron dos tests: El de Kolmogorov-Smirnov con *bootstrap* (Marsaglia *et al.*, 2003; Sekhon, 2010) y el de Anderson-Darling (Scholz y Stephens, 1987). Todos los análisis estadísticos de los datos fueron realizados usando el lenguaje R (Muenchen y Hilbe, 2010; R Development Core Team, 2010).

III.4. Resultados

En primer lugar, los modelos de uso común (GEV, GPD, Gamma, Weibull, Gumbel, Exponencial, Log-Normal II y III) y los dos nuevos modelos (la modificación de Gumbel y de la Log-Logística) se ajustaron a la distribución de la precipitación diaria de todos los meses del año, para cada observatorio. La figura 2 muestra algunos ejemplos de ajustes correspondientes a diferentes climas de España: A Coruña (Galicia), que presenta clima oceánico, Tenerife-Los Rodeos (Islas Canarias), con clima subtropical, y el observatorio de Navacerrada (Madrid), con clima de montaña. En la mayoría de los casos, los dos nuevos modelos se ajustaron mejor que el resto de modelos al rango entero de la precipitación. Si bien,

la distribución log-normal III y GPD opuesta se ajustaron mejor para las precipitaciones más bajas que para las más altas.

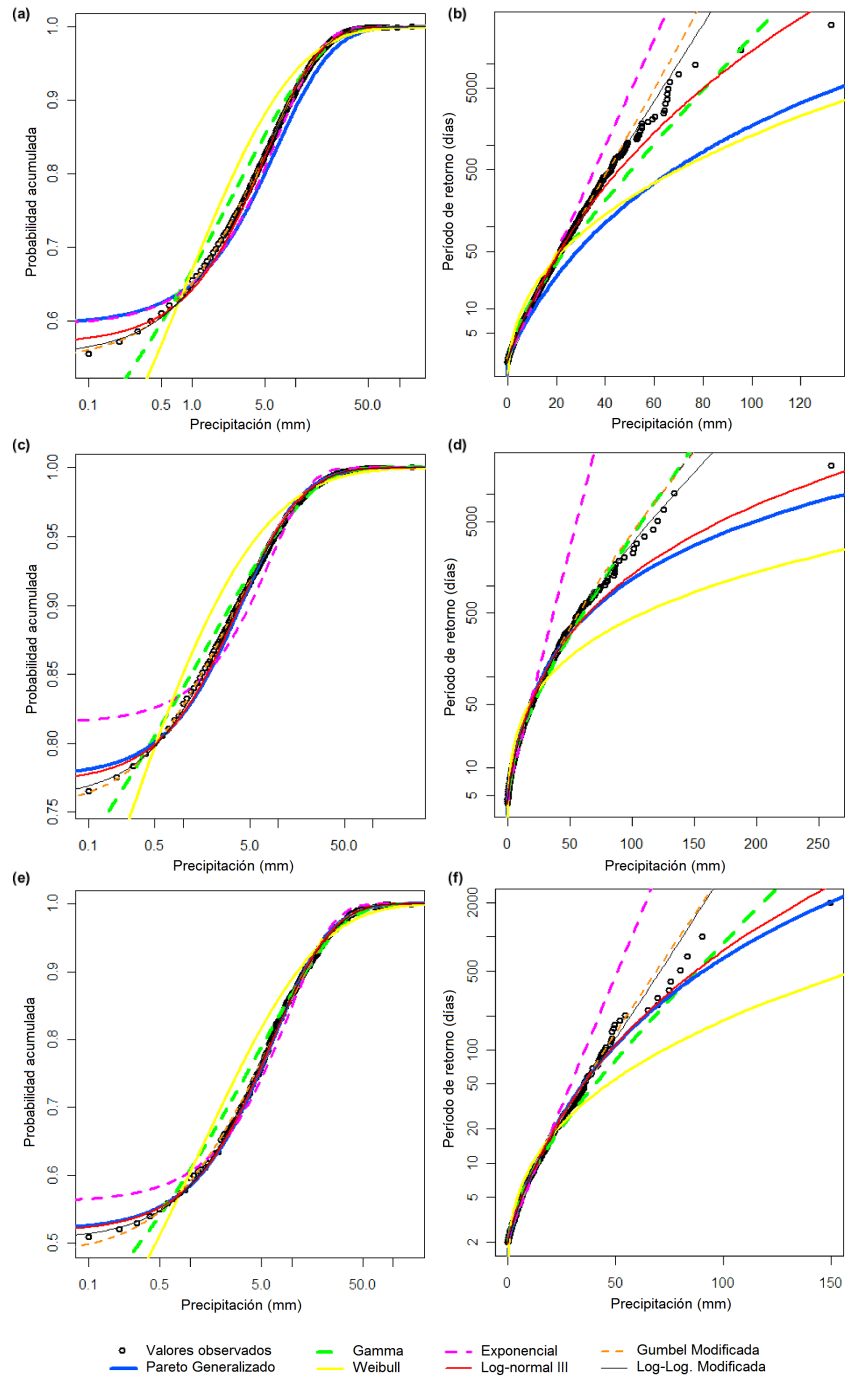


FIG. III.2. Ejemplos de probabilidad acumulada (izquierda) y período de retorno (derecha), de acuerdo con los valores empíricos y varios modelos ajustados, para

tres observatorios con diferentes climas: (a) y (b) Tenerife - Los Rodeos (clima subtropical); (c) y (d) A Coruña (clima oceánico); y (e) y (f) Navacerrada (clima de montaña).

A partir de la comparación entre la precipitación diaria observada y simulada de todos los observatorios y meses se obtuvo el Error Absoluto Normalizado Medio (NMAE). Las distribuciones de Gumbel (clásica), GEV, Log-normal II y Pareto mostraron un NMAE mayor que 1 para la mayoría de observatorios. Especialmente para la precipitación más alta y baja; por lo tanto esos modelos fueron descartados desde el principio. Sin embargo, los demás modelos obtuvieron un NMAE más bajo que 1 para todos los observatorios, excepto el modelo Weibull que presentó un NMAE elevado para algunos observatorios (Fig. III.2). La versión modificada del modelo de Gumbel opuesto y del Log-Logístico mostraron un NMAE menor que 0,10; mientras que los modelos restantes (Log-Normal III, GPD y GEV opuesto) presentaron un NMAE situado generalmente entre 0,10 y 0,20 (Fig. III.3). El AIC indicó que la modificación del Log-Logístico es en general mejor que la versión modificada de Gumbel. Sin embargo, para algunos observatorios y algunos meses, el modelo modificado de Gumbel puede ser mejor (Anexo 1).

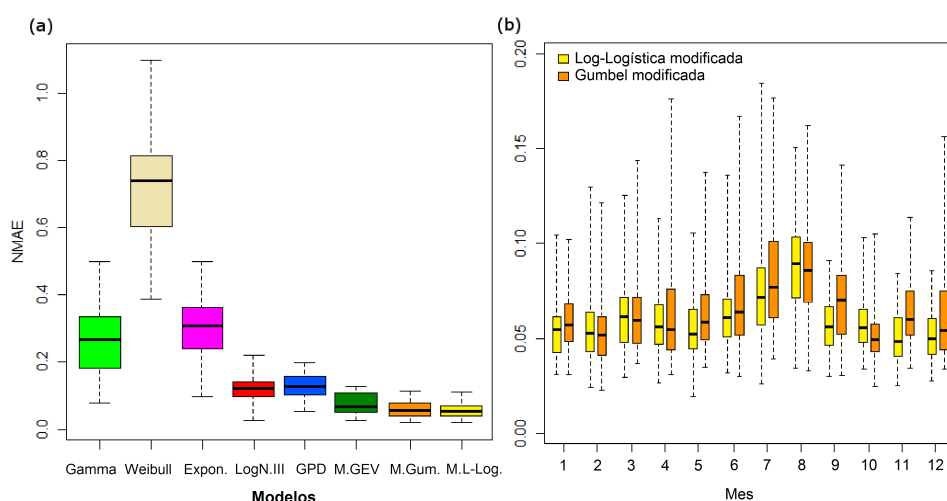


FIG. III.3. Comparación del Error Absoluto Normalizado Medio (NMAE) de la precipitación simulada respecto a la precipitación observada, para el conjunto de 108 observatorios, con ajuste de: (a) Varios modelos aplicados a la precipitación diaria de todo el año: Gamma, Weibull, Exponencial, Log-normal III, Distribución Generalizada de Pareto (GPD), Valor Extremo Generalizado Opuesto (M. GEV), Gumbel Modificada (M. Gum.) y Log-Logística Modificada (M. L-Log.). (b) Comparación entre el NMAE de Gumbel Modificada y Log-Logística Modificada, para los meses del año.

Por último, se llevó a cabo una comparación de la bondad de ajuste de los dos mejores modelos (modificación de Gumbel y del modelo alternativo) mediante la prueba de Kolmogorov-Smirnov y Anderson-Darling (Fig. III.4). La comparación muestra que el ajuste de ambos modelos a los datos observados no es significativamente diferente ($> 0,05$). El ajuste es peor para los meses de verano, debido en parte a los pocos días de precipitación.

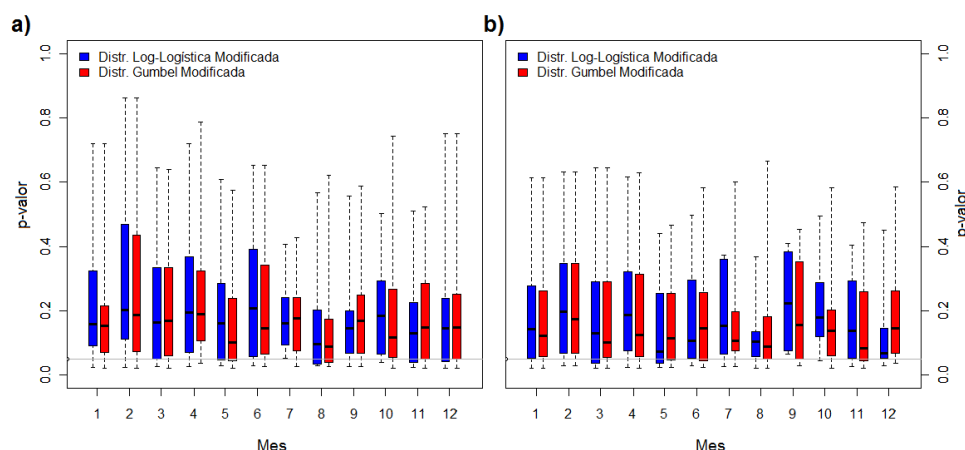


FIG. III.4. Comparación entre las distribuciones de probabilidad de la precipitación observada y simulada, de cada observatorio, usando las nuevas versiones de los modelos Gumbel y Log-Logístico. P-valor de: (a) test de Kolmogorov-Smirnov con remuestreo de *bootstrap* y (b) test de Anderson-Darling con ajuste para colas.

III.5. Discusión

III.5.1. Ventajas y limitaciones

En este capítulo hemos centrado nuestra atención en toda la precipitación diaria, en vez de estudiar la probabilidad de una precipitación específica, por ejemplo el máximo de precipitación diaria en un mes o un año (Bridges y Haan, 1972; Etoh *et al.*, 1986; Wilson y Toumi, 2005; Durman *et al.*, 2006; Eslamian y Feizi, 2006). Sin embargo, las funciones clásicas de probabilidad no pueden abarcar con precisión todo el intervalo de precipitación, desde la frecuencia de días sin precipitación hasta la frecuencia más alta de los extremos. Esto se debe a que, por lo general, las funciones de probabilidad utilizan dos o tres parámetros, lo cual parece ser insuficiente para ajustar los datos observados (Hanson y Vogel 2008). Los nuevos modelos de probabilidad desarrollados en este trabajo pueden representar con precisión todo el intervalo de probabilidad de la precipitación diaria (Fig. III.2 y Fig. III.5), ya que utilizan cuatro parámetros.

Para estudiar las limitaciones en la predicción de los extremos, todos los ajustes se realizaron excluyendo los dos valores mayores y los dos más

bajos. Con ello se obtuvo un NMAE = 0,15 para la predicción de dichos extremos (por ejemplo, Fig. II.5). De hecho, si representamos las curvas de probabilidad diaria de todos los meses del año (Fig. III.6), el error absoluto normalizado es similar para todas las escalas de precipitación, excepto para el valor mínimo medible (0,1 mm).

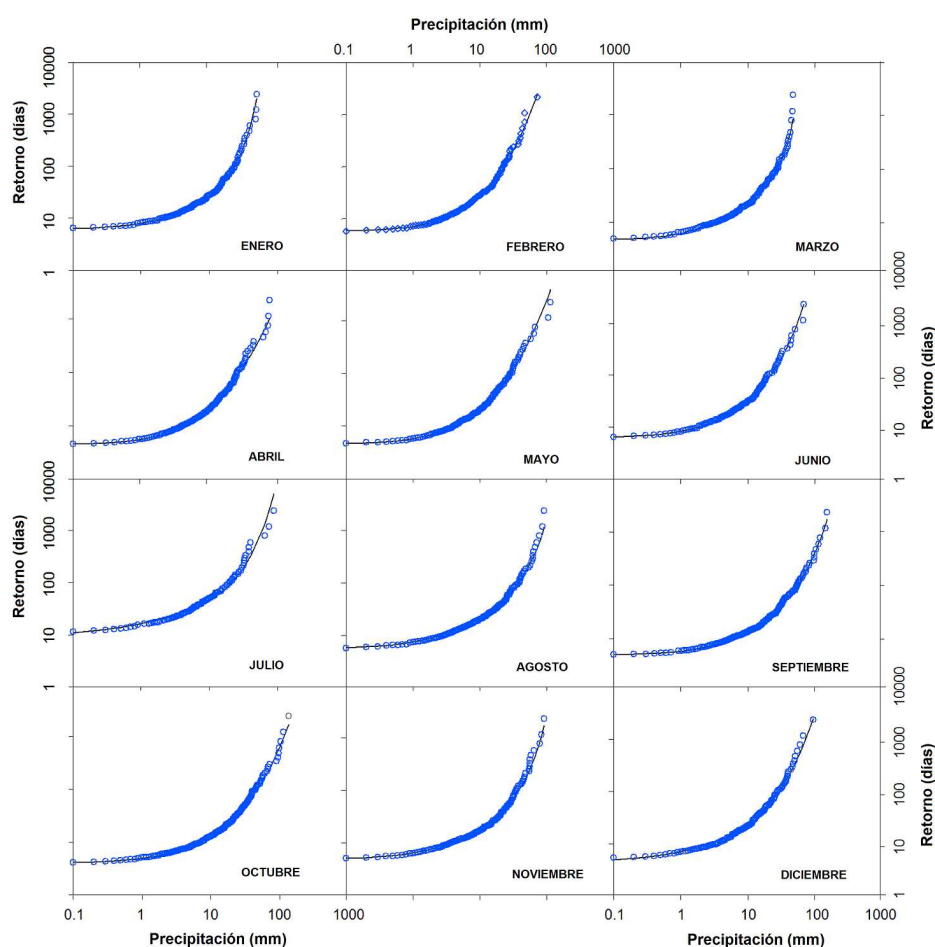


FIG. III.5. Período de retorno frente a la precipitación diaria para los meses del año y para el observatorio de El Prat (Barcelona): valores observados (círculos azules) y valores predichos (línea negra), según el modelo Log-Logístico modificado.

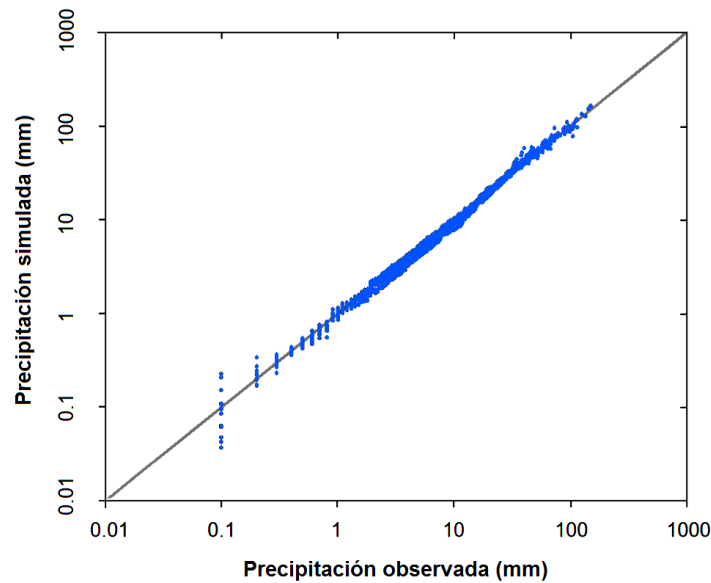


FIG. III.6. Validación de la precipitación diaria de todos los meses del año para el observatorio de El Prat (Barcelona).

Sin embargo, el uso de cuatro parámetros puede dar lugar a dificultades en el ajuste, ya que los algoritmos informáticos requieren valores iniciales muy cercanos a los “mejores estimadores” de los parámetros. Si no se inician adecuadamente, los algoritmos encuentran ajustes incorrectos. Además, la gran flexibilidad de los nuevos modelos puede causar diferencias significativas en ajustes de series de datos similares. Por lo tanto, puede haber cierta interdependencia entre los parámetros de los modelos, lo cual resta solidez a la interpretación o la comparación entre los diferentes ajustes. Por ejemplo, se encontró una correlación entre P_0 y P_1 con $R^2 = 0,49$, y entre w y $\log(P_1)$ con $R^2 = 0,60$, lo que sugiere la existencia de entrelazamientos parciales durante los ajustes. Por estas razones, podemos encontrar algunos casos en los que otros modelos se ajustan mejor a la precipitación diaria que los dos nuevos modelos propuestos. Esto ocurre especialmente para los observatorios con lluvias menos abundantes.

Por otro lado, una importante ventaja de estos modelos es que pueden transformar completamente una serie de precipitación diaria en una serie de probabilidad: serie del período de retorno (RPS). La RPS es una “serie temporal desnuda”, es decir, la información acerca de la media y de la variabilidad de la precipitación diaria recae principalmente en los parámetros mensuales extraídos. Sin embargo, hay un porcentaje de la variabilidad que es diferente entre dos RPS con igual longitud, debido a algunos factores que a priori son al azar. Esto significa que si dos RPS son muy largos, tienden a ser iguales en términos de su distribución de probabilidad, especialmente para los climas similares. Por ejemplo, una

RPS obtenida con la estandarización parcial será similar a la RPS obtenida por estandarización total, salvo en pequeños eventos aleatorios relacionados con la duración corta de la serie. Otro porcentaje de la variabilidad de la RPS puede ser debida a las características climáticas. Por ejemplo, dos series de precipitación pueden tener la misma distribución de probabilidad de la precipitación diaria, pero pueden diferir en la frecuencia de los períodos secos y húmedos así como en la alternancia de éstos.

Además, los dos nuevos modelos pueden ser utilizados para caracterizar y corregir la distribución de probabilidad de un conjunto genérico de datos de precipitación diaria. Su uso podría introducir mejoras en el *downscaling* estadístico de los modelos climáticos (Kysely, 2002; Widmann *et al.*, 2003) y por lo tanto eso se traduciría en mejoras para la predicción de cambios en los regímenes de precipitación (Kidson y Thompson, 1998; Perkins *et al.*, 2007). Del mismo modo, estos modelos también podrían ayudar a perfeccionar los modelos de pronóstico del tiempo y mejorar la calibración de los radares meteorológicos (Marks *et al.*, 2010).

III.5.2. Relaciones climáticas

Los modelos probabilísticos que se han desarrollado en este trabajo pueden ser usados para estudiar los patrones climáticos de la precipitación. Siguiendo con el mismo ejemplo de los apartados anteriores, en la figura III.7 se muestra la evolución mensual de la frecuencia de la precipitación diaria del observatorio de El Prat. En este sentido, los parámetros de ajuste para una distribución de probabilidad de precipitación deberían de reflejar adecuadamente las características del clima.

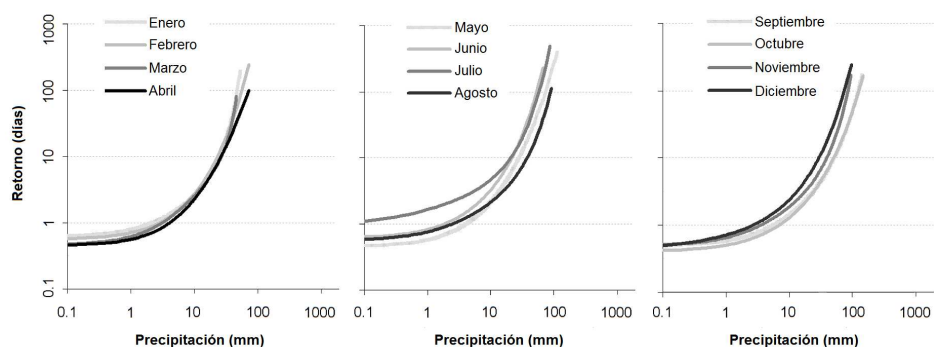


FIG. III.7. Comparación de curvas de período de retorno de la precipitación diaria, para los meses del año y para el observatorio de El Prat (Barcelona).

Debido a que el clima predominante en España es el mediterráneo (con un verano seco), se espera que los parámetros de ajuste de los modelos probabilísticos muestren ciertas diferencias para los meses del año. La Figura III.8 muestra esta propiedad, especialmente para el

parámetro de "escala" (P_1) y "suavidad" (k) de la distribución Log-logística modificada. Como cabía esperar, la "escala" de la distribución de probabilidad es menor y la "suavidad" es más alta para los meses menos lluviosos. Eso significa que en verano la probabilidad de precipitaciones débiles es más alta. Sin embargo, el parámetro de la "forma" (w) es ligeramente inferior en verano, lo que implica una mayor probabilidad (o menor período de retorno) de lluvia extrema. En definitiva, es coherente con las características del clima mediterráneo, predominante en España.

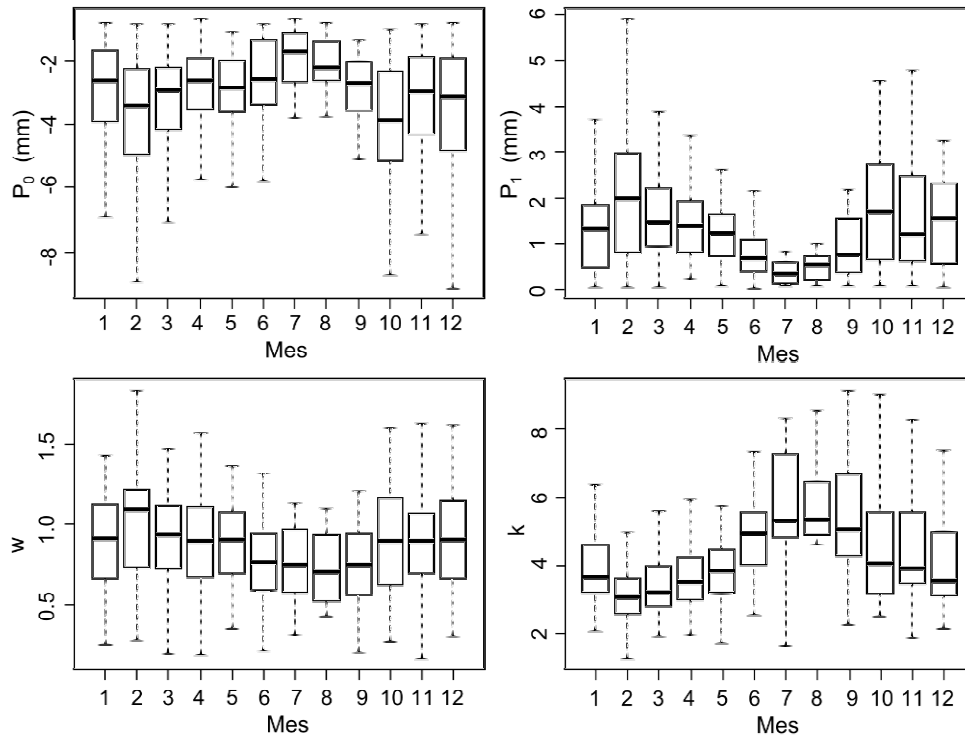


FIG. III.8. Valores mensuales de los cuatro parámetros ajustables de la versión modificada de la distribución Log-Logística (P_0 , P_1 , w , y k), aplicados para los 108 observatorios del estudio.

Las estaciones del año del clima mediterráneo pueden representar diferentes climas; por ejemplo, el verano se asemeja al clima desértico, la primavera al clima oceánico, y el invierno a los climas fríos. Por lo tanto, el buen ajuste de los nuevos modelos de probabilidad para los diferentes meses del año indica un buen funcionamiento para los diferentes climas. De hecho, muchos de los observatorios estudiados presentan otros climas (subtropical, oceánico y montaña) y también fueron modelados adecuadamente. De acuerdo con los resultados, los nuevos modelos funcionan mejor para los climas y meses más lluviosos (figuras III.3b y III.4), ya que la mayor probabilidad acumulada suaviza los errores aleatorios.

III.6. Conclusiones

III.6.1. Sobre la bondad de ajuste

La mayoría de las distribuciones de probabilidad usadas habitualmente (Valor Extremo Generalizado, Distribución Generalizada de Pareto, Gamma, Gumbel, Weibull, Exponencial y Lognormal) no fueron capaces de ajustarse adecuadamente al conjunto de la precipitación diaria. Por consiguiente, fue necesario usar modelos con tres o cuatro parámetros, y se probaron dos nuevos modelos: una generalización de la distribución Gumbel opuesta y una generalización de la distribución Log-Logística. Estos modelos presentaron un error medio entorno al 5%, y en cualquier caso inferior al 10% en la mayoría de las estaciones analizadas, para todas las escalas de la precipitación. Los tests de bondad de ajuste de Kolomorov-Smirnov y de Anderson-Darling concluyen que las distribuciones simuladas no son significativamente diferentes de las observadas ($p\text{-valor} > 0.05$).

III.6.2. Sobre las relaciones climáticas

Los modelos probabilísticos desarrollados en este trabajo reflejan de forma coherente el ciclo estacional, por lo que pueden ser utilizados para caracterizar los patrones climáticos de la precipitación. Sin embargo, en el análisis mensual de los ajustes, se producen mayores errores para los meses de verano, debido al escaso número de días con precipitación. Por esa razón, los nuevos modelos se ajustan mejor para los climas lluviosos.

**CAPÍTULO IV. SEGUNDO TEST: CORRECCIÓN
DE MODELOS CLIMÁTICOS EN EL PAÍS VASCO**

IV.1. Introducción

Para el análisis de los posibles cambios del régimen climático de la precipitación, existen una serie de modelos numéricos que tratan de simular las condiciones climáticas de la circulación atmosférica global. El Panel Intergubernamental del Cambio Climático (IPCC, Christensen *et al.*, 2007) proporciona diferentes escenarios climáticos globales, basados en las posibles evoluciones de la actividad económica mundial. En concreto se distinguen entre cuatro grupos de escenarios de emisiones de Gases de Efecto Invernadero (GEIs), en relación a los combustibles fósiles, según el modelo socioeconómico predominante en el futuro:

- A1: Escenarios de Rápido Crecimiento Global. La familia de escenarios A1 se divide en tres grupos según las distintas direcciones de la evolución del sistema energético: fuentes de energía intensivas de origen fósil (A1FI), de origen no fósil (A1T) o un equilibrio entre todas las fuentes (A1B).
- A2: Escenarios de Crecimiento Regional. En este caso se proyecta un crecimiento más lento pero continuado de la población mundial.
- B1: Escenarios de Crecimiento Sostenible Global. El crecimiento alcanza su nivel más alto a mediados del siglo XXI para disminuir posteriormente.
- B2: Escenario de Crecimiento Sostenible Regional. En este caso la población crece constantemente pero a un ritmo menor que en los escenarios de tipo A2, con cambios tecnológicos menos rápidos pero más variados que en los A1 y B1.

A priori todos los escenarios anteriores son equiprobables; sin embargo, el máximo de extracción de los combustibles fósiles ("peak oil") se prevé para antes de 2030 (Alekklett *et al.*, 2010; Carlson, 2011). Por esa razón, parece más probable que se dé alguno de los escenarios que proyectan una combinación con las energías renovables.

En este sentido, los miembros del proyecto ENSEMBLES (Niehörster *et al.*, 2008), impulsado por la Comisión Europea, han apostado fuertemente por centrar sus esfuerzos en el escenario A1B, que señala un rápido crecimiento global combinando con energías renovables. El proyecto ENSEMBLES está desarrollando un sistema de predicción climática por conjuntos basado en las últimas técnicas de modelización de alta resolución, de ámbito mundial y regional. El objetivo del citado proyecto es cuantificar y reducir la incertidumbre en la representación y la predicción del clima, para así maximizar el aprovechamiento de los resultados, vinculándolos a la agricultura, la salud, la seguridad alimentaria, energía, recursos hídricos, seguros y gestión del riesgo climático. En este marco global, en el País Vasco se está llevando a cabo el proyecto K-Egokitzen, que pretende evaluar los impactos producidos por el cambio climático regional y proponer posibles vías de adaptación, en respuesta a los futuros escenarios medioambientales.

Según la mayoría de proyecciones climáticas para el siglo XXI, el norte de la Península Ibérica se encuentra en una zona de frontera entre las disminuciones de la precipitación (sur) y el aumento de las mismas (norte) (Christensen *et al.*, 2007; Goubanova y Li, 2007). La expansión de la célula de Hadley (Lu *et al.*, 2007) podría hacer ascender la zona de subsidencia tropical (áreas desérticas) hacia el sur peninsular. Al mismo tiempo, se espera que el calentamiento del aire haga aumentar el contenido de agua precipitable en el norte de Europa, provocando una mayor intensidad de las precipitaciones (Déqué *et al.*, 2007; Goubanova y Li, 2007; Khon *et al.*, 2007). Sin embargo, el País Vasco se encuentra entre las dos zonas, por lo que resulta necesario realizar una regionalización de la señal climática de diferentes modelos, en relación a las características climáticas propias de la zona. Una forma de regionalizar la señal climática consiste en la calibración de los modelos con datos de los observatorios (Palmer *et al.*, 2008; Benestad, 2010).

Por todo ello, el objetivo principal de este capítulo es obtener tanto la distribución de probabilidad como la distribución espacio-temporal de las precipitaciones intensas, a partir de la corrección probabilística de diferentes proyecciones climáticas para este siglo. Se pretende construir una herramienta que sirva para el análisis del riesgo de inundación potencial y por lo tanto para cuantificar los posibles impactos derivados de ello.

IV.2. Materiales

IV.2.1. Área de estudio

Este estudio se centra en el País Vasco (Fig. IV.1) por la especial relevancia del cambio climático en aspectos como la planificación urbana y de los recursos hídricos. En particular, el riesgo de inundaciones por precipitación extrema unido al ascenso global del nivel del mar puede conllevar pérdidas materiales significativas en las zonas costeras vascas, especialmente por el hecho de que en dichas zonas se concentra la mayor parte de la población y muchas de las actividades antrópicas (Chust *et al.*, en prensa). Además, el riesgo de sequías también es de interés no sólo para las actividades humanas sino también para la conservación de las numerosas áreas naturales del País Vasco (Fig. IV.1), entre las que destaca el estuario del río Oka. Este estuario, situado en la Reserva de la Biosfera de Urdaibai (Vizcaya), es uno de los biológicamente más diversos y mejor conservados del País Vasco, junto con el del Bidasoa (Borja *et al.*, 2004).

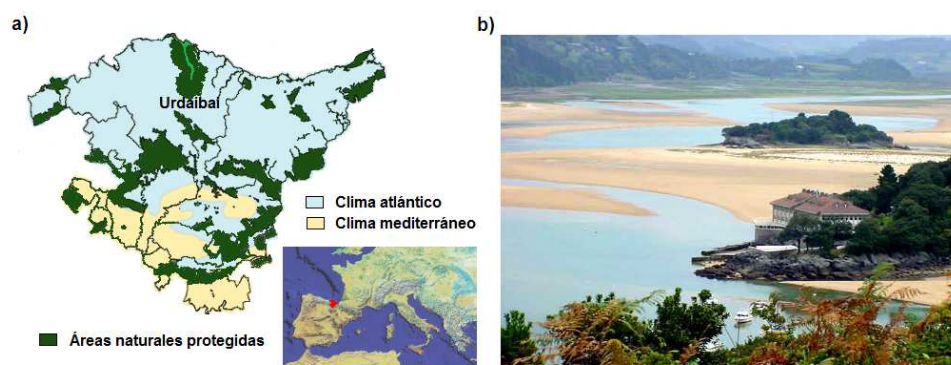


Figura IV.1. a) Mapa de los climas dominantes del País Vasco y las áreas naturales protegidas, según el Departamento de Ordenación del Territorio y Medio Ambiente. b) Fotografía del estuario del río Oka en Urdaibai (Vizcaya).

En cuanto a los aspectos climáticos, su ubicación en el norte de la Península Ibérica provoca que predomine la influencia de las ondas de chorro polar u ondas de Rossby, pero en algunas ocasiones se ve afectado también por el chorro subtropical (Alves y Verdière 1999, Peliz *et al.* 2002). El País Vasco, como la mayor parte de la cornisa cantábrica, se caracteriza por un relieve muy montañoso relativamente próximo al mar. Esas condiciones orográficas provocan que exista un gran contraste entre las precipitaciones de la vertiente cantábrica y las de la vertiente sur (Garmendia *et al.*, 1989).

La variabilidad climática en el País Vasco también es importante; la precipitación anual oscila entre los 400 mm en el sur de Álava, con clima mediterráneo, y los 2500 mm en el noreste de la Guipúzcoa, con clima oceánico (Capel-Molina 2000). Este último clima se caracteriza por precipitaciones regulares durante todo el año (en algunos casos con más del 50% de los días de lluvia). En cambio, las precipitaciones del clima mediterráneo se caracteriza por veranos secos y lluvias intensas concentradas en unos pocos días (Martín-Vide, 2004; Lana *et al.*, 2004), sobre todo en otoño y primavera. A lo largo del año, el número de días con precipitación por lo general oscila entre el 20 y el 30% de media, aunque puede alcanzar valores por debajo del 20% en julio y en algunos casos más del 50% en abril (Capel-Molina, 2000).

IV.2.2. Datos de precipitación observada

En este trabajo se han empleado un total de 156 observatorios de la Agencia Estatal de Meteorología. De éstos, 67 observatorios presentan una disponibilidad diaria (Fig. IV.2a). Sin embargo, la disponibilidad de muchos observatorios se interrumpe durante algunos años, por lo que el número de estaciones útiles varía fuertemente con el tiempo, obteniendo dos períodos

con máximos destacables, uno entre 1972 y 1980 y otro entre 1986 y 1995 aproximadamente (Fig. IV.2c).

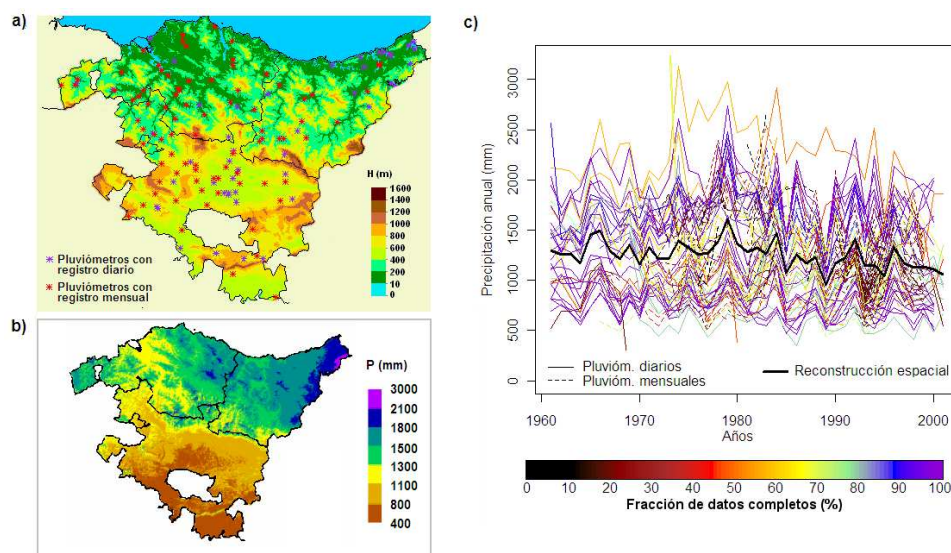


Figura IV.2. a) Localización de los observatorios empleados en este estudio (AEMET). b) Reconstrucción espacial de la precipitación media 1961-2000 (Moncho *et al.*, 2009b). c) Series temporales de los observatorios y su porcentaje de registros completos.

Los observatorios con registros mensuales presentan una media de 30% de huecos, y los registros diarios presentan el 15% de huecos. Para el período de referencia 1961-2000 sólo 19 estaciones disponen de una serie completa, mientras que el volumen total de datos de las 156 estaciones es de poco más del 50% respecto al volumen máximo posible. Por todo ello fue necesario emplear también una serie de referencia representativa de la precipitación del País Vasco (línea negra de Fig. IV.2b), reconstruida espacialmente mediante un modelo digital del terreno y un ajuste lineal múltiple (Moncho *et al.*, 2009b).

IV.2.3. Modelos regionales del clima

Los modelos empleados en este estudio son las salidas numéricas de 7 Modelos Regionales del Clima o RCM (Fig. IV.3), disponibles del proyecto europeo de ENSEMBLES-rt3 (<http://ensemblesrt3.dmi.dk/>). Dichos modelos presentan una resolución temporal diaria y una resolución espacial de 25 km x 25 km, y están forzados con el escenario futuro SRES-A1B, correspondiente a un rápido crecimiento global con una aportación combinada de las diferentes fuentes energéticas. La elección del escenario SRES-A1B por parte de los miembros de ENSEMBLES-rt3 se justifica porque se trata de un escenario intermedio entre las diferentes

proyecciones del aumento de la concentración de gases de efecto invernadero (Christensen *et al.*, 2007; Niehörster *et al.*, 2008).

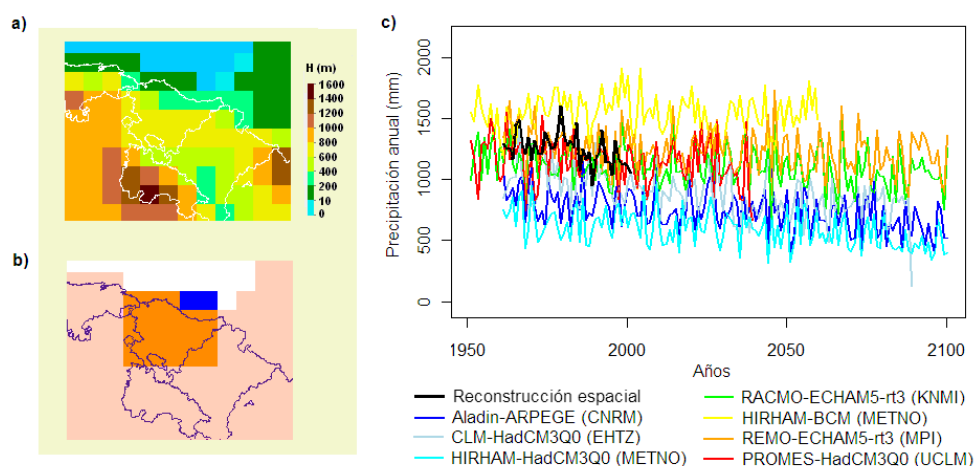


Figura IV.3. a) Modelo digital del terreno típico de los modelos regionales de ENSEMBLES-rt3. b) Elección de los puntos de malla de interés para el País Vasco, donde se distingue entre puntos de mar (azul) y puntos de tierra (naranja). c) Series temporales de la media espacial de 7 modelos regionales del clima.

En concreto, se tomó la media de 20 puntos de malla de los siguientes 7 modelos regionales:

- a) Aladin-ARPEGE: Centro Nacional de Estudios Meteorológicos (CNRM), Francia.
- b) CLM-HadCM3Q0: Instituto Suizo de Tecnología, Departamento de Geografía (ETHZ), Suiza.
- c) HIRHAM-HadCM3Q0: Instituto Meteorológico de Noruega (METNO).
- d) HIRHAM-BCM: Instituto Meteorológico de Noruega (METNO).
- f) RECMO-ECHAM5-rt3: Instituto Meteorológico de los Países Bajos (KNMI).
- g) REMO-ECHAM-rt3: Instituto Max-Planck de Meteorología (MPI), Alemania.
- h) PROMES-HadCM3Q0: Universidad de Castilla-La Mancha (UCLM), España.

IV.3. Metodología

IV.3.1. Distribución de probabilidad

Como se ha comentado en los capítulos anteriores, a partir de la curva empírica de la probabilidad acumulada (ECDF) de cada mes, se

puede transformar la serie temporal de precipitación de un modelo climático y corregirla con la ECDF de un observatorio, para un mismo intervalo temporal de referencia. Sin embargo, estas curvas pueden presentar irregularidades importantes, por lo que se propone ajustar diferentes modelos teóricos a la probabilidad empírica, de tal modo que se intente pulir dichas irregularidades. En este capítulo se usaron los dos modelos teóricos desarrollados anteriormente, en concreto, se trata de la versión generalizada de Gumbel (ec. II.1) y la versión generalizada de la distribución Log-Logística (ec. II.5).

IV.3.2. Estandarización probabilística

Si se extraen los parámetros mensuales de la distribución de probabilidad de una serie temporal, es posible construir una serie temporal del "período de retorno de cada día", es decir, el período de retorno asociado a cada valor de la precipitación diaria. Esta serie resultante, a partir de ahora se llamará Serie de Período de Retorno o RPS. Por lo tanto, definiremos este proceso como *estandarización total*, frente a la *estandarización parcial*, que se define como el proceso de generar un RPS con los parámetros extraídos de una parte de la extensión total de la serie temporal (por ejemplo, del período de referencia).

Un aspecto importante del proceso de estandarización es que se pueda partir de una serie de precipitación lo suficientemente detallada en cuanto a precisión de los valores para lograr transformaciones más flexibles. Por ejemplo, la salida numérica de precipitación de un modelo climático tiene muchos valores inferiores a 0,1 mm por lo que contiene más detalle en la distribución de probabilidad que un observatorio meteorológico (en el que en general la resolución es de 0,1 mm).

IV.3.3. Ampliación de series temporales

Para construir una distribución espacial y temporal completa de la precipitación del área de estudio, es necesario disponer del mayor número de series diarias posibles con una longitud extensa. Es posible extraer los parámetros mensuales de la distribución de probabilidad de cada una de las series temporales y comparar las RPS resultantes (o su logaritmo) para un mismo período de referencia. Si dos series temporales son similares, se puede transformar una en la otra con los parámetros obtenidos. Por esta razón, a fin de extender una serie temporal de precipitación, se puede utilizar otra serie similar (serie de control) con más extensión.

Por lo tanto, la RPS de la serie de control se obtuvo mediante estandarización parcial, es decir, con los parámetros obtenidos de la serie de control pero en el período común a la serie temporal estudiada. Una vez "estandarizada parcialmente" toda la serie de control, se aplicaron los parámetros de ajuste de la serie temporal estudiada. De esta manera se

fuerza a que la serie de control se parezca más a la serie estudiada en el período de referencia común. Por último, la transformación resultante de la serie de control sirvió para ampliar la extensión de la serie temporal de cada observatorio.

La elección de la serie de control para cada observatorio se realizó tomando como criterio aquella que tuviera un RPS con mayor correlación de Pearson y menor Error Medio Absoluto Normalizado (MNAE), respecto a la serie temporal que se quería simular. Además, se rechazaron todos los ajustes que presentaron un MNAE igual o superior al 20%.

IV.3.4. Transformación probabilística de los modelos climáticos

Debido a la baja resolución espacial de los Modelos Regionales del Clima, la orografía de una región pequeña como el País Vasco está muy suavizada. Por ejemplo, el pico más alto de la rejilla está entorno a 800m, frente a los casi 1600m de Aitxuri (Guipúzcoa). Por esa razón, se espera una débil correspondencia entre los puntos de malla y los observatorios.

Se puede decir que, en términos de valores absolutos, las simulaciones son poco realistas para la escala local, por ello es necesario aplicar algún tipo de regionalización matemática (Fernández-Ferrero *et al.*, 2009; Piani *et al.*, 2010). Si bien, se espera que la señal del cambio climático se refleje de alguna forma en el cambio de la distribución de probabilidad de la precipitación.

En este trabajo se ha optado por recoger la señal climática mediante el promediado de los puntos de malla más representativos espacialmente del País Vasco. Esa señal climática se encuentra muy suavizada por la baja resolución de los modelos climáticos, por lo que es necesario *sintonizarla* localmente mediante una transformación de la distribución de probabilidad (Anexo 2).

La transformación consiste en trasladar la forma de la distribución de probabilidad observada a la precipitación simulada por los modelos. Para ello, en primer lugar se “desnuda” la serie temporal de los modelos mediante una estandarización parcial. El período de referencia que se toma es lógicamente el período común respecto al observatorio que se pretende simular en cada caso. La RPS obtenida de cada modelo se “reviste” con los parámetros ajustados a la serie temporal del observatorio.

En este caso, la señal climática se mantiene inmutada, ya que se encuentra transmitida en la RPS. En una serie temporal de períodos de retorno se puede observar si existe algún cambio en el futuro sobre la frecuencia de ocurrencia de un tipo de evento con una determinada probabilidad *a priori*. Ello es independiente de si la RPS se estima mediante estandarización parcial o total, ya que la distribución temporal de la RPS puede ser *a priori* muy diferente en un período u otro del clima.

Sin embargo, si intentamos trasladar la curva de probabilidad de la precipitación de un observatorio con pocos datos, entonces nos podemos encontrar con disparidades en los resultados, ya que no se garantiza la

representatividad climática en el caso de series temporales muy cortas. Por ello es importante completar las series observadas con otras más largas. Después de completar todas las series cortas, el proceso de transformación probabilística de los modelos presenta menos problemas. Por lo tanto se trata de una estandarización probabilística completa (Fig. IV.4).

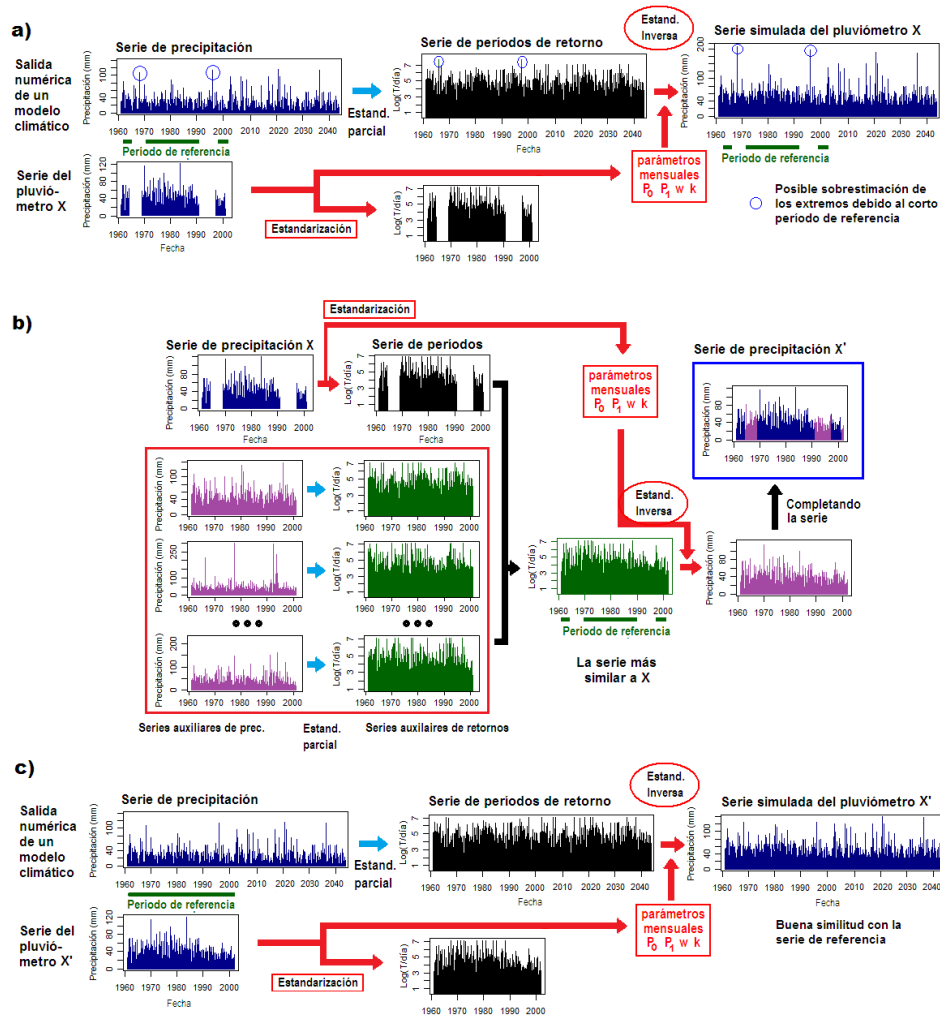


Figura IV.4. Resumen del método de estandarización probabilística completa: a) Esquema de la transformación probabilística de la serie temporal de un modelo regional climático, en base a un pluviómetro X. b) Rellenado de huecos del pluviómetro X mediante la transformación probabilística de las series temporales de otros pluviómetros y obtención de la serie rellenada X'. c) Transformación probabilística similar al de a) pero con la serie temporal X'.

IV.3.5. Análisis de los extremos de precipitación

Una vez transformada la distribución de probabilidad de cada modelo climático, se proyectó la precipitación esperada para un periodo de retorno de 100 años tanto para el pasado (1961-2000) como para principios de este siglo (2001-2040). Por lo tanto, para cada observatorio se estimó el cambio relativo de la precipitación esperada según los 7 modelos climáticos. El error estimado de dicho cambio se calculó suponiendo una distribución normal del error, es decir, el intervalo de confianza del 95% se estimó en relación a dos veces la desviación típica de los cambios proyectados por los 7 modelos.

Para representar un mapa sobre el cambio relativo de la precipitación extrema, se realizaron dos pasos. En primer lugar, se estimó una regresión lineal múltiple entre una serie de predictores de los observatorios simulados (con registros diarios) y los cambios relativos predichos para la precipitación. Los predictores que se tomaron fueron: las coordenadas geográficas, la altura y la orientación. A partir de la regresión múltiple anterior, se estimaron los valores del cambio de la precipitación extrema y su error para los observatorios con registro mensual. En segundo lugar, se realizó una interpolación mediante Thin Plate Splines (TPS) con dos dimensiones (Green y Silverman, 1994; Furrer *et al.*, 2010), con todos los observatorios, tanto los simulados directamente como los estimados mediante regresión lineal múltiple. Por lo tanto, se realizó un estudio de la calidad de la interpolación mediante validación cruzada para los dos pasos seguidos. Para ello se empleó la correlación de Pearson y el Error Medio Absoluto Relativo (RMAE, ver glosario).

Por último se estimó la precipitación sub-diaria a partir del método del índice n (Moncho *et al.*, 2009a). En concreto, se usó la definición implícita en la ecuación II.11. Se estimó el índice n para todos los observatorios y simulaciones a partir de la precipitación de al menos 3 días seguidos. Finalmente, se estudiaron los posibles cambios de la distribución de probabilidad a lo largo del siglo XXI.

Todos los tratamientos de los datos y de los resultados se llevaron a cabo mediante paquetes estadísticos basados en el lenguaje R (R Development Core Team, 2010).

IV.4. Resultados

IV.4.1. Rellenado y transformación probabilística

En primer lugar se realizó un proceso de ampliación y de relleno de huecos mediante la estandarización parcial de diferentes series temporales de control. Se tomaron todos los observatorios de al menos 40 años (11 de 67 observatorios) y se transformaron de acuerdo con los parámetros de probabilidad de las series más cortas, correspondientes a los modelos de

las ecuaciones IV.2 y IV.3. De los 56 observatorios con menos de 40 años de datos, sólo se logró completar correctamente 34 de ellas, ya que el resto no pasaron las pruebas mínimas de ajuste (se obtuvo un error normalizado medio superior al 20%).

Para las estaciones ampliadas correctamente, el proceso las prolongó entre un 3 y un 64% (la media fue 21%), hasta completar los 40 años del período de referencia (1961-2000). Se obtuvo que la correlación de Pearson (R) entre cada serie observada y la que más se asemeja oscilaba entre un mínimo de 0.61 y un máximo de 0.90 (con media 0.74). El error absoluto normalizado medio (MNAE) de la ECDF de las series simuladas se situó entre el 0.02 y 0.12 (con media 0.05) respecto a la correspondiente serie observada.

Por lo tanto, en total se consiguieron 45 observatorios con al menos 40 años de datos. Las series temporales de dichos observatorios se estandarizaron de nuevo mediante los dos modelos de probabilidad (ecuación IV.2 y IV.3). Para cada mes climático de cada observatorio se tomó el ajuste teórico con menor error absoluto normalizado, que en un 72% de los casos fue la versión generalizada de la Distribución Log-Logística (Fig. IV.5).

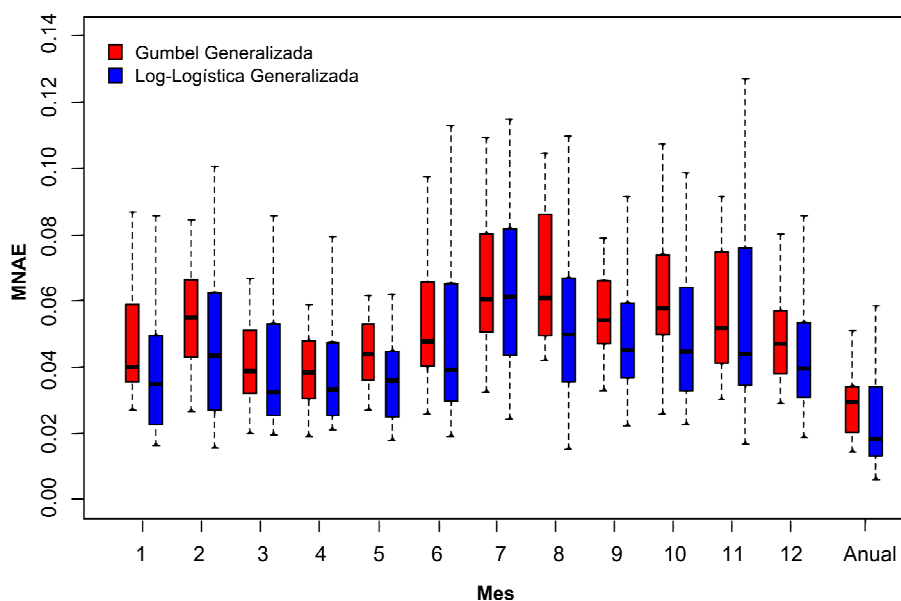


Figura IV.5. Comparación del Error Absoluto Normalizado Medio (MNAE) de los dos modelos probabilísticos empleados para la estandarización de los 45 observatorios con al menos 40 años de datos.

Con los parámetros obtenidos de la estandarización, se transformaron las series de precipitación de los 7 modelos regionales del clima simulando cada uno de los observatorios. En la Figura IV.6 se

muestran algunos ejemplos en los que las curvas de períodos de retorno terminan ajustándose a la curva empírica de cada estación. En dichos casos se observa como la corrección de la distribución de probabilidad anual se transmite adecuadamente a los modelos climáticos. Eso incluye también la frecuencia de días sin precipitación, que está implícito en el extremo inferior de las curvas de probabilidad (inversa del período de retorno mínimo). Nótese que, debido a que los modelos regionales del clima presentan un mayor número de días de precipitación, el cambio de la transformación del período de retorno mínimo (Δd) será positivo.

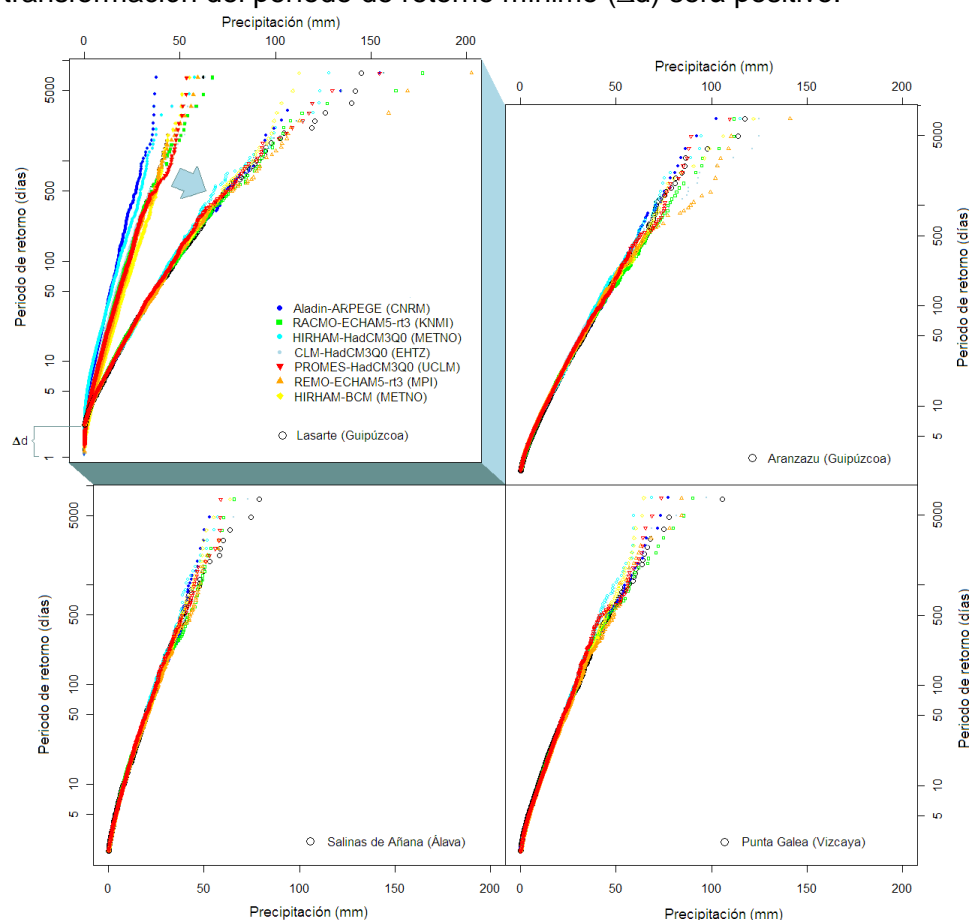


Figura IV.6. Ejemplo de transformación de la curva de períodos de retorno de la precipitación diaria (para todo el año) de los 7 modelos climáticos del estudio. En la parte de arriba a la izquierda se aprecia el cambio de las curvas de los modelos antes y después de la transformación en las 7 series temporales que tratan de simular la precipitación de Lasarte (Guipúzcoa). En la parte de arriba a la derecha, se representa el resultado final de las 7 simulaciones para Aranzazu (Guipúzcoa), mientras que abajo se repite lo mismo para Salinas de Añana (Álava) y Punta Galea (Vizcaya), de izquierda a derecha.

IV.4.2. Validación de la transformación probabilística

Si se compara la media climática mensual y la desviación típica diaria de cada serie simulada respecto a la correspondiente serie climática observada, entonces se aprecia que el valor relativo del error absoluto medio (RMAE) oscila generalmente entre el 1 y el 15%, dependiendo de la estación del año y del modelo climático (Fig. IV.7). El valor promedio del RMAE para todos los observatorios simulados se sitúa por debajo del 5% en todos los casos y en todos los meses, exceptuando el Aladin-ARPEGE para los meses de verano, con un error medio entorno el 10%. Algo similar ocurre con el sesgo relativo (RBIAS) que también es inferior al 5% en la mayoría de los casos, exceptuando de nuevo el modelo Aladin-ARPEGE. El número de días sin precipitación simulado presenta un error relativo todavía inferior, cercano al 1%, y con un sesgo medio despreciable (Fig. IV.8). Estos errores relativos contrastan con los obtenidos antes de corregir las series: por ejemplo, los errores en la media se situaban generalmente entre el 15 y el 30%, mientras que el error en los días sin precipitación llegaba a superar el 50% en algunos casos.

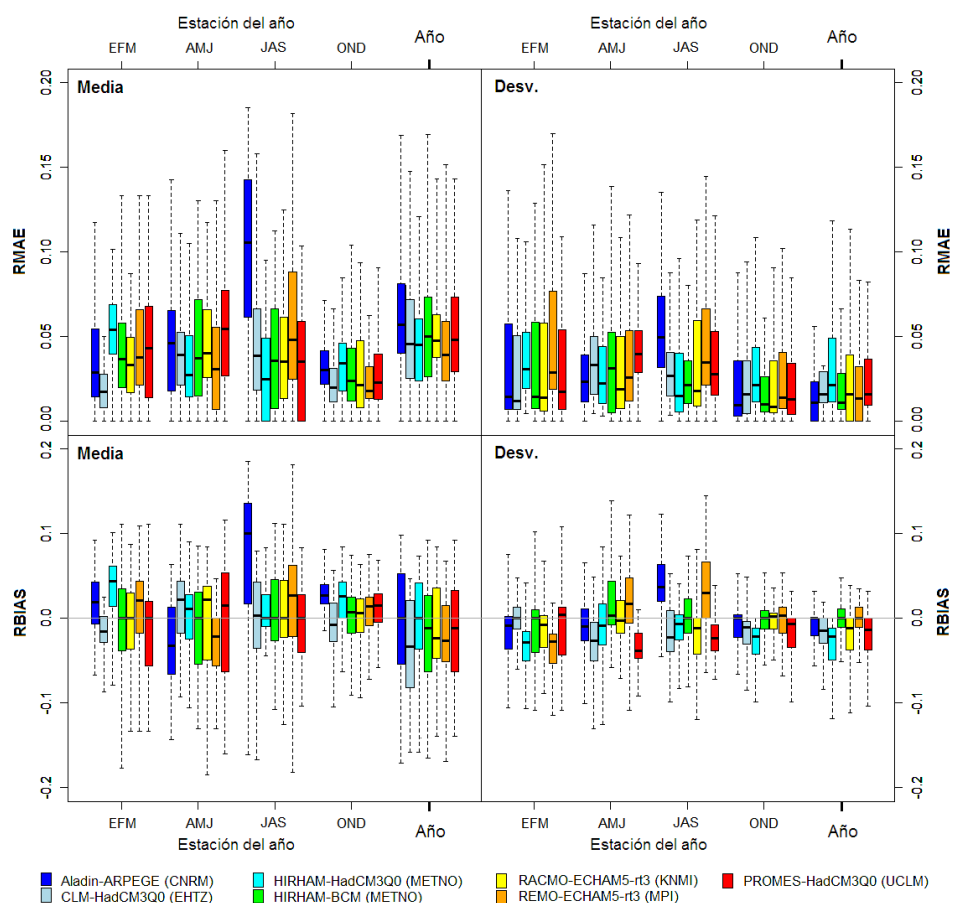


Figura IV.7. Error absoluto medio relativo (RMAE) y sesgo relativo (RBIAS) de las medias y desviaciones típicas de las series temporales simuladas. Cada barra representa el conjunto de todos los observatorios según cada uno de los 7 modelos regionales.

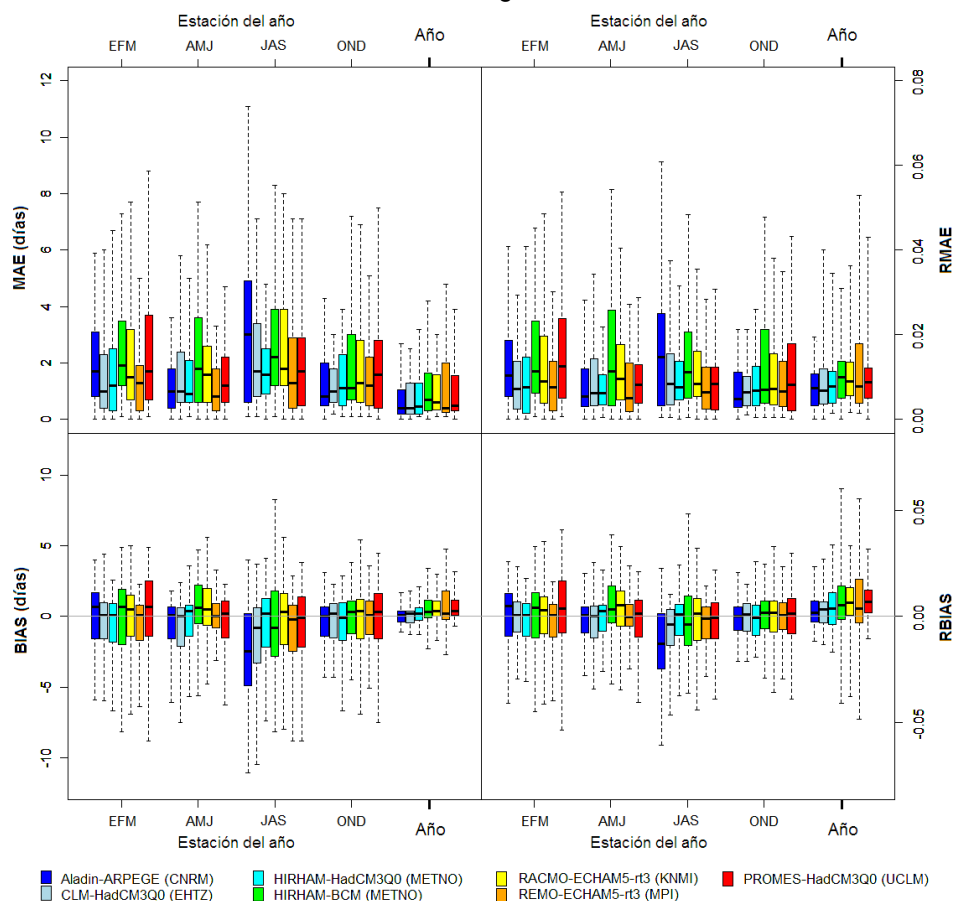


Figura IV.8. Valores absolutos y relativos del error medio (MAE) y sesgo (BIAS) de los días sin precipitación de las series simuladas. Cada barra representa el conjunto de todos los observatorios según cada uno de los 7 modelos regionales.

Además del bajo error relativo, el conjunto de las series simuladas presenta una elevada correlación con los observatorios de referencia. Así por ejemplo, la correlación de las medias mensuales es de en torno a $R^2 = 0,99$ para todos los modelos, con una pendiente que oscila entre el 1,05 y 1,14 dependiendo del modelo (Tabla IV.1 y Fig. IV.9). La correlación es ligeramente más baja para la simulación de la desviación típica diaria y el número de días sin precipitación, con valores de R^2 que oscilan entre 0,96 y 0,98, pero la pendiente es más cercana a 1, ya que oscila entre 0,95 y 1,02.

IV. Segundo test: Corrección de modelos climáticos en el País Vasco

Tabla IV.1. Correlación y pendiente de los valores climáticos de la media, desviación típica diaria y número de días sin precipitación, de las series simuladas mediante los modelos, frente a los observatorios de referencia.

Modelo regional	Correlación R^2			Pendiente		
	Precip. Media	Desv. típica	Días sin preci.	Precip. Media	Desv. típica	Días sin preci.
Aladin-ARPEGE	0,99	0,98	0,97	1,07	0,97	0,99
CLM-HadCM3Q0	0,99	0,96	0,97	1,07	0,95	1,01
HIRHAM-HadCM3Q0	0,99	0,98	0,98	1,14	1,00	0,98
HIRHAM-BCM	0,99	0,98	0,97	1,05	0,97	0,99
RACMO-ECHAM5-rt3	0,99	0,97	0,97	1,08	0,99	1,02
REMO-ECHAM5-rt3	0,99	0,97	0,97	1,08	1,01	0,96
PROMES-HadCM3Q0	0,99	0,97	0,96	1,11	0,97	0,96

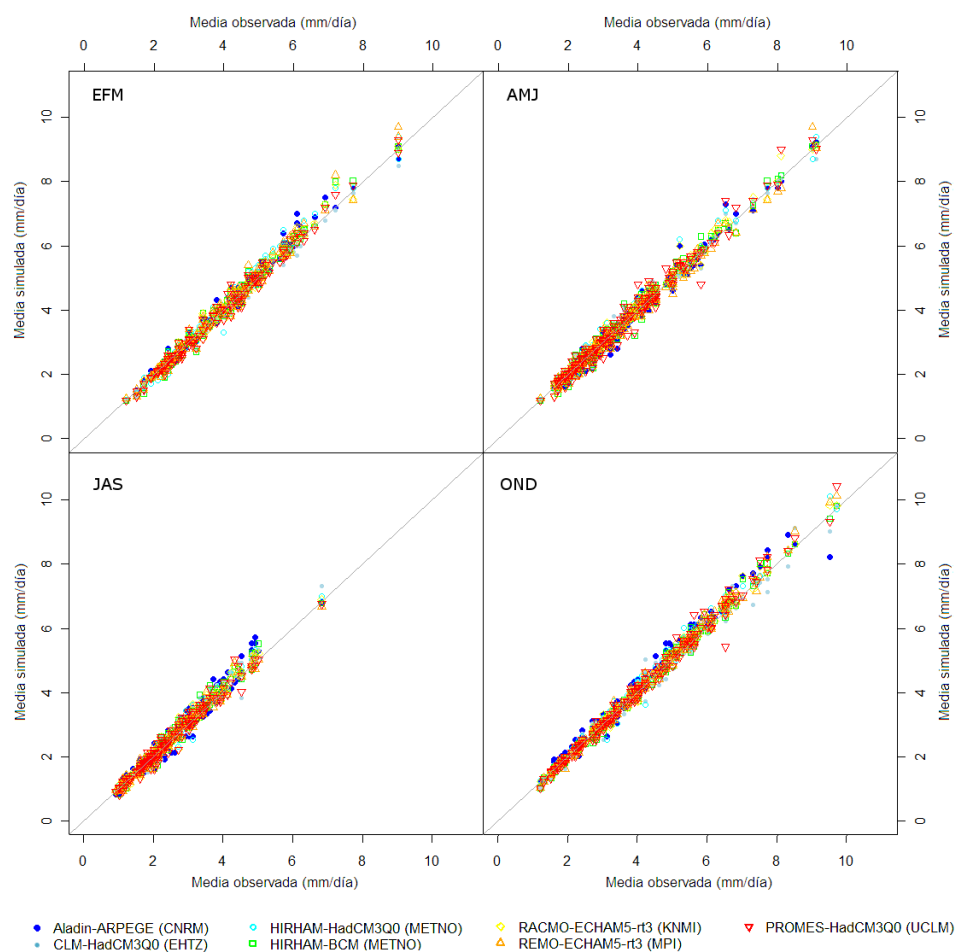


Figura IV.9. Comparación entre los valores observados y simulados de las medias mensuales de todos los observatorios.

IV.4.3. Análisis del cambio de los extremos de precipitación

Para analizar los posibles cambios de la precipitación extrema, el estudio se centró en el período de retorno de 100 años. Es decir, a partir de las curvas de probabilidad teóricas ajustadas para los observatorios simulados, se prolongó el valor de la curva para un retorno de 100 años. Se realizó una comparación entre el valor observado para el período de referencia (1961-2000) y el valor simulado para el mismo período, así como el valor simulado para el período 2001-2040 y 2041-2080. En la Figura IV.10 se aprecia un aumento de dicha precipitación de en torno el 30% en la mayoría de observatorios y modelos. Para el período 2001-2040 la variación no es tan clara, aunque algunos observatorios del oeste del País Vasco coinciden en dar un aumento de la precipitación extrema para la mayoría de modelos analizados.

Para observar mejor la distribución de esos cambios, se realizó una interpolación de los valores extremos proyectados para el período 2001-2040. La regresión lineal múltiple entre los valores del cambio relativo y los predictores de los observatorios (coordenadas, altura y orientación) obtuvo una correlación de $R^2 = 0,83$ y un Error Absoluto Medio Relativo de $RMAE = 0,2$. Para el intervalo de confianza del cambio de la precipitación también se realizó una regresión múltiple con los mismos predictores. La validación cruzada mostró que la correlación era más baja que en el caso anterior, en torno a $R^2 = 0,48$, pero el error bajó a $RMAE = 0,15$.

Con todo lo anterior, se observó un área en el oeste del País Vasco con un cambio proyectado de hasta el +30%, con un intervalo de confianza superior al 99% (Fig. IV.11). Si bien, puesto que el error relativo de esta última interpolación es del 15%, cabe esperar que la probabilidad del cambio sea menor.

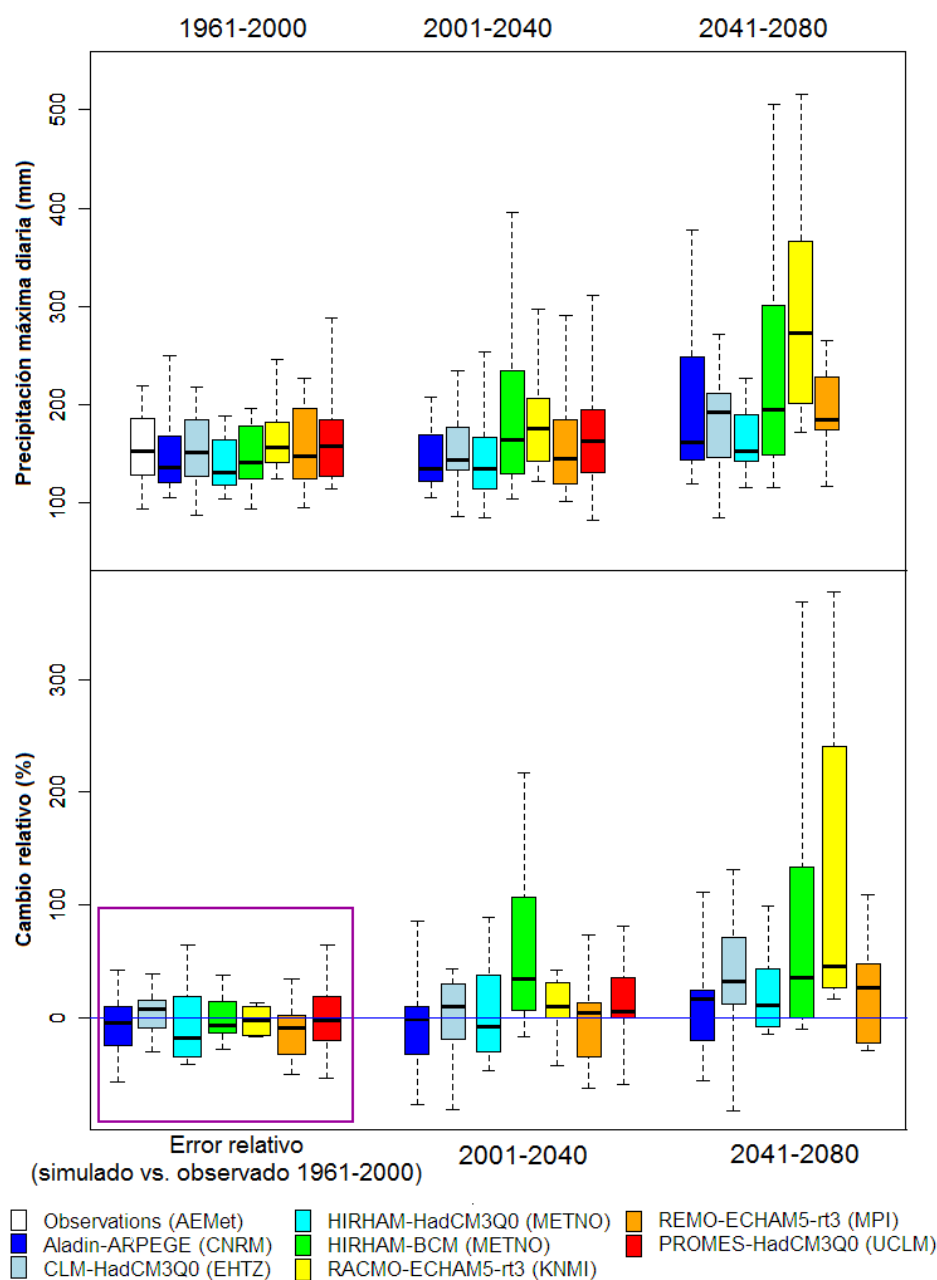


Figura IV.10. Proyección de la precipitación máxima diaria para un período de retorno de 100 años según las curvas teóricas ajustadas a los observatorios simulados. Arriba se observa la proyección del valor absoluto para los tres periodos (1961-2000, 2001-2040 y 2041-2080). Abajo se muestra el cambio relativo para 2001-2040 y 2041-2080 y el error de la simulación en comparación con las series temporales observadas.

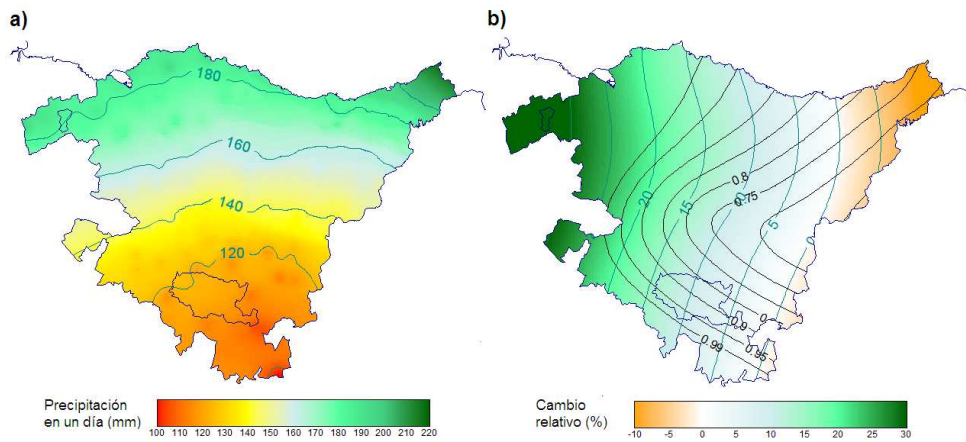


Figura IV.11. a) Precipitación máxima en un día simulada para 1961-2000 con un período de retorno de 100 años. Los valores se estiman según las curvas teóricas ajustadas a los observatorios simulados, y se interpolan mediante un ajuste espacial múltiple. b) Cambio relativo de la precipitación descrita en a), para el período 2001-2040 respecto 1961-2000. Las líneas negras representan los intervalos de confianza estimados a partir de la estadística de los 7 modelos regionales del clima.

Respecto al índice n , se estudió la distribución de probabilidad empírica para toda la precipitación y para aquella que superó los 50 mm en un día (Fig. IV.12). Para el conjunto de la precipitación no se observan cambios significativos. Esto se debe a que la tipología de la precipitación es tan diversa que enmascara cualquier cambio del índice n , al menos para los períodos estudiados en el País Vasco. Sin embargo, para la precipitación que supera los 50 mm en un día, sí se observan cambios en la probabilidad de los índices entre 0,6 y 0,8. Eso implica un desplazamiento de la media en 3 ± 1 centésimas (de 0,57 a 0,60). Es decir, no sólo aumentaría la precipitación máxima diaria (Fig. IV.10 y IV.11) sino que además cabría esperar una mayor concentración en el tiempo, a escala subdiaria.

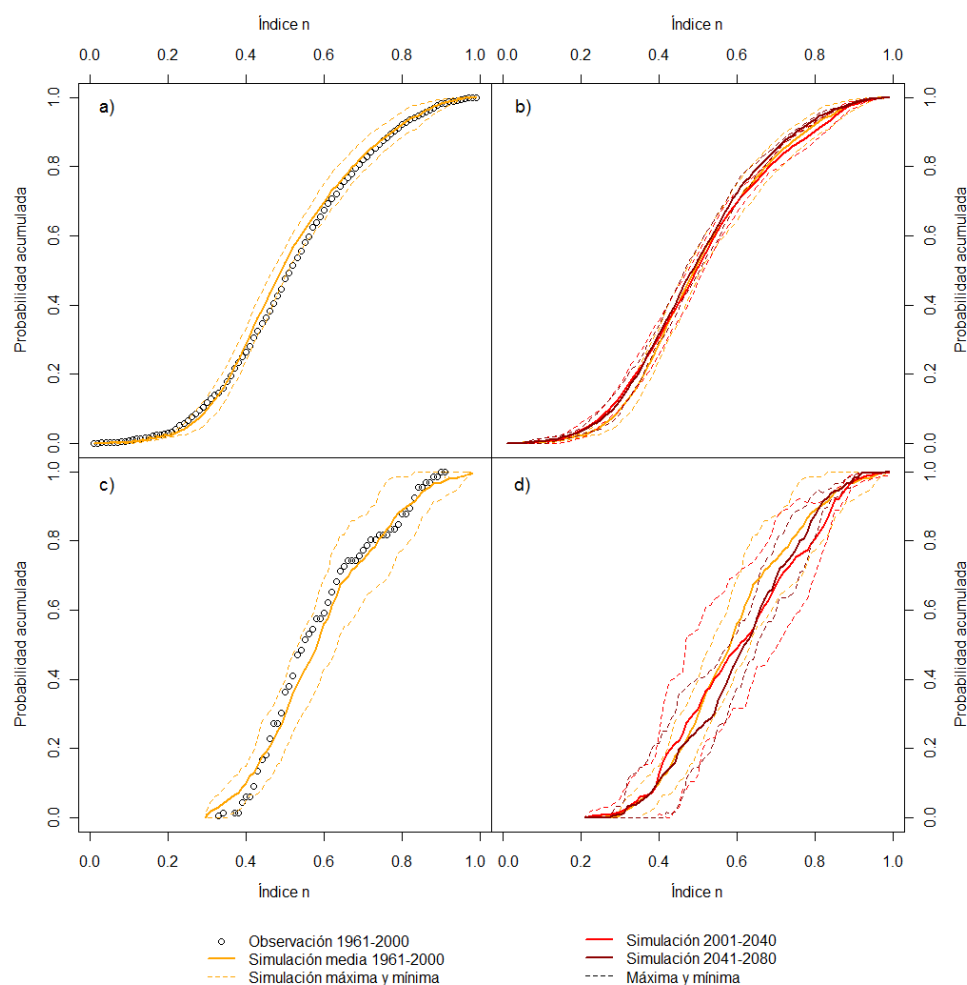


Figura IV.12. a) Distribución de probabilidad empírica del índice n para el conjunto de todos los observatorios en el período 1961-2000, e intervalo de las simulaciones de los 7 modelos regionales. b) Proyección del cambio para 2001-2040 y 2041-2080. c) y d) Respectivamente, lo mismo que a) y b) pero sólo las precipitaciones que superaron los 50 mm en un día.

IV.5. Discusión

IV.5.1. Elección de la metodología

Para corregir las distribuciones de probabilidad de los modelos regionales del clima, en este trabajo se optó por emplear curvas teóricas ajustadas a los observatorios, ya que las curvas empíricas pueden incorporar irregularidades adicionales a la curva de probabilidad corregida de los modelos. Además, las curvas teóricas permiten prolongarse fácilmente para estimar los extremos de precipitación, mientras que las curvas empíricas no. Sin embargo, existen algunos problemas en el empleo de los ajustes teóricos. Uno de ellos es el problema de la estacionalidad y otro el de la estacionariedad.

El primero se refiere a la diferencia de la probabilidad de la precipitación dependiendo de la estación del año e incluso del mes o de la quincena del mes. Eso significa que puede resultar necesario ajustar una distribución de probabilidad para cada estación del año o para cada mes (o a escalas inferiores) hasta conseguir que la serie temporal de los períodos de retorno (RPS) sea totalmente no-estacional (sin ninguna autocorrelación). El problema de la separación estacional de las distribuciones de probabilidad es la decisión de los cortes. En este trabajo se han tomado los meses oficiales, por tres razones:

- 1) Se considera que con 12 secciones es suficiente para que la RPS resultante sea no-estacional (Elshamy *et al.*, 2009).
- 2) Porque las fronteras de los cortes son bastante suaves, es decir, no hay tanta diferencia entre la probabilidad de un mes y el siguiente (tal vez exceptuando julio).
- 3) Si se escoge un mes natural (30 días móviles) habría 365 distribuciones de probabilidad diferentes por cada observatorio y modelo, lo cual dificulta el tiempo de cálculo, y según los puntos anteriores no parece necesario.

El segundo problema se refiere a que las relaciones empíricas entre las variables del pasado no tienen porqué ser estacionarias en el futuro. Este problema no afecta en gran medida al método de la transformación probabilística por dos razones:

- 1) Cuando se estandariza parcialmente una serie temporal (futuro con pasado), se obtiene una RPS que contiene toda la señal de cambio climático (ya que se puede analizar el cambio de la frecuencia de ciertos períodos de retorno). Cuando las 12 distribuciones de probabilidad de un observatorio son añadidas a la RPS de un modelo regional del clima, se refleja un cambio en las distribuciones de probabilidad del futuro en relación con el pasado. Por lo tanto, en realidad no se está forzando el futuro con una probabilidad del

pasado, ya que la información esencial de la probabilidad viene dada por la RPS.

- 2) Aunque en el futuro, por ejemplo, un mes de septiembre se parezca un poco más a un mes de agosto que ahora, ese posible cambio estará contenido también en la RPS, ya que se observaría que pasa de ser no-estacional en el pasado a presentar ciertas diferencias estacionales en el futuro.

En definitiva, todo parece indicar que las relaciones físicas entre la probabilidad de la precipitación y los escenarios futuros de cambio climático están salvaguardadas por la estandarización en forma de RPS.

Como hemos visto, una RPS es una “serie temporal desnuda”, es decir, la información sobre la media y la variabilidad de la precipitación diaria recae principalmente en los parámetros mensuales extraídos. Sin embargo, hay un porcentaje de la variabilidad que difiere entre los dos RPS de igual longitud, debido a las características de las rachas secas y lluviosas, así como debido a algunos factores que a priori son aleatorios. Eso significa que si dos RPS son muy largos tienden a ser iguales en términos de su distribución de probabilidad, especialmente para los climas similares. Por ejemplo, en principio, una RPS obtenida con estandarización parcial (suficientemente larga) será similar a la RPS obtenida mediante estandarización total, salvo en pequeños eventos aleatorios relacionados con la corta extensión de la serie.

IV.5.2. Transformación de la probabilidad

Las distribuciones de probabilidad empleadas en este trabajo fueron escogidos frente a otros porque presentan mejores resultados para diferentes climas (Anexo 1). De hecho, para el País Vasco, las dos distribuciones se ajustan satisfactoriamente a la precipitación mensual, con un error medio normalizado que en general es inferior al 6% (Fig. IV.5). Sin embargo en verano el error de los ajustes es algo mayor, debido al gran número de días con precipitación leve, que hace disparar el error relativo.

En cualquier caso, las series temporales simuladas se asemejan adecuadamente a las series observadas, ya que el MNAE es en general inferior al 5% (Fig. IV.7 y IV.8). Sin embargo, eso no es garantía de que las distribuciones de probabilidad se parezcan lo suficiente como para decir que son estadísticamente indistinguibles. Una forma de medir esto es el test de Kolmogorov-Smirnov (KS; Marsaglia *et al.*, 2003, y Sekhon, 2010) y el test de Anderson-Darling (AD; Scholz y Stephens, 1987). Los resultados muestran que el p-valor de la mayoría de observatorios se sitúa por encima del 0,05, y por lo tanto se concluye que las distribuciones de probabilidad simulada son en general indistinguibles respecto a las observadas.

Sin embargo, si analizamos el p-valor del KS por trimestres (Fig. IV.13), se aprecia que en los meses de verano casi la mitad de los

observatorios se sitúan por debajo de la significancia del 0,05. Algo similar ocurre con el test AD. Eso puede explicarse por el gran número de días sin precipitación. Es decir, estos tests dan mucha confianza a valores con una alta frecuencia de ocurrencia, por lo que las pequeñas diferencias en el número de días secos provocan que el p-valor se aproxime a cero.

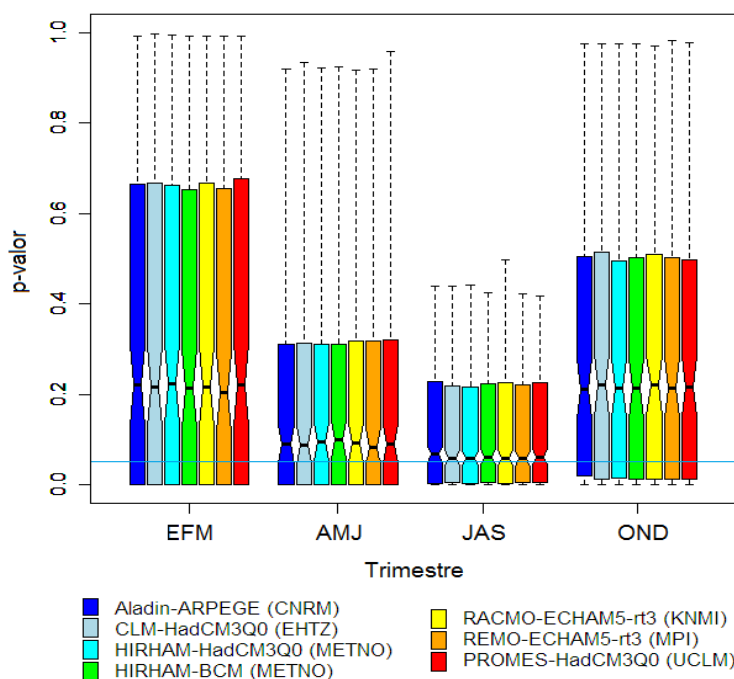


Figura IV.13. P-valor del test de Kolmogorov-Smirnov de la precipitación simulada frente observada, para los 45 observatorios con al menos 40 años.

Uno de los aspectos más sensibles de la metodología de RPS es que mantiene prácticamente intactas las características de las rachas secas y húmedas. Es decir, una transformación probabilística puede cambiar la escala de la precipitación y el número de días secos, pero no la disposición relativa y la alternancia de días secos con días de lluvia. Por lo tanto es importante comprobar que los modelos son capaces de reproducir las condiciones observadas de las rachas secas y húmedas.

En este sentido, en la Figura IV.14 se aprecia que los modelos regionales (RCM) reproducen adecuadamente las características climáticas de las rachas húmedas para el período de referencia 1961-2000: Número de días seguidos con lluvia (longitud de racha), precipitación máxima diaria (pico de la racha), precipitación acumulada de las rachas e índice n . Cabe destacar que en general la proyección para 2001-2040 no presenta cambios destacables respecto al período de referencia; sin embargo, se aprecia una ligera reducción de la longitud de las rachas húmedas en todos los trimestres exceptuando el de invierno. Esta reducción junto al ligero

aumento de los picos en otoño proyectados para algunos observatorios implica un leve incremento del índice n , como ya se apreciaba en la figura IV.12.

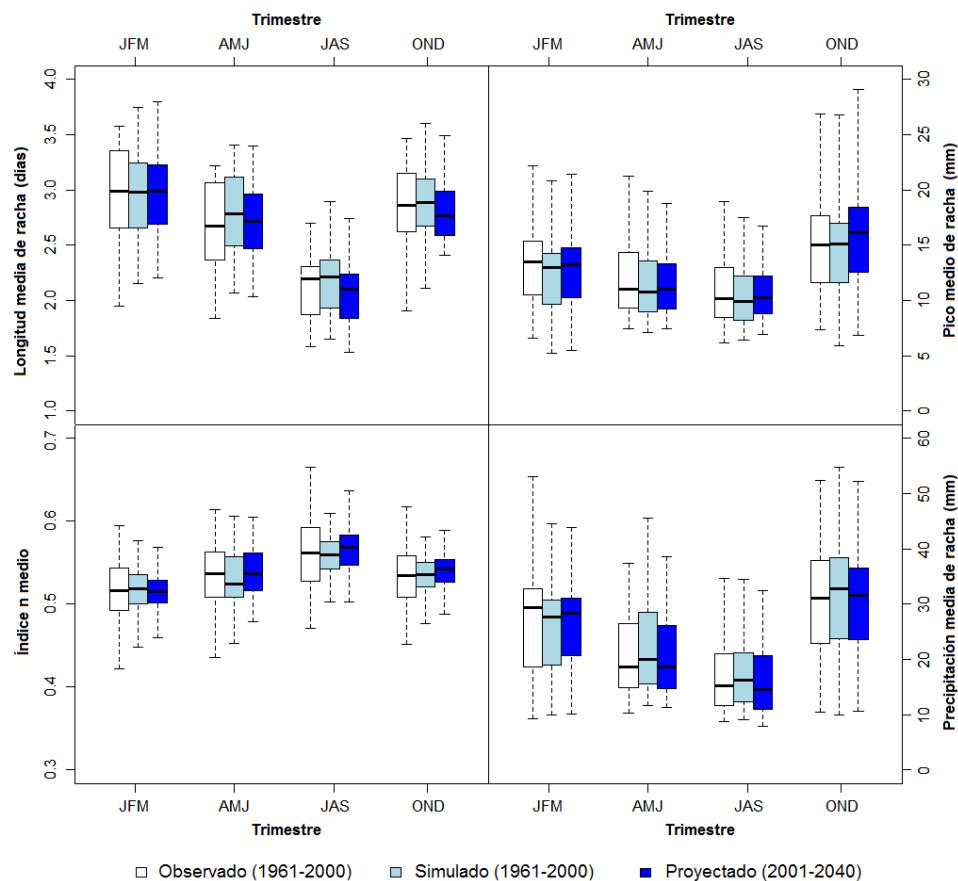


Figura IV.14. Comparación de las rachas húmedas observadas y simuladas para 1961-2000 y las proyectadas para 2001-2040. Cada punto del *boxplot* representa el valor climático de un observatorio real o simulado. Arriba a la izquierda, se representa la longitud media de las rachas (en días); arriba a la derecha, se representa la media del pico máximo diario de las rachas (en milímetros). Abajo a la derecha se representa la acumulación media de las rachas (en milímetros), y abajo a la izquierda se compara el índice n medio asociado a las rachas.

La reducción de las rachas húmedas es más pronunciada en verano y otoño. En cambio, la precipitación media de las rachas disminuye más claramente en verano que en otoño, aunque no para todos los observatorios. Esto se explica porque en los meses de verano la intensidad de la precipitación se mantiene aproximadamente constante (Fig. IV.14).

En cuanto a las rachas secas, la longitud media se reproduce adecuadamente para el período de referencia (Fig. IV.15). Sin embargo, las distribuciones de probabilidad de la duración de racha seca presentan ciertas diferencias respecto a las observaciones, aunque lógicamente

respetan el ciclo estacional. Para el período 2001-2040 se proyecta un claro aumento de la longitud de las rachas secas en primavera y verano de más de 0,5 días por racha. Nótese que el aumento de la duración de las rachas secas es coherente con la ligera disminución de la longitud de las rachas húmedas proyectadas para 2001-2040 (Fig. IV.14).

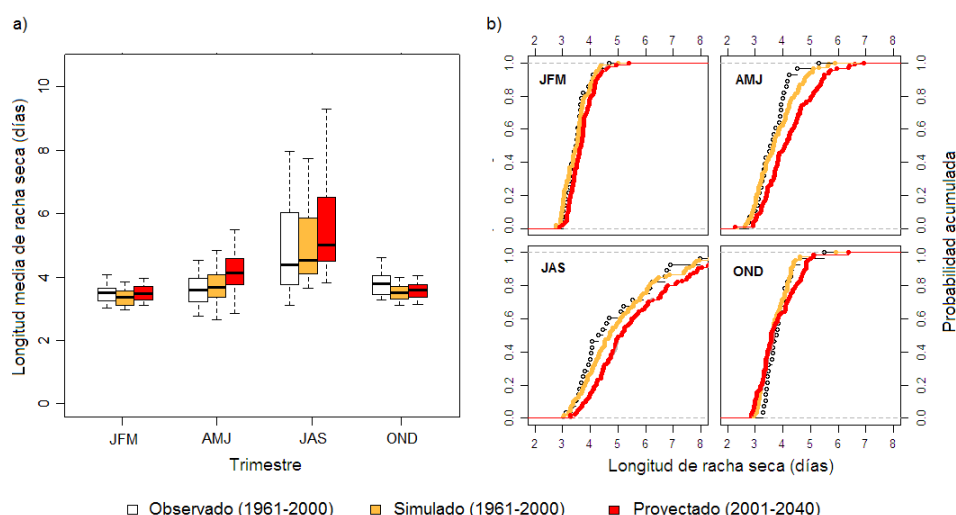


Figura IV.15. Evolución de las rachas secas para el período 2001-2040 en relación al período 1961-2000, tanto observado como simulado: a) longitud media de las rachas secas, y b) probabilidad acumulada de la longitud de rachas secas.

IV.5.3. Análisis del cambio de la precipitación extrema

El aumento de la precipitación intensa en el País Vasco puede estar relacionado a una posible mediterrización del clima. El aumento del número de días secos (Fig. IV.15) y la mayor concentración en poco tiempo de la precipitación (Fig. IV.12 y IV.14) apoyan esa hipótesis.

La posible mediterrización del clima del País Vasco puede estar supeditada a la expansión de la célula de Hadley (Lu *et al.*, 2007). El calentamiento global proyectado bajo el escenario A1B provoca una dilatación de dicha célula y un consiguiente ascenso hacia el norte de la zona de subsidencia subtropical. Las implicaciones físicas de dicho fenómeno se resumen básicamente en dos: 1) Menor precipitación en verano debido a la mayor presencia de anticiclones en niveles altos. 2) Mayor intensidad de la precipitación otoñal en áreas del mediterráneo debido al mayor calentamiento estival y por consiguiente mayor cantidad de agua disponible en la atmósfera.

IV.6. Conclusiones

IV.6.1. Sobre la metodología

Tanto la versión modificada de la distribución de Gumbel como la versión generalizada de la distribución Log-Logística se ajustan satisfactoriamente a la probabilidad empírica de la precipitación diaria en el País Vasco. Si bien, esta última distribución de probabilidad mostró mejores resultados, con un error medio del 4% frente al 5% obtenido por la distribución de Gumbel.

La transformación probabilística de los modelos climáticos corrige adecuadamente la distribución de probabilidad y la hace indistinguible de la de los observatorios en el período de referencia. De hecho, los errores de la serie simulada son en general inferiores al 5% tanto para la media, como para la desviación típica. El error en el número de días sin precipitación es inferior al 1% en la mayoría de los casos.

Además, dicha transformación no afecta directamente a la señal climática de los modelos, sino que se limita a sintonizarla coherentemente con la frecuencia probabilística de la precipitación de los observatorios. Por lo tanto, la metodología es útil para cuantificar los posibles cambios del clima a escala local; pero no sólo en referencia a la cantidad de precipitación, sino también en el modo en el que ésta se distribuye en el tiempo (rachas secas, húmedas, índice n).

IV.6.2. Sobre el cambio de la precipitación intensa

Los resultados de este trabajo desprenden que, a lo largo del siglo XXI y bajo el escenario A1B, cabe esperar cambios en la distribución temporal de la precipitación a favor de una intensificación. De hecho, aunque podría acortarse la duración de los días seguidos con lluvia, ésta se concentraría en un menor tiempo aumentando así la intensidad del máximo diario en otoño. Esta idea se ve reforzada porque el índice n proyectado para el resto del siglo XXI muestra un aumento ligero, pasando de 0,57 a 0,60.

En cuanto a la precipitación más extrema, la mayoría de modelos proyecta un claro aumento. Así por ejemplo, la precipitación con un período de retorno de 100 años aumentará hasta un 30% en 2001-2040 respecto a 1961-2000, en áreas del oeste de Vizcaya. Para el período 2041-2080, el aumento puede incluso superar el 40% en la mayor parte del territorio.

Además de la intensificación de la precipitación, que es más clara en otoño, los modelos climáticos proyectan un aumento de la duración de las rachas secas y una disminución ligera de la precipitación en verano. Por lo tanto, estos cambios señalan una posible mediterrización del clima del País Vasco.

CAPÍTULO V. CONCLUSIONES GENERALES

V.1. Ajuste a los climas de España

V.1.1. Sobre la metodología

Los dos modelos de probabilidad desarrollados en este trabajo se ajustan satisfactoriamente a todo el intervalo de precipitaciones en España. En concreto, el error normalizado medio (MNAE) es del 5% para el conjunto de observatorios (en promedio) e inferior a 10% en el 95% de los casos. Los tests de bondad de ajuste de Kolomorov-Smirnov y de Anderson-Darling indican que las distribuciones simuladas no son significativamente diferentes de las observadas ($p\text{-valor} > 0.05$). Por lo tanto, esta metodología desarrollada puede servir para mejorar la corrección de los errores sistemáticos de modelos climáticos y meteorológicos, así como para completar series temporales o estudiar inhomogeneidades.

V.1.2. Sobre las relaciones climáticas

Los nuevos modelos probabilísticos reflejan de forma coherente el ciclo estacional, por lo que también pueden ser utilizados para caracterizar los patrones climáticos de la precipitación. Si bien, el intervalo de error de los ajustes es mayor para los meses de verano, debido al escaso número de días con precipitación. Por esa razón, los nuevos modelos son especialmente aplicables para la mayor parte de los climas, exceptuando los más áridos.

V.2. Corrección de escenarios de precipitación del País Vasco

V.2.1. Sobre la metodología

Tanto la versión modificada de la distribución de Gumbel como la versión generalizada de la distribución Log-Logística se ajustan satisfactoriamente a la probabilidad empírica de la precipitación diaria en el País Vasco. Si bien, esta última distribución de probabilidad mostró mejores resultados, con un error medio del 4% frente al 5% obtenido por la distribución de Gumbel.

La transformación probabilística de los modelos climáticos corrige adecuadamente la distribución de probabilidad y la hace indistinguible de los observatorios en el período de referencia. Los errores de la serie simulada son en general inferiores al 5% tanto para la media, como para la desviación típica. El error en el número de días sin precipitación es inferior al 1% en la mayoría de los casos.

Además, dicha transformación no afecta directamente a la señal climática de los modelos, sino que se limita a sintonizarla coherentemente con la frecuencia probabilística de la precipitación de los observatorios. Por

lo tanto, la metodología es útil para cuantificar los posibles cambios del clima a escala local; pero no sólo en referencia a la cantidad de precipitación, sino también en el modo en el que ésta se distribuye en el tiempo (rachas secas, húmedas, índice n).

V.2.2. Sobre el cambio de la precipitación intensa

Los resultados de este trabajo desprenden que, a lo largo del siglo XXI y bajo el escenario A1B, cabe esperar cambios en la distribución temporal de la precipitación a favor de una intensificación. De hecho, aunque podría acortarse la duración de los días seguidos con lluvia, ésta se concentraría en un menor tiempo aumentando así la intensidad del máximo diario en otoño. Esta idea se ve reforzada porque el índice n proyectado para el resto del siglo XXI muestra un aumento ligero, pasando de 0,57 a 0,60.

En cuanto a la precipitación más extrema, la mayoría de modelos proyecta un claro aumento. Así por ejemplo, la precipitación con un período de retorno de 100 años aumentará hasta un 30% en 2001-2040 respecto a 1961-2000, en áreas del oeste de Vizcaya. Para el período 2041-2080, el aumento puede incluso superar el 40% en la mayor parte del territorio.

Además de la intensificación de la precipitación, que es más clara en otoño, los modelos climáticos proyectan un aumento de la duración de las rachas secas y una disminución ligera de la precipitación en verano. Por lo tanto, estos cambios señalan una posible mediterrización del clima del País Vasco.

V.3. Sobre el índice n

En este trabajo se ha reforzado matemáticamente el concepto de la intensidad de la precipitación en un sentido más amplio. El índice n ha servido para unificar criterios a la hora de analizar el tipo de precipitación y sus posibles cambios.

Teniendo en cuenta que la relación entre la escala diaria y subdiaria puede aproximarse a una función potencial con exponente n , cabe esperar que las mismas distribuciones de probabilidad puedan emplearse para diferentes escalas de tiempos. Por lo tanto, se abren nuevas posibilidades en el estudio sobre las estructuras temporales y probabilísticas de las rachas de días lluviosos.

GLOSARIO DE TÉRMINOS Y ACRÓNIMOS

A1. Escenario de Rápido Crecimiento Global. Los escenarios de tipo A1 se refieren a un contexto de crecimiento económico global muy rápido. La población mundial alcanza su nivel más alto a mitad del siglo XXI y disminuye posteriormente, produciéndose una rápida introducción de nuevas tecnologías más eficaces. Se supone una convergencia entre las regiones respecto a los ingresos *per cápita*, y una mayor interacción entre las sociedades. La familia de escenarios A1 se divide en tres grupos según las distintas direcciones de la evolución del sistema energético: fuentes de energía intensivas de origen fósil (A1FI), de origen no fósil (A1T) o un equilibrio entre todas las fuentes (A1B). Este grupo de escenarios contrasta con las familias del A2, B1 y B2, que presentan un crecimiento más lento.

A1B. Escenario de Rápido Crecimiento Global y de Equilibrio Energético. El escenario A1B pertenece a la familia de escenarios A1, y se caracteriza por combinar tanto energías de origen fósil como de origen no fósil, de tal modo de que no existe una dependencia excesiva por una fuente concreta.

A2. Escenario de Crecimiento Regional. Los escenarios de tipo A2 se refiere a un contexto de desarrollo heterogéneo. Es decir, se caracteriza por la preservación de las identidades locales, los modelos propios de desarrollo y la autosuficiencia. La convergencia entre las diferentes regiones es muy lenta, lo cual provoca un aumento continuo de la población mundial.

AEMET. Agencia Estatal de Meteorología. Antiguo Instituto Nacional de Meteorología, actualmente se encuentra adscrito a la Secretaría de Estado de Cambio Climático dentro del Ministerio de Medio Ambiente, y Medio Rural y Marino.

B1. Escenarios de Crecimiento Sostenible Global. Los escenarios de tipo B1 describen un contexto de desarrollo convergente, que alcanza su nivel más alto a mediados del siglo XXI para disminuir posteriormente, de un modo similar a los escenarios A1 pero con cambios más rápidos en las estructuras económicas, hacia un desarrollo más sostenible. Se incentiva la economía de la información, de los servicios y la introducción de tecnologías limpias. Sin embargo no se desarrollan nuevas iniciativas políticas respecto al clima.

B2. Escenario de Crecimiento Sostenible Regional. Los escenarios de tipo B2 se corresponden con un crecimiento basado en las soluciones locales de sostenibilidad económica, social y ambiental. En este caso la población crece constantemente pero a un ritmo menor que en los escenarios de tipo A2, con cambios tecnológicos menos rápidos pero más variados que en los A1 y B1.

BIAS. *Sesgo medio.* Es la diferencia media entre los N valores simulados o predichos (P_i) y los N valores observados (O_i) de una variable:

$$BIAS = \frac{\sum_{i=1}^N (P_i - O_i)}{N}$$

Bootstrap. Método especial del test de Kolmogorov-Smirnov para estimar un p-valor (o significancia) más preciso, basado en bucles (Marsaglia et al. 2003, Sekhon, 2010).

Célula de Hadley. Es una corriente atmosférica, perteneciente a la circulación global, que presenta fuertes ascensos de aire en la zona ecuatorial (cinturón de depresiones) y fuertes descensos de aire en la zona subtropical (cinturón de anticiclones), situada entorno a los grandes desiertos.

CNRM. Centre National de Recherches Météorologiques. Centro Nacional de Estudios Meteorológicos (Francia).

CNRM-RM4.5. Modelo regional climático (RCM) elaborado por el CNRM.

Downscaling. Regionalización de un campo numérico, por ejemplo de un modelo meteorológico o climatológico. La regionalización puede ser **dinámica**, mediante simulaciones físicas que incluyen la dinámica atmosférica, o **estadística**, mediante conexiones físico-estadísticas entre variables predictoras y variables predichas.

ECDF. Empirical Cumulative Distribution Function. Curva de distribución de probabilidad acumulada empírica.

ECMWF. *European Centre for Medium-Range Weather Forecasts.* Centro Europeo de Predicción a Medio Plazo.

Ensemble. Conjunto o predicción por conjuntos. Este término se usa habitualmente en predicción meteorológica y climatológica para referirse a la combinación de diferentes modelos para disminuir las incertidumbres de las predicciones.

ENSEMBLES. Proyecto financiado por la Comisión Europea que pretende desarrollar un sistema de predicción por conjuntos para el estudio del cambio climático basado en las últimas técnicas de modelización de alta resolución, de ámbito mundial y regional. <http://www.ensembles-eu.org/> (Niehörster *et al.*, 2008)

ENSEMBLES-RT3. Subproyecto de ENSEMBLES conocido como “Research Theme 3 (RT3)” o bien “Formulation of very high resolution Regional Climate Model Ensemble for Europe”. Dicho subproyecto pretende proporcionar mejores herramientas para la modelización climática en alta resolución a escalas espaciales de 50 km hasta 10 km <http://ensemblesrt3.dmi.dk/> (Niehörster *et al.*, 2008).

ERA40. Reanálisis de las condiciones atmosféricas y de superficie globales, desde septiembre de 1957 hasta agosto de 2002, llevado a cabo por el ECMWF, a partir de datos observados de radiosondas, globos, aviones, boyas y satélites, entre otros.

GEI. Gas de Efecto Invernadero. Cualquier gas atmosférico que provoca una ganancia de energía respecto a si no lo hubiese. Ejemplos: Vapor de agua, CO₂, metano, CFCs.

Hadley. Ver **célula de Hadley**.

Hietograma. Gráfica que relaciona la medición de la precipitación con los intervalos de tiempo con los que se mide. En la mayoría de los casos, el intervalo de tiempo es de anchura constante ya que se corresponde con la frecuencia con la que el instrumento registra la precipitación.

Hietograma sintético o de diseño. Es un hietograma teórico de un evento de precipitación, con una cierta distribución temporal (p.e., índice n) y con un determinado período de retorno.

IDF. Intensidad-Duración-Frecuencia. Curva que resulta de conectar los puntos representativos de la intensidad media de la precipitación para diferentes intervalos de duración pero correspondientes a un mismo período de retorno (Témez, 1978).

IPCC. *Intergovernmental Panel of Climate Change*. Organismo multinacional encargado de coordinar la discusión científica sobre el Cambio Climático, y de negociar la política de medidas de mitigación y adaptación respecto al mismo. <http://www.ipcc.ch/>

KNMI. *Koninklijk Nederlands Meteorologisch Instituut*. Real Instituto Meteorológico de los Países Bajos.

KNMI-RAKMO2. Modelo regional climático (RCM) elaborado por el KNMI.

METNO. *Norwegian Meteorologisk institutt*. Instituto Meteorológico de Noruega.

METNO-HIRHAM. Modelo regional climático (RCM) elaborado por el METNO.

N. Índice de la tipología de la precipitación, según la regularidad temporal de la intensidad media máxima (Moncho *et al.*, 2009a, 2011).

M. Exponente que relaciona la precipitación con los períodos de retorno, usando una distribución de probabilidad de Pareto (Moncho *et al.*, 2011).

MAE. *Mean Absolute Error*. Se calcula con la media de las diferencias, en valor absoluto, de los N valores simulados (P_i) y los N valores observados (O_i):

$$MAE = \frac{\sum_{i=1}^N |P_i - O_i|}{N}$$

MNAE. *Mean Normalized Absolute Error.* Se calcula con la media de las diferencias, en valor absoluto, de los N valores simulados (P_i) y los N valores observados (O_i), dividiendo por cada valor observado (O_i) (Brière *et al.*, 2007):

$$MNAE = \frac{1}{N} \sum_{i=1}^N \left| \frac{P_i - O_i}{O_i} \right|$$

Período de retorno. Es el tiempo que cabe esperar entre dos eventos de la misma intensidad, así como también es el tiempo que cabe esperar entre ahora y un evento con una determinada intensidad. Para la precipitación, se suele tomar como variable a la cantidad de lluvia registrada en un día, o la cantidad máxima diaria en un año. El período de retorno empírico de una intensidad X_o se calcula mediante la inversa de la probabilidad acumulada de que se registre un evento con una intensidad X igual o superior a X_o .

P-valor. Valor de significancia. Para el caso de los tests de indistinguibilidad (por ejemplo Kolomarov-Smirnov y Anderson-Darling), un p-valor igual a 0,05 hace referencia a un rango de confianza del 95% de que dos poblaciones de datos sean diferentes, por lo tanto, un p-valor superior a 0,05 significa que dos poblaciones son estadísticamente indistinguibles (ya que la confianza de que sean diferentes es inferior al 95%). Si dos poblaciones de datos son iguales, entonces se obtendrá un p-valor = 1.

Reanálisis. Simulación de las condiciones atmosféricas pasadas mediante un modelo de circulación global, a partir de diferentes fuentes de datos de entrada (por ej. estaciones, radiosondas, boyas, barcos). Los más utilizados son el ERA40 y el NCEP/NCAR (llevados a cabo por el *National Centers for Environmental Prediction* y el *National Center for Atmospheric Research*).

RPS. *Return Period Series.* Serie de los períodos de retorno obtenida a partir de la probabilidad asociada a cada día de precipitación.

RBIAS. *Relative bias.* Se calcula con la suma de las diferencias de los N valores simulados (P_i) y los N valores observados (O_i), dividiendo finalmente entre la suma de los valores observados (Yu *et al.*, 2006):

$$RBIAS = \frac{\sum_{i=1}^N (P_i - O_i)}{\sum_{i=1}^N O_i}$$

RMAE. *Relative Mean Absolute Error.* Se calcula con la suma de las diferencias, en valor absoluto, de los N valores simulados (P_i) y los N valores

observados (O_i), dividiendo finalmente entre la suma de los valores observados (Yu *et al.*, 2006):

$$RMAE = \frac{\sum_{i=1}^N |P_i - O_i|}{\sum_{i=1}^N O_i}$$

SRES. *Special Report on Emissions Scenarios*. Informe de escenarios de emisiones realizado por el IPCC para el TAR, tercer informe de evaluación del Cambio Climático (Cubasch *et al.* 2001) así como también en el AR4, cuarto informe de evaluación (Meehl *et al.*, 2007). En el SRES se describen las diferentes familias de escenarios futuros (A1, A2, B1 y B2).

UCLM. Universidad de Castilla–La Mancha (España).

UCLM-PROMES. Modelo regional climático (RCM) elaborado por la UCLM.

BIBLIOGRAFÍA

- AEMET (2011): "Interpretación: predicción por provincias". [www.aemet.es
http://www.aemet.es/documentos/es/eltiempo/prediccion/comun/prediccion_interpretacion.pdf](http://www.aemet.es/documentos/es/eltiempo/prediccion/comun/prediccion_interpretacion.pdf)
- ABRAMOWITZ, M. Y STEGUN, I. A. (1972): *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover.
- AKAIKE, H. (1974): "A new look at the statistical model identification". *IEEE Trans. Automat. Cont.*, 19: 716-723.
- ALEKLETT, K.; HOOK, M.; JAKOBSSON, K.; LARDELLI, M.; SNOWDEN, S.; SODERBERGH, B. (2010): "The Peak of the Oil Age - Analyzing the world oil production Reference Scenario in World Energy Outlook 2008". *Energy Policy* 38: 1398-1414.
- ALVES, M., y VERDIÈRE, A. C. (1999): "Instability dynamics of a subtropical jet applications to the Azores Front-Current System-eddy driven mean flow". *J. Phys. Oceanogr.*, 29: 837-864.
- BARRY, C. A. (1983): *Pareto Distributions*. International Co-operative Publishing House, Burtonsville, Maryland.
- BEGUERIA, S. (2005): "Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value". *J. Hydrol.*, 303: 215-230.
- BENESTAD, R. E. (2010): "Downscaling precipitation extremes. Correction of analog models through PDF predictions". *Theor. Appl. Climatol.*, 100: 1-21.
- BEN-ZVI, A. (2009) Rainfall intensity-duration-frequency relationships derived from large partial duration series. *J. Hidrol.*, 367: 104-114.
- BORJA, A.; AGUIRREZABALAGA, F.; MARTÍNEZ, J.; SOLA, J.C.; GARCÍA-ARBERAS, L., y GOROSTIAGA, J.M., (2004): "Benthic communities, biogeography and resources Management". In: Borja, A., Collins, M. (Eds.), *Oceanography and Marine Environment of the Basque Country*. Elsevier Oceanography Series 70, pp. 455-492.
- BRIÈRE, C.; ABADIE, S.; BRETEL, P.; LANG, P. (2007). "Assessment of TELEMAC system performances, a hydrodynamic case study of Anglet, France". *Coastal Engineering*, 54: 345-356.
- BURNHAM, K. P.; ANDERSON, D. (2002): *Model Selection and Multi-Model Inference*. Springer-Verlag, NY, ISBN-10: 0387953647.
- CAPEL-MOLINA, J. J. (2000): *El Clima de la Península Ibérica*. Ed. Ariel, Barcelona.
- CARLSON, W. B. (2011): "The Modeling of World Oil Production Using Sigmoidal Functions". *Energy Sources, Part B: Economics, Planning, and Policy* 6: 178-186.

- CERVENY, R.S.; LAWRIEMORE, J.; EDWARDS, R., y LANDSEA, C. (2007): "Extreme Weather Records: Compilation, Adjudication and Publication". *Bull. Amer. Meteor. Soc.*, 88: 853-860.
- CHUST, G., A. CABALLERO, M. MARCOS, P. LIRIA, C. HERNÁNDEZ, AND Á. BORJA (2010). Regional scenarios of sea level rise and impacts on Basque (Bay of Biscay) coastal habitats, throughout the 21st century. *Estuarine, Coastal and Shelf Science* 87:113-124.
- CHUST, G.; BORJA, A.; CABALLERO, A.; IRIGOIEN, X.; SÁENZ, J.; MONCHO, R.; MARCOS, M.; LIRIA, P.; HIDALGO, J.; VALLE, M.; VALENCIA, V. "Climate Change impacts on the coastal and pelagic environments in the southeastern Bay of Biscay". *Clim. Res.*, En Prensa, Ms. No. 201004021.
- CHRISTENSEN, J.H.; HEWITSON, B.; BUSUIOC, A.; CHEN, A.; GAO, X.; HELD, I.; JONES, R.; KOLLI, R.K.; KWON, W.-T.; LAPRISE, R.; MAGAÑA RUEDA, V.; MEARN, L.; MENÉNDEZ, C.G.; RÄISÄINEN, J.; RINKE, A.; SARR, A.; WHETTON, Y.P. (2007): *Regional Climate Projections. In: Climate Change: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [SOLOMON, S.; QIN, D.; MANNING, M.; CHEN, Z.; MARQUIS, M.; AVERYT, K.B.; TIGNOR, M.; MILLER, H.L. (eds.)]. Cambridge University Press, Cambridge, United Kingdom y New York, NY, USA.
- CUBASCH, U.; MEEHL, G. A.; BOER, G. J.; STOUFFER, R. J.; DIX, M.; NODA, A.; SENIOR, C. A.; RAPER, S. y YAP, K. S. (2001): "Projections of future climate change". In: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* [HOUGHTON J.T., DING Y., GRIGGS, D. J.; NOUGUER, M.; van der LINDEN, P.; DAI, X.; MASKELL, K.; JOHNSON, C. I. (eds.)]. Cambridge University Press, ISBN 0521 01495 6.
- DÉQUÉ, M.; ROWELL, D. P.; LÜTHI, D.; GIORGI, F.; CHRISTENSEN, J. H.; ROCKEL, B.; JACOB, D.; KJELLSTRÖM, E.; DE CASTRO, M. y VAN DEN HURK, B. (2007): "An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections". *Clim. Change*, 81: 53-70.
- DURMAN, C. F.; GREGORY, J. M.; HASSELL, D. C.; JONES, R. G., y MURPHY, J. M., (2006): "A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates". *Quart. J. Roy. Meteor. Soc.*, 127: 1005–1015.
- ELSHAMY, M. E.; SEIERSTAD, I. A., y SORTEBERG, A. (2009): "Impacts of climate change on Blue Nile flows using bias-corrected GCM scenarios". *Hydrol. Earth Syst. Sci.*, 13: 551–565.

- ESLAMIAN, S.S y FEIZI, H. (2006): "Maximum monthly rainfall analysis using L moments for and arid region is Isfahan Province", *Iran. J Appl Met. Clim.*, 46: 494-503.
- FERNÁNDEZ-FERRERO, A.; SÁENZ, J.; IBARRA-BERASTEGI, G., Y FERNÁNDEZ J. (2009): "Evaluation of statistical downscaling in short range precipitation forecasting". *Atmospheric Research*, 94: 448-461
- FURRER, R.; NYCHKA, D., y SAIN, S. (2010). "fields: Tools for spatial data". R package version 6.3. <http://CRAN.R-project.org/package=fields>
- GARMENDIA, M. I.; PÉREZ, C.; RODRÍGUEZ, C., y GARMENDIA, J. (1989): "Factores determinantes de la precipitación anual en la vertiente cantábrica". *Acta Samaltina de Ciencias*, 67: 113-117, 3 Ref.
- GOUBANOVA, K. y LI, L. (2007): "Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations". *Global and Planetary Change*, 57 (2007), 27-42.
- GRANGER, R. J. (1989): "An examination of the concept of potential evaporation". *J. Hydrol.*, 111: 9-19.
- GREEN, P.J., y SILVERMAN, B. W. (1994): *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall, London.
- HANSON, L. S. y VOGEL, R. (2008): *The probability distribution of daily rainfall in the United States*. World Environmental and Water Resources Congress 2008, Ahupua A: 1-10. doi 10.1061/40976(316)585.
- HARDY, M. (2010): "Pareto's Law". *Math. Intel.*, 32: 38-43.
- HOGG, R. VI. y CRAIG. A. T. (1978): *Introduction to Mathematical Statistics*, 4th edition. Macmillan, New York.
- INES, A.V.M, y HANSEN, J. W. (2006): "Bias correction of daily GCM rainfall for crop simulation studies". *Agr. For. Met.*, 138: 44-53.
- JEEVANANDA, S.R. (1995): "Sensitivity of some potential evapotranspiration estimation methods to climate change". *Agric. Forest. Meteorol.*, 77: 121-125.
- JOHNSON, N. L.; KOTZ, S., y BALAKRISHNAN, N. (1994): "Continuous univariate distributions". Wiley Series in John Wiley & Sons: *Probability and Mathematical Statistics: Applied Probability and Statistics*. New York. MR1299979, ISBN 978-0-471-58495-7.
- KHON, VI. Ch.; MOKHOV, I.I.; ROECKNER, E., y SEMENOV, VI.A. (2007): "Regional changes of precipitation characteristics in Northern Eurasia from simulations with global climate model". *Global and Planetary Change*, Vol. 57, nº 1-2 (May 2007), 118-123.
- KIDSON, J.W. y THOMPSON, C.S. (1998): "A comparison of statistical and model-based downscaling techniques for estimating local climate variations". *J. Climate*, 11: 735-753.

- KYSELY, J. (2002): "Comparison of extremes in GCM-simulated, downscaled and observed central-European temperature series". *Clim. Res.*, 20: 211–222.
- LANA, X.; MARTÍNEZ, M.D.; SERRA, C., y BURGUEÑO, A. (2004): "Spatial and temporal variability of the daily rainfall regime in Catalonia (northeastern Spain), 1950–2000". *Int. J. Clim.*, 24: 613–641.
- LU, J.; VECCHI, G.A., y REICHLER, T. (2007): "Expansion of the hadley cell under global warming". *Geophysical Research Letters*, 34: L06805.
- MARSAGLIA, G.; TSANG, W.W. y WANG J. (2003): "Evaluating Kolmogorov's distribution". *Journal of Statistical Software*, 8/18.
- MARTÍN-VIDE, J. (2004): "Spatial distribution of a daily precipitation concentration index in peninsular Spain". *Int. J. Clim.*, 24: 959–971.
- MARKS, D. A.; WOLFF, D. B.; CAREY, L. D., y TOKAY, A. (2010): "Quality control and calibration of the dual-Polarization radar at Kwajalein, RMI". *J. Atmos. Oceanic. Technol.* (in pres.).
- MCGUINNESS, J. L. y LESLIE, H. P. (1972): "Maximum potential evapotranspiration frequency". *J. Irrigat. Drain. Div.*, 2: 207–214.
- MEEHL, G.A.; STOCKER, T. F.; COLLIN, W.D.; FRIEDLINGSTEIN, P.; GAYE, A. T.; GREGORY, J. M.; KITOH, A.; KNUTTI, R.; MURPHY, J. M.; NODA, A.; RAPER, S. C. B.; WATTERSON, I.G.; WEAVER, A. J. y ZHAO, Z.-C. (2007): Global Climate Projections. In: "*Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate*".
- MONCHO, R.; BELDA, F. y CASELLES, V. (2011a): "Distribución probabilística de los extremos globales de precipitación", *Nimbus*, (en revisión).
- MONCHO, R.; BELDA, F. y CASELLES, V. (2009a): "Climatic study of the exponent n of the IDF curves of the Iberian Peninsula", *Tethys*, 6: 3–14.
- MONCHO, R., CHUST, G. y CASELLES, V. (2009b): "Análisis de la precipitación del País Vasco en el período 1961–2000 mediante reconstrucción espacial". *Nimbus*, nº 23–24: 149–170.
- MONCHO, R. (2009): *Validación y calibración de los modelos climáticos regionales para el estudio de la precipitación extrema en el País Vasco*. Memoria del Diploma de Estudios Avanzados. Facultat de Física, Universitat de València, Burjassot.
- MUENCHEN, R. A., y HILBE, J. M. (2010): *R for Stata Users. Statistics and computing*, Springer, ISBN: 978-1-4419-1317-3.
- NIEHÖRSTER, F.; FAST, I.; HUEBENER, H., y CUBASCH, U. (2008): "The stream one ENSEMBLES projections of future climate change". *Ensembles Technical Report nº 3*, 2008.
- OLCINA, J. (2008): "Riesgo de inundaciones y ordenación del territorio en España". *Documents d'Anàlisi Geogràfica*, 51: 178–180.

- ÖZTÜRK, A. (1981): "On the study of a probability distribution for precipitation Totals". *J. Appl. Met.*, 20: 1499-1505.
- PALMER, T. N.; DOBLAS-REYES, F. J.; WEISHEIMER, A., y RODWELL, M. J. (2008): "Toward seamless prediction: Calibration of climate change projections using seasonal forecasts". *Bull. Am. Met. Soc.*, 89: 459-470.
- PARK, J. S., y JUNG H. S. (2002): "Modelling Korean extreme rainfall using a Kappa distribution and maximum likelihood estimate". *Theor. Appl. Climatol.*, 72: 55-64.
- PELIZ A.; ROSA T.L.; SANTOS A.M.P., y PISSARRA J.L. (2002): "Fronts, jets, and counterflows in the Western Iberian upwelling system". *J. Mar. Syst.*, 35: 61-77.
- PERKINS, S. E.; PITMAN, A. J.; HOLBROOK, N. J., y MCANENEY, J. (2007): "Evaluation of the ar4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions". *J. Climate* 20: 4356-4376.
- PIANI C.; HAERTER, J. O., y COPPOLA, E. (2010): "Statistical bias correction for daily precipitation in regional climate models over Europe". *Theor. Appl. Climatol.*, 99:187-192.
- RAUE, A.; KREUTZ, C.; MAIWALD, T.; BACHMANN, J.; SCHILLING, M.; KLINGMÜLLER, U., y TIMMER, J. (2009): "Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood". *Bioinformatics*, 25: 1923-9.
- R DEVELOPMENT CORE TEAM (2010): "R: A language and environment for statistical computing". *R Foundation for Statistical Computing*, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- SEKHON, J. S. (2010): "Matching: Multivariate and Propensity Score Matching with Balance Optimization". *R package version 4.7-11*. URL <http://CRAN.R-project.org/package=Matching>.
- SCM (2011): "Divulgació meteorològica: Entendre les prediccions". Servei Meteorològic de Catalunya. <http://www.meteo.cat/servmet/index.html>
- SCHOLZ, F.W., y STEPHENS, M.A. (1987): "K-sample Anderson-Darling Tests". *Journal of the American Statistical Association*, 399: 918-924.
- TÉMEZ, J. R. (1978): *Cálculo Hidrometeorológico de caudales máximos en pequeñas cuencas naturales*. Dirección General de Carreteras. Madrid.
- TERINK, W.; HURKMANS, R.T.W.L.; TORFS, P.J.J.F., y UIJLENHOET, R. (2010): "Evaluation of a bias correction method applied to downscaled precipitation and temperature reanalysis data for the Rhine basin". *Hydrol. Earth Syst. Sci.*, 14: 687-703.
- WIDMANN, M.; BRETHERTON, C. S., y SALATHÉ, E. P. (2003): "Statistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor". *J. Climate.*, 16: 799-816.

WILSON, P. S.; TOUMI, R. (2005): "A fundamental probability distribution for heavy rainfall". *Geophys. Res. Lett.*, 32: L14812.

YU, S.; EDER, B.; DENNIS, R.; CHU, S.-H., y SCHWARTZ, S.E. (2006): "New unbiased symmetric metrics for evaluation of air quality models". *Atmospheric Science Letters*, 7: 26–34.

ANEXO 1. Alternative model for precipitation probability distribution

Climate Research, 51: 23-23

(Aceptado en 2011, publicado en 2012)

*Revista con un factor de impacto de 2,25 y que
ocupa la posición 18 de un total de 70 revistas
en el tema de “meteorología y ciencias atmosféricas”*

Alternative model for precipitation probability distribution

Roberto Moncho^{*,1,2}, Vicente Caselles², Guillem Chust³

¹Fundación para la Investigación del Clima. C/ Gran Vía, N. 22, 6th Izda., Madrid, 28013 (Spain)

²Departament de Física de la Terra i Termodinàmica, Universitat de València. C/ Dr. Moliner, N. 50, 46100 Burjassot (Valencia, Spain)

³AZTI-Tecnalia, Marine Research Division. Txatxarramendi ugarte z/g, 48395 Sukarrieta (Vizcaya, Spain)

*Email: robert@temps.cat

ABSTRACT

In climatology, there is difficulty in describing the probability distribution of rainfall because there are many days without precipitation, which usually causes the most likely daily rainfall to be zero. None of the widely used models is able to adjust the overall variation of daily precipitation. Therefore, this article proposes an alternative model for the climate probability of precipitation. This model, based on four parameters, has been applied to daily rainfall throughout all the months of the year and for 108 observatories in Spain. This model provides better results than the commonly used probability models (Generalized Extreme Value, Pareto and Generalized Pareto Distribution, Gamma, Gumbel, Weibull, Exponential and Log-normal). In particular, it had a mean absolute error of < 10% in most of the stations analyzed. Thus, this alternative model could be used to correct the probability distribution of daily precipitation for weather forecasting and climate models.

Key words: Probability distribution, return period, precipitation, Spain

1. Introduction

The prediction of possible climate change in precipitation is an issue of current interest (Kharin & Zwiers 2000), especially to detect changes in extreme rainfall and drought for agriculture and environment. However, climate models present problems to adequately reproduce the probability distribution of precipitation (Ebert 2001, Wilson & Toumi 2005, Perkins et al. 2007). Probably, the problem is intrinsic to the irregular nature of rainfall.

In most of the weather stations, the maximum probability of daily rainfall is zero; thus, the climate probability curve decreases from low to high rainfall. Rainfall does not behave as a normal distribution because precipitation has no negative values. Moreover, even if we try to complete the domain of the function of probability with evaporation as a mode of negative precipitation, distribution would still be asymmetric. This is because the potential evaporation rarely reaches 20 mm in one day in most of the observatories (McGuinness & Leslie 1972, Granger 1989, Jeevananda 1995), whereas precipitation > 50 mm in one day are recorded more frequently in some countries (Cerveny et al. 2007). In particular, extreme precipitation is frequent in the Mediterranean climate, when compared with the annual precipitation. This causes difficulties in fitting the probability distribution to daily precipitation.

Possibly for that reason, the commonly used models (Generalized Extreme Value, Generalized Pareto Distribution, Gamma, Gumbel, Weibull) do not fit the overall variation of daily precipitation (very low and very high precipitation simultaneously) (Bridges & Haan 1972, Öztürk 1981, Etoh et al. 1986, Begueria 2005). All these models use two or three parameters; hence, it might be necessary to use probability models with more parameters (Park & Jung 2002, Hanson & Vogel 2008).

The objective of this paper is to present an alternative model for probability distribution of daily precipitation (including days without rain and with rainfall of < 0.1 mm). The new model aims to characterize and correct the probability distribution of a generic set of daily precipitation data. This is an important requirement for improving the systematic correction of dynamic and statistical downscaling (Kysely 2000). This model should also improve the calibration of operational weather models and other tools of prediction and analysis of precipitation, such as weather radars.

2. Materials

2.1. Study area

The study area was Spain, including the two archipelagos (Fig. 1), which was the pilot area for the first analysis of the methodology developed. The prevailing climate in Spain is the Mediterranean, although other climatic conditions can be found regionally, such as the oceanic climate area (in the north of the Iberian Peninsula), mountain climate (in Sierra Nevada and

Pyrenees), and dry subtropical climate (in the Canary Islands). Spain has a varied climate due to its geographical location and characteristics of the territory. The Iberian Peninsula is located within the influence of the polar jet waves, or Rossby waves, and it is also occasionally affected by the subtropical jet (Alves & Verdière 1999, Peliz et al. 2002).

Therefore, the climatic conditions are heterogeneous; e.g., the annual accumulation of rainfall is between 150 mm in the southeast of Andalusia and 2500 mm in the northeast of the Basque Country, with oceanic climate (Capel-Molina 2000). This latter climate is characterized by regular precipitations throughout the year (in some cases, with >50% of the days of rain). On the other hand, precipitation of the Mediterranean climate is characterized by dry summers and heavy rainfall concentrated in a few days (Martin-Vide 2004, Lana et al. 2004), especially in autumn and spring. Throughout the year, the number of days with precipitation usually ranges between 10 and 30% on an average, although it can attain values <10% in July, and in some case, >50% in April (Capel-Molina, 2000). Therefore, the different seasons, to some extent, represent different climates (e.g. Mediterranean summer climate is similar to dry climates and spring to humid climates). For these reasons, Spain is a suitable pilot area to test the methodology developed in this work.

2.2. Precipitation data

The precipitation data used for this study correspond to 108 observatories with daily resolution from the State Meteorological Agency (AEMET) of Spain (Fig. 1). The selected time series satisfy a length of at least 20 years and 90% of days with records. Among these observatories, 62 had a time series with more than 50 years and 98% of days with records. The data showed a high frequency of negligible precipitation (i.e. between 0 and 0.1 mm); in particular, they were between 0.6 and 16% of days with < 0.1 mm. These values changed to 0, and hence, the study took into account that the zeros actually represent the class of [0, 0.1) mm. In addition, some erroneous daily values (accumulated for several days), being <0.1%, were found in the series analyzed; however, they were not outliers, and hence, did not affect the statistics. In these cases, we divided them by the number of daily accumulated precipitation.

3. Methodology

The frequency of daily rainfall for a certain group of stations or grid points of a model can be analyzed with the mean or maximum precipitation of the group. For this study, the probability for each observatory was analyzed separately. The cumulative probability $\pi(p \geq P)$ is defined as the probability that a station or group of stations registers a precipitation $\geq P$ (on a day of the year or of a particular month). The occurrence frequency for each threshold of rainfall, P , can be estimated by counting the records of

daily rainfall of the station or a set of stations, and the cumulative probability $\pi(p \geq P)$ is estimated with the normalized sum of those frequencies. The return period T or expected time (in days) between two precipitation events with $p \geq P$ is given by the inverse of the cumulative probability:

$$T(P) \equiv \frac{1}{\pi(p \geq P)} \quad (1)$$

To fit the probability of daily precipitation (of full year or a particular month), first we tested the commonly used models: Generalized Extreme Value (GEV), Pareto and Generalized Pareto Distribution (GPD), Gamma, Gumbel, Weibull, Exponential and Log-normal II and III (Bridges & Haan 1972, Öztürk 1981, Etoh et al. 1986, Begueria 2005). Second, other models were also tested, such as the modified version of GEV:

$$\pi(\lambda_0 \geq \lambda) = \exp \left[- (1 + \varepsilon \lambda)^{\frac{1}{\varepsilon}} \right] \quad (2)$$

and the modified versions of Gumbel model:

$$\pi(\lambda_0 \geq \lambda) = \exp \left[- \exp(\lambda^w + k) \right] \quad (3)$$

where ε , w , and k are adjustable parameters and λ is the relative precipitation, which depends on an adjustable value P_o for the most probable precipitation as well as an adjustable value of normalization, P_1 :

$$\lambda \equiv \frac{P - P_o}{P_1} \quad (4)$$

Therefore, the cumulative probability of the occurrence of precipitation $p \geq P$ satisfies the same probability distribution (Equation (2)), i.e., $\pi(p \geq P) = \pi(\lambda_0 \geq \lambda)$. Alternatively, we propose another function of accumulated probability $\pi(\lambda)$, according to:

$$\pi(\lambda_0 \geq \lambda) \approx \frac{1}{1 + \lambda^{w + \lambda^w e^{-k}}} \quad (5)$$

where w and k are the positive parameters (see Annex 1), and λ is the relative precipitation, defined positive. It must be noted that we can make a change of variables as $r \equiv \exp(-k)$, although we prefer to use k , because it is a parameter that varies more smoothly than r . Hence, four adjustable parameters for the return period (with Equations (1), (4), and (5)) have been used:

$$T \approx 1 + \left(\frac{P - P_o}{P_1} \right)^{w + e^{-k} \left(\frac{P - P_o}{P_1} \right)^w} \quad (6)$$

where the parameter w is the “shape” k is the “smoothness” P_1 is the “scale”, and P_o is the “location” The parameter P_o represents the minimum value (return period = 1), describing how far precipitation (P) is from the minimum value. The parameter P_1 is defined as an amplitude that indicates

the difference between the expected precipitation in 2 days P_2 and the minimum value P_0 , i.e., $P_1 \equiv P_2 - P_0$. The inverse of Equation (6) can be approximated to Pareto distribution for short intervals (see Annex 2).

In addition, from Equation (6), it is possible to construct a series of the “daily period return”, i.e., the return period of each value of daily precipitation. This resulting series is henceforth called return-period series (RPS). This process is called total standardization, as opposed to partial standardization, which is defined as the process of generating an RPS with parameters extracted from a portion of the total length of the time series.

To adjust all the functions, we used the maximum likelihood inference, in particular, with the Profile Log-Likelihood (PLL) (Raue et al. 2009). However, the quantity minimized into PLL was not exactly the square of the difference between the empirical probability (π) and the probability predicted by the model fit (π'). These were normalized by the empirical probability (both π and $1 - \pi$, in particular with $(1 - \pi) \cdot \pi$) to minimize the error at the tails of the probability distribution.

To compare the goodness-of-fit of the models, we used the measure of the Normalized Mean Absolute Error (NMAE) and the Akaike Information Criterion (AIC). The NMAE is a better index than the others for precipitation because it compares the relative errors for the lowest and greatest precipitation. According to the Akaike criterion, the model with the lowest AIC is that which best fits to the dataset (Akaike 1975, Burnham & Anderson 2002). In addition, to compare the probability distribution of daily precipitation predicted by the models and the observed precipitation distribution, two tests were employed: Bootstrap Kolmogorov-Smirnov test (Marsaglia et al. 2003, Sekhon 2010) and Anderson-Darling test (Scholz and 1987). As the precipitation data showed many ties, we used an estimation of p value adjusted for ties in Anderson-Darling test and bootstrap resampling in Kolmogorov-Smirnov test (Sekhon 2010). Statistical analysis of the data was calculated by using the R-language (Muenchen & Hilbe 2010, R Development Core Team 2010).

4. Results

First, the commonly used models (Generalized Extreme Value, Pareto and Generalized Pareto Distribution, Gamma, Weibull, Gumbel, Exponential Log-normal II and III) and the three proposed models (Modified GEV, Modified Gumbel, and Alternative Model) were adjusted for empirical distribution function of daily precipitation of every month of the year, for each observatory. Figure 2 shows some fits for three example observatories corresponding to the Spanish climates: A Coruña (Galice) presents the oceanic climate, Tenerife-Los Rodeos (Canary Islands) presents the subtropical climate, and Navacerrada observatory (Madrid) presents the mountain climate. In most of the cases, modified Gumbel and alternative model fitted the high and low precipitation better than the other models. On

the other hand, log-normal III and Generalized Pareto Distribution fitted the low precipitation better than the high precipitation.

The NMAE was obtained by comparing the observed and predicted precipitation for all stations and months. Gumbel, GEV, Log-normal II, and Pareto distributions showed a total NMAE of > 1 for the predicted daily precipitation for most of the observatories, especially for the highest and lowest rainfall; therefore, these models were discarded from the beginning. However, all the other models showed NMAE of < 1 for all values of precipitation and for all stations, except Weibull model that showed high NMAE for some observatories (Fig. 3). The modified Gumbel and alternative model showed the best NMAE (< 0.10), while the other models, such as Log-normal III, GPD, and Modified GEV presented an NMAE that generally ranged between 0.10 and 0.20 (Fig. 3). The AIC indicated that, in general, the alternative model is better, because it showed lower values than those of the Gumbel model. Figure 4 shows the AIC for three observatories (Igeldo, Valencia and Prat), which indicates that in most of the months, the alternative model fits better than the Gumbel model.

Finally, a comparison of the goodness-of-fit of the two best models (modified Gumbel and alternative model) was undertaken using Kolmogorov-Smirnov test and Anderson-Darling test (Fig. 5). A value of $p = 0.05$ refers to a range of 95% confidence that the two data populations are different; therefore, $p > 0.05$ means that the two populations are statistically indistinguishable. The comparison showed that both the models present good p values with both the tests (> 0.05), although in some cases, the alternative model fitted better than the modified Gumbel Model. The fit was worse for the summer months, partly due to the few days of precipitation.

5. Discussion

5.1. Advantages and limitations

This study focused on the examination of the overall daily precipitation variation, instead of investigating the probability of a specific (e.g. maximum) daily precipitation in a month or a year (e.g., Bridges & Haan 1972, Etoh et al. 1986, Durman et al. 2006, Eslamian & Feizi 2006). Some authors investigated wider ranges of probability of daily precipitation, although without reaching the full range of occurrence (Wilson & Toumi 2005). However, the standard probability functions cannot accurately fit the entire range of precipitation, from the frequency of days without precipitation to the frequency of highest extremes. This is because usually the probability function uses two or three parameters, which appears to be insufficient to fit the observed data (Hanson & Vogel 2008). The alternative model of probability developed in this work can accurately represent the entire range of probability of daily precipitation (e.g., Figs. 2 and 6) because it uses four parameters.

Figure 6 is an example of the cumulative probability of monthly precipitation according to Prat observatory (Tarragona). The fitted line shown in Fig. 6 corresponds to the alternative model. The adjustment was made for all values of precipitation, except for the two highest and the two lowest values; thus, extreme values fit well even if they are excluded from the determination of the parameters (NMAE = 0.15). In fact, if we represent all standardized absolute errors throughout the year (Fig. 7), the normalized absolute error is similar for all scales of precipitation, except for the minimum measurable value (0.1 mm). This is an important advantage over other models, due to the use of four parameters.

However, the use of four parameters can lead to difficulties in adjustment by computer algorithms, because it requires the initial values very close to the best estimates of the parameters. If not adequately initiated, the algorithms can find incorrect fits. However, the great flexibility of the alternative model can cause significant differences in the adjustments to similar datasets. Therefore, there may be some interdependence between the parameters, which subtracts robustness from the interpretation or comparison between different settings. For example, in this work, a correlation between P_0 and P_1 has been found, with $R^2 = 0.49$, and between w and $\log(P_1)$ with $R^2 = 0.60$, which suggests partial links during adjustment. For these reasons, in some cases, other models fit better to the daily precipitation than the proposed alternative model. This is true especially for the less rainy observatories, in which the modified Gumbel, among the other models, sometimes presented better results (e.g. see Fig. 3b in August and October).

On the other hand, an important advantage of these models is that it can completely transform a series of daily rainfall into a series of probability: RPS. An RPS is "naked time-series," i.e., information about the mean and variability of daily precipitation falls mainly on the monthly parameters extracted. However, there is a percentage of variability, which is different between two RPS of equal length, because some a priori factors are random. This means that if two RPS are very long, then they tend to be equal in terms of their probability distribution, especially for similar climates. For example, an RPS obtained with partial standardization (long enough) will be similar to that obtained with total standardization, except for small random events related to the short length of the series. Another percentage of RPS variability may be due to climatic features; e.g., two sets of precipitation can have the same probability distribution, although they may differ in the prevailing duration of dry and wet periods and alternation between them.

In addition, this model can be used to characterize and correct the probability distribution of a generic set of daily precipitation data. Its use could introduce improvements in the statistical downscaling of climate models (Widmann 2003), and thus improvements in predicting changes in the precipitation regimes (Kidson et al. 1998, Perkins et al. 2007). This model could also lead to enhanced forecasting weather models and

improved calibration of weather radars (Marks et al. 2010). Furthermore, to capture the rainfall data containing zeros, some authors have proposed to use mixed distribution (Smith et al. 2010). However, the mixed distribution is not suitable, because it is not robust for climatic analysis in other applications, such as to correct climate models.

5.2. Climatic relationships

The probabilistic model that has been developed can be used to study the climate patterns of precipitation. For example, Fig. 8 shows the monthly evolution of the frequency of daily precipitation at Prat meteorological station. In this sense, the parameters fitted for a probability distribution of precipitation should adequately reflect the climate characteristics.

As the Mediterranean climate (with a dry summer) is predominant in Spain, it is expected that the four parameters fitted for the alternative model reflect the differences in the months of the year, especially in the summer. Figure 9 shows this property, especially for the “scale” (P_1) and “smoothness” (k). Logically, the “scale” of the probability distribution is lower and the “smoothness” is higher for less rainy months. This means that the probability of weak rainfall is highest for summer. However, the “shape” parameter (w) is slightly lower in summer, which implies a higher probability (or lowest return period) of extreme rainfall. This is consistent with the characteristics of Mediterranean climate prevalent in Spain.

Accordingly, a correlation between the monthly “scale” (P_1) and the daily average precipitation (P_{day}) was found, with $R^2 = 0.63$ and slope of 0.6 ± 0.1 (Fig. 10a). When this ratio was extracted from the “scale” parameter (P_1), we observed a higher correlation with the “shape” (w). In particular, we found that w is approximately equal to $b \cdot (P_1/P_{day})^c$, where $b = 1.1 \pm 0.1$ and $c = 0.34 \pm 0.05$, with $R^2 = 0.90$ (Fig. 10b). This function is important because it relates the shape of the probability distribution and the ratio between the scale and average precipitation. In other words, the average is related as $P_{day} \sim P_1 \cdot (b/w)^{1/c}$ (see Annex 1).

The seasons in Spain may represent different climates; e.g., the summer of coastal Mediterranean climate is similar to the desert climate, spring is similar to oceanic climate, and the winter of the continental-Mediterranean climate could be similar to cold climates. Therefore, the fit of the alternative probability model for the different months of the year suggests a good performance for different climates. In fact, most of the observatories studied in Spain, representing the ocean, mountain, and subtropical climates, were adequately modeled (Figs. 2 and 3). According to the results (Figs. 3–5), the alternative model works better for rainy climates, because the empirical probability distribution shows a low random error, and thus, it appears to be a function with very regular derivatives.

6. Conclusions

Most of the widely used models (Generalized Extreme Value, Generalized Pareto Distribution, Gamma, Gumbel, Weibull, Exponential and Log-normal) were unable to precisely adjust the set of daily precipitation, including the frequency distribution. It was therefore necessary to use models with three or four parameters, and we tested three models: a modified Gumbel and Generalized Extreme Value model, and an alternative model of probability. The alternative model with four parameters presented the best results for most of the observatories. To be precise, it obtained a mean absolute error of <10% in most of the stations analyzed, and it was similar for all scales of precipitation. The fit was worse for the summer months, due to the few days of precipitation; therefore, probably, the alternative model might work best for rainy climates. This model can be used to characterize the climate patterns of precipitation; hence, it can also be used to correct the probability distribution of a generic set of daily precipitation.

Acknowledgments. This work is a part of the doctoral thesis submitted at the Department of Earth Physics of the University of Valencia. This work is supported by the Department of Environment, Regional Planning, Agriculture and Fisheries of the Basque Government (K-Egokitzen II project, Eortek Funding Program). Furthermore, we acknowledge the State Meteorological Agency of Spain (AEMET) and Hydrographics Confederations of Ebro (CHE) and Júcar (CHJ) for providing the data for this study. In particular, we thank José Ángel Nuñez, Head of the Department of Climatology AEMET delegation in Valencia, and Margarita Martín, AEMET delegate in the Basque Country, for their helpful comments. Finally, it is fair to acknowledge the support of Maddalen Mendizabal (Tecnalia) for raising the issue of probability of daily precipitation.

ANNEX 1

Density function

From Equation (4), we can deduce that the probability density function, $f(\lambda; w, k)$, is given by:

$$f(\lambda; w, k) \equiv \frac{d}{d\lambda} \pi(\lambda; w, k) = \frac{\lambda^{w+\lambda^w e^{-k}} - 1}{\left(1 + \lambda^{w+\lambda^w e^{-k}}\right)^2} \left[w + (1 + w \ln \lambda) \lambda^w e^{-k} \right] \quad (A1)$$

where w is the “shape parameter” and k is the “smoothness parameter,” while λ is the relative precipitation, defined in Equation (3). In Fig. A1 and A2, it must be noted that the parameter k is only related to the probability of medium and high values of λ , while w is related to the shape of all curves of the probability density. The two parameters (w and k) and the variable λ are defined as always positive.

Some statistical characteristics are described as follows. The mean of variable λ can be estimated with the product of the density probability and λ , using methods of numerical integration. It must be noted that for high values of w , the form of probability density function seems to be a Gaussian curve (see Fig. A1), with mean $\mu \approx 1$. This is because, if $k \gg 1$, the Equation (A1) corresponds approximately to the log-logistic distribution (Ashkar & Mahdi 2006). Therefore, the mean will be similar; in particular, we found that:

$$mean(\lambda) \approx \left(\frac{\frac{\pi}{w}}{\sin\left(\frac{\pi}{w}\right)} \right)^{g(w,k)} \quad (A2)$$

where $g(w, k)$ is a smooth function that tends to 1 for high values of w and k . It must be noted that the mean of λ is approximately $(\pi/w)^g$ for low values of w . On the other hand, the derivative of Equation (A1) shows that the mode is a nonlinear long function of w and k ; however, it is always between 0 and 1, and the influence of k is very low. For this reason, the function of mode can be approximated by the log-logistic mode (McLaughlin 2001):

$$mod(\lambda) \approx \begin{cases} 0 & si \quad w < 1 \\ \left(\frac{w-1}{w+1} \right)^{\frac{1}{w}} & si \quad w \geq 1 \end{cases} \quad (A3)$$

where $mod(\lambda)$ is the mode of λ .

ANNEX 2

Inverse function

Equation (4) has no explicit inverse, and therefore, it is necessary to use programmed algorithms or find an approximate inverse function:

$$\frac{P - P_o}{P_1} = (T - 1)^{m(T-1)} \quad (\text{A4})$$

where $m(T - 1)$ is an adjustable function of return period, which is similar to the exponent of Equation (A4). For short intervals of periods, Equation (A4) can also be approached to Pareto distribution (Hardy, 2010), according to:

$$\frac{P}{P_1} \approx T^{m_o} \quad (\text{A5})$$

where P_o and m_o are the adjustable parameters. For a rainfall event with a duration different of a day, we can consider the duration as an expected time for a given rainfall accumulation in the context of the event. Thus, it is possible to approximate the temporal distribution of maximum accumulation to another power distribution:

$$\frac{P}{P_1} \approx t^{h_o} \quad (\text{A6})$$

where t is the duration, and P_1 and h_o are the adjustable parameters. By adding the expressions (A5) and (A6), an Intensity-Duration-Frequency (IDF) curve is obtained (Ben-Zvi, 2009):

$$P(T, t) \approx P_1(T_1, t_1) \left(\frac{T}{T_1} \right)^{m_o} \left(\frac{t}{t_1} \right)^{h_o} \quad (\text{A7})$$

where $P(T, t)$ is the expected rainfall accumulation for a T -return and t -duration, $P_o(T_o, t_o)$ is the expected rainfall accumulation for a T_o -return and t_o -duration, and m_o and h_o are the adjustable parameters.

REFERENCES

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Cont* 19: 716-723
- Alves M, Verdière AC (1999) Instability dynamics of a subtropical jet applications to the Azores Front-Current System-eddy driven mean flow. *J Phys Oceanogr* 29: 837-864
- Ashkar F, Mahdi S (2006) Fitting the log-logistic distribution by generalized moments, *J. Hydrol* 328: 694–703
- Begueria S (2005) Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. *J Hydrol* 303: 215–230
- Ben-Zvi A (2009) Rainfall intensity-duration-frequency relationships derived from large partial duration series. *J Hydrol* 367: 104-114
- Burnham KP, Anderson D (2002) *Model Selection and Multi-Model Inference*. Springer-Verlag, NY, ISBN-10: 0387953647
- Bridges TC, Haan CT (1972): Reliability of precipitation probabilities estimated from the gamma distribution. *Mon Wea Rev* 100: 607-611
- Capel Molina JJ (2000) *El Clima de la Península Ibérica*. Ed. Ariel, Barcelona
- Cervený RS, Lawrimore J, Edwards R, Landsea C (2007): *Extreme Weather Records: Compilation, Adjudication and Publication*. *Bull Amer Meteor Soc* 88: 853-860
- Durman CF, Gregory JM, Hassell DC, Jones RG, Murphy JM (2006) A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates. *Quart J Roy Meteor Soc* 127: 1005–1015
- Ebert EE (2001) Ability of a poor man's ensemble to predict the probability and distribution of precipitation. *Mon Wea Rev* 129: 2461-2480
- Eslamian SS, Feizi H (2006) Maximum monthly rainfall analysis using L moments for and arid region is Isfahan Province, Iran. *J Appl Met Clim* 46: 494-503
- Etoh T, Murota A, Nakanishi M (1986) SQRT-Exponential type distribution of maximum, hydrologic frequency modelling. *Proceedings of the International Symposium on flood frequency and risk analyses*, 14-17 May 1986, Louisiana State University, Shing VP (Ed), Reidel Pub Com USA, 253-264
- Granger RJ (1989) An examination of the concept of potential evaporation. *J Hydrol* 111: 9-19
- Hanson LS, Vogel R (2008) The probability distribution of daily rainfall in the United States. *World Environmental and Water Resources Congress 2008*, Ahupua A: 1-10. doi 10.1061/40976(316)585

- Hardy M (2010) Pareto's Law. *Math Intel* 32: 38-43
- Jeevananda SR (1995) Sensitivity of some potential evapotranspiration estimation methods to climate change. *Agric Forest Meteorol* 77: 121-125
- Kharin VV, Zwiers FW (2000) Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. *J Climate* 13: 3760-3788
- Kidson JW, Thompson CS (1998) A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *J Climate* 11: 735-753
- Kysely J (2002) Comparison of extremes in GCM-simulated, downscaled and observed central-European temperature series. *Clim Res* 20: 211-222.
- Marks DA, Wolff DB, Carey LD, Tokay A (2010) Quality control and calibration of the dual-Polarization radar at kwajalein, RMI. *J Atmos Oceanic Technol* (in pres.)
- Marsaglia G, Tsang WW and Wang J (2003) Evaluating Kolmogorov's distribution. *Journal of Statistical Software*, 8/18
- Martín-Vide J (2004) Spatial distribution of a daily precipitation concentration index in peninsular Spain. *Int J Clim* 24: 959-971
- McLaughlin MP (2001) A Compendium of Common Probability Distributions, http://www.causascientia.org/math_stat/Dists/Compendium.pdf (2011-02-15)
- McGuinness JL, Leslie HP (1972) Maximum potential evapotranspiration frequency. *J Irrigat Drain Div* 2: 207-214
- Muenchen RA, Hilbe JM (2010) *R for Stata Users. Statistics and computing*, Springer, ISBN: 978-1-4419-1317-3
- Lana X, Martínez MD, Serra C, Burgueño A (2004) Spatial and temporal variability of the daily rainfall regime in Catalonia (northeastern Spain), 1950–2000. *International Journal of Climatology* 24: 613–641
- Öztürk A (1981) On the study of a probability distribution for precipitation Totals. *J Appl Met* 20: 1499-1505
- Park JS, Jung HS (2002) Modelling Korean extreme rainfall using a Kappa distribution and maximum likelihood estimate. *Theor Appl Climatol* 72: 55-64
- Peliz A, Rosa TL, Santos AMP, Pissarra JL (2002) Fronts, jets, and counterflows in the Western Iberian upwelling system. *J Mar Syst* 35: 61-77
- Perkins SE, Pitman AJ, Holbrook NJ, McAneney J (2007): Evaluation of the ar4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *J Climate* 20: 4356-4376

- Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* 25: 1923–9.
- R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Sekhon JS (2010). Matching: Multivariate and Propensity Score Matching with Balance Optimization. R package version 4.7-11. URL <http://CRAN.R-project.org/package=Matching>.
- Scholz F.W. and Stephens M.A. (1987), K-sample Anderson-Darling Tests, *Journal of the American Statistical Association*, 399: 918–924
- Smith JA, Villarini G and Baeck ML (2010): Mixture Distributions and the Hydroclimatology of Extreme Rainfall and Flooding in the Eastern US. *Journal of Hydrometeorology* 2010: e-View
- Widmann M, Bretherton CS, Salathé EP (2003) Statistical precipitation downscaling over the northwestern United States using numerically simulated precipitation as a predictor. *J Climate* 16: 799–816
- Wilson PS, Toumi R (2005) A fundamental probability distribution for heavy rainfall. *Geophys Res Lett* 32: L14812

Figure captions

FIG. 1. Location of the meteorological stations used in this study and the annual precipitation for period of 1950–2000, interpolated with Thin-Plate Spline method.

FIG. 2. Examples of cumulative probability and return period, according to the empirical values and several models fitted, for three observatories with different climates: (a) and (b) Tenerife - Los Rodeos (subtropical climate); (c) and (d) A Coruña (oceanic climate); and (e) and (f) Navacerrada (mountain climate).

FIG. 3. Comparison of the Normalized Mean Absolute Error (NMAE) of the predicted precipitation *versus* the observed prediction of the empirical distribution function, for the set of 108 observatories, with the adjustment of: (a) Several models applied for the annual precipitation: Gamma, Weibull, Exponential, Log-normal III, Generalized Pareto Distribution (GPD), Modified Generalized Extreme Value (M. GEV), Modified Gumbel and Alternative Model. (b) Comparison of modified Gumbel and Alternative Model, applied for the monthly precipitation.

FIG. 4. Comparison of the monthly AIC of the adjustments to Modified Gumbel and Alternative Model of accumulated probability for three observatories: (a) Igeldo (Gipuzkoa), (b) Valencia (Valencia), and (c) Prat (Barcelona).

FIG. 5. The P value corresponding to (a) Kolmogorov-Smirnov test with bootstrap resampling and (b) Anderson-Darling test with adjustment for ties, to compare the empirical and predicted precipitation using modified Gumbel and alternative model.

FIG. 6. Return period *versus* daily precipitation for the months of the year for the observatory of Prat (Barcelona): observed values (gray circles) and predicted values (black line).

FIG. 7. Validation of the daily precipitation for the months of the year for the observatory of Prat (Barcelona).

FIG. 8. Comparison of the return curves of daily precipitation for the months of the year for the observatory of Prat (Barcelona).

FIG. 9. Monthly values of four adjustable parameters of Alternative Model (P_0 , P_1 , w , and k) applied to 108 observatories of Spain.

FIG. 10. (a) Relationship between “scale” parameter (P_1) of the annual fit and the daily average precipitation (P_{day}). (b) Relationship between “form” parameter and the ratio P_1 / P_{day} .

FIG. A1. Some examples of the probability density curve, with $w = 1$ and $k = 1, 2, 10$; and with $k = 2$ and $w = 1, 2, 10$.

FIG. A2. Some examples of the cumulative probability curve, with $w = 1$ and $k = 1, 2, 10$; and with $k = 2$ and $w = 1, 2, 10$.

Figures

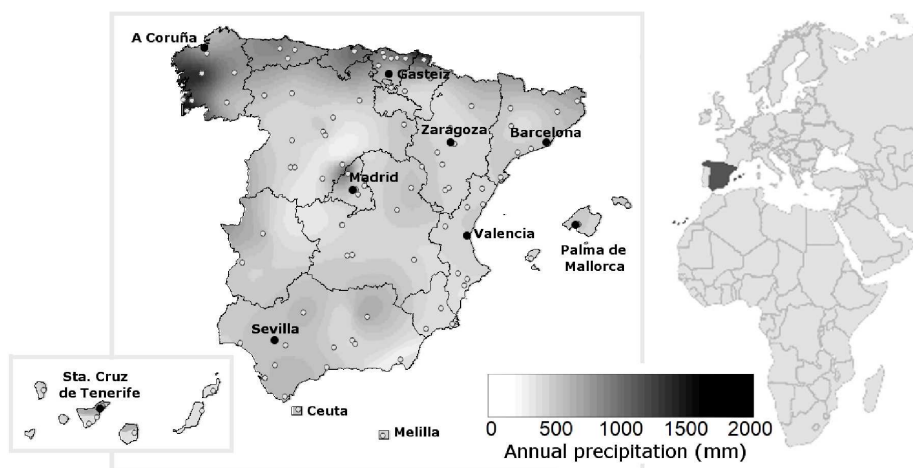


FIG. 1

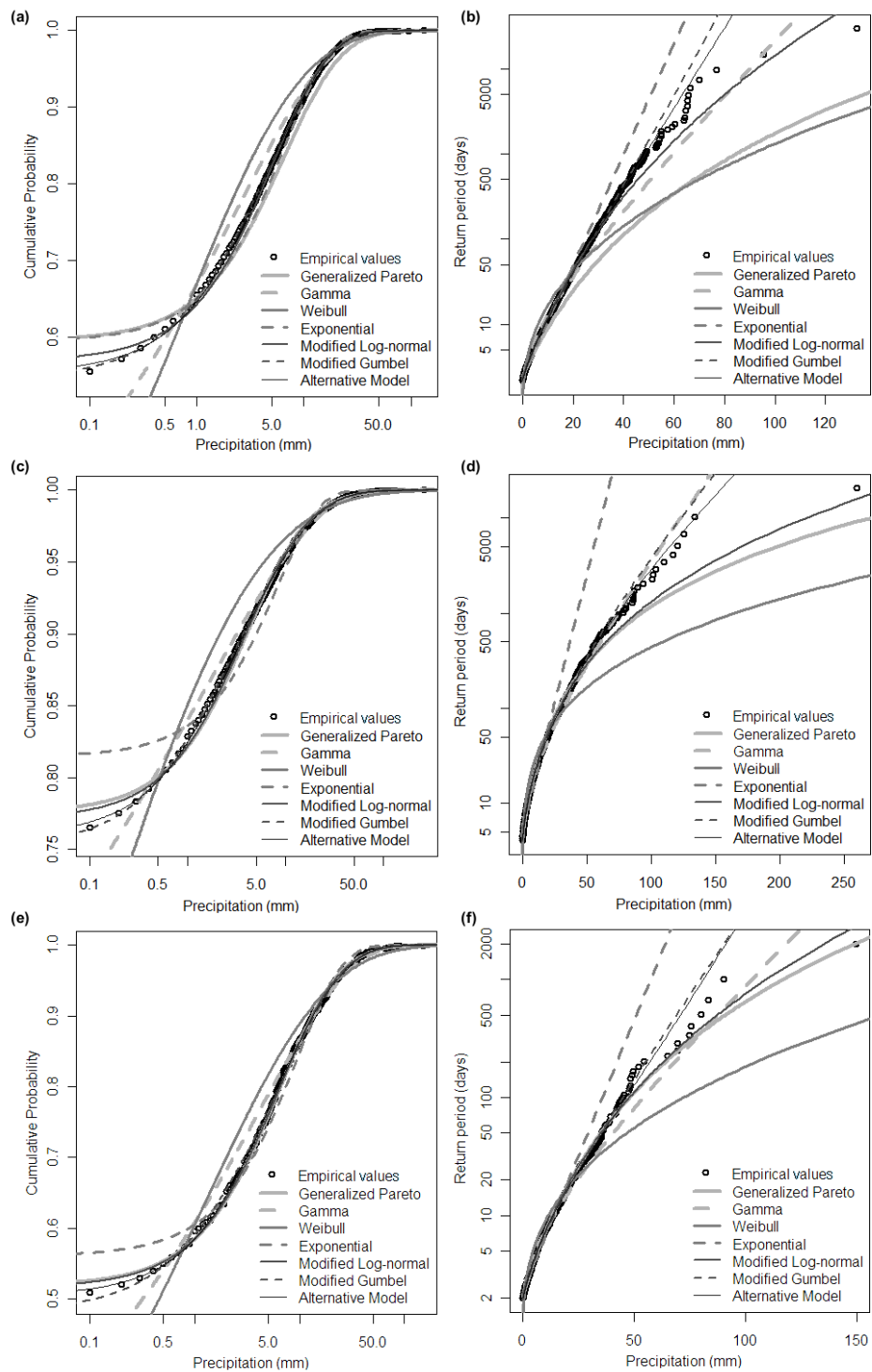


FIG. 2

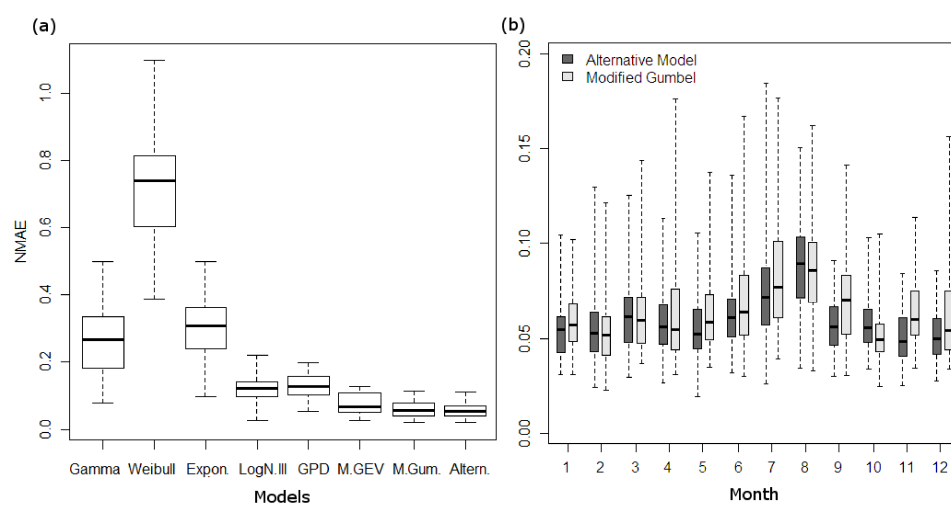


FIG. 3

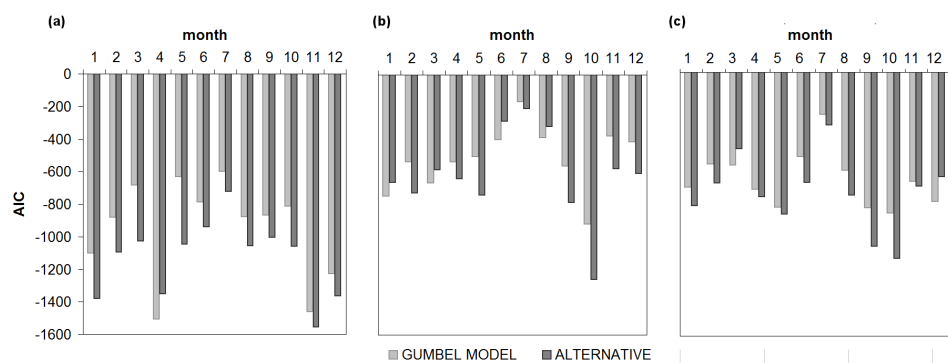


FIG. 4

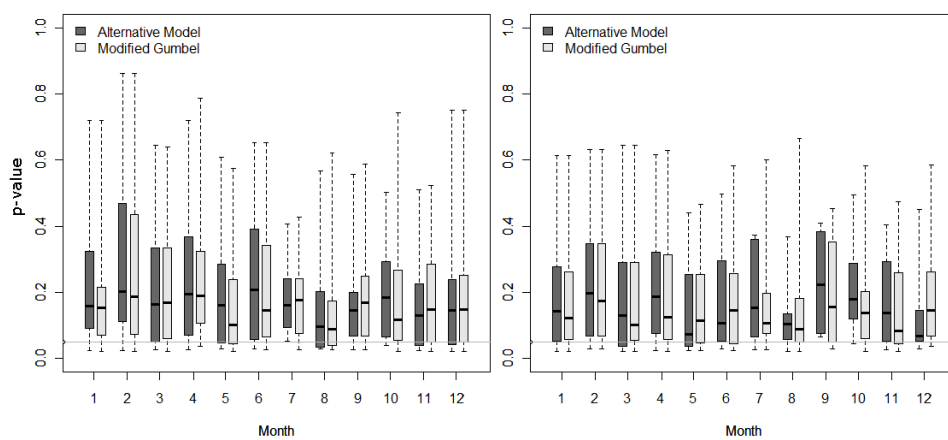


FIG. 5

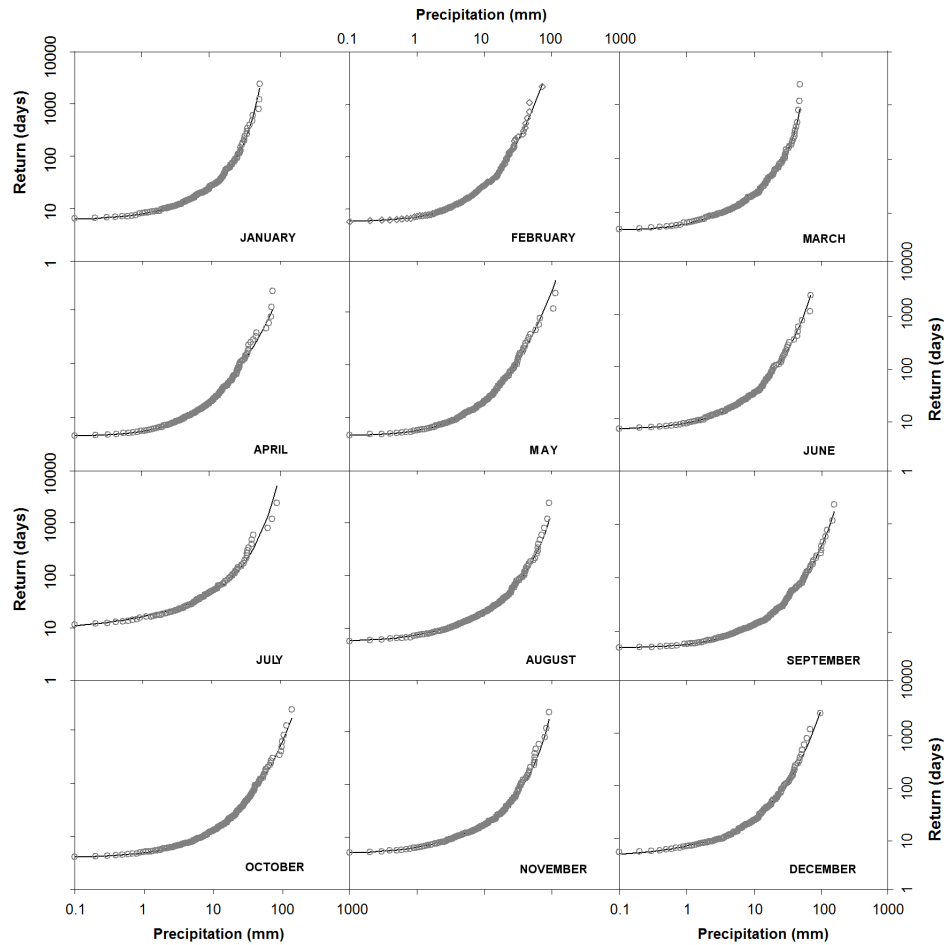


FIG. 6

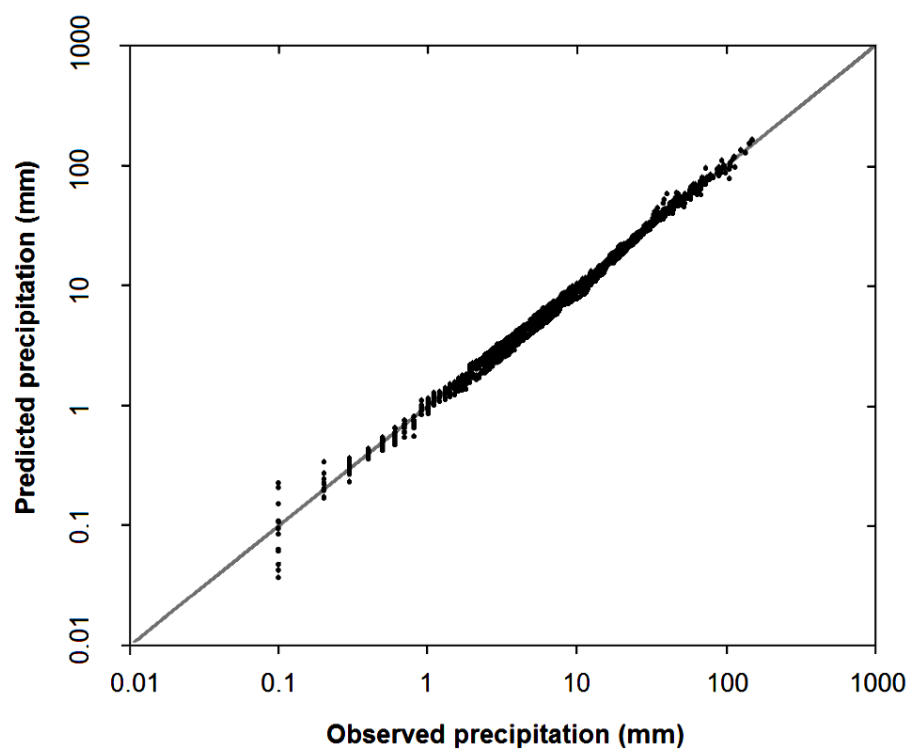


FIG. 7

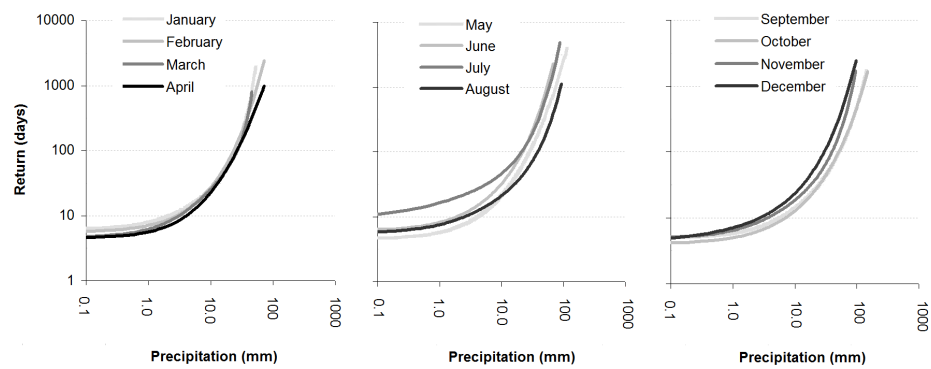


FIG. 8

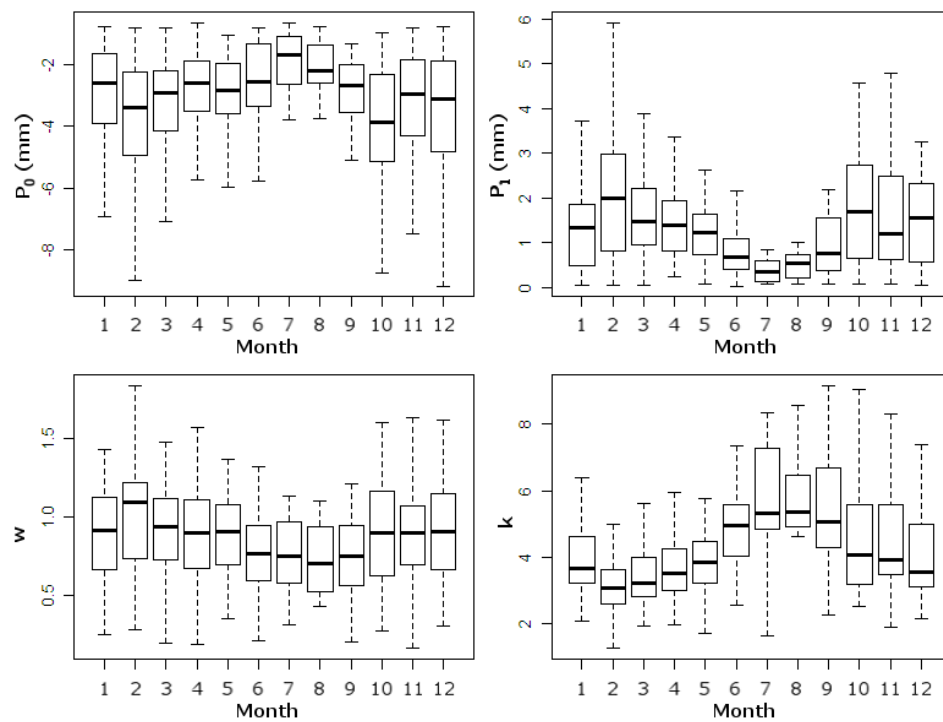


FIG. 9

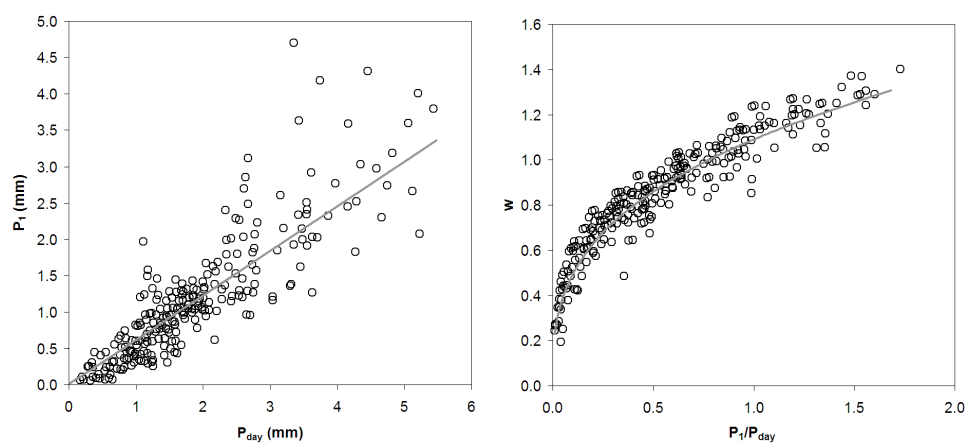


FIG. 10

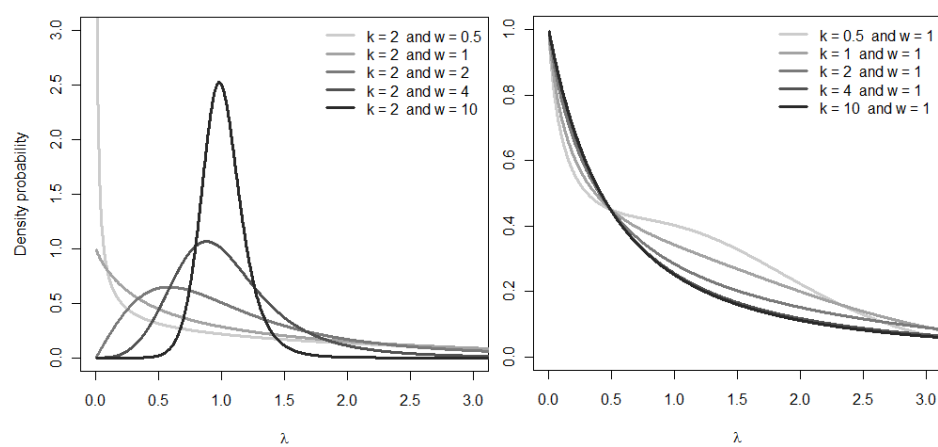


FIG. A1

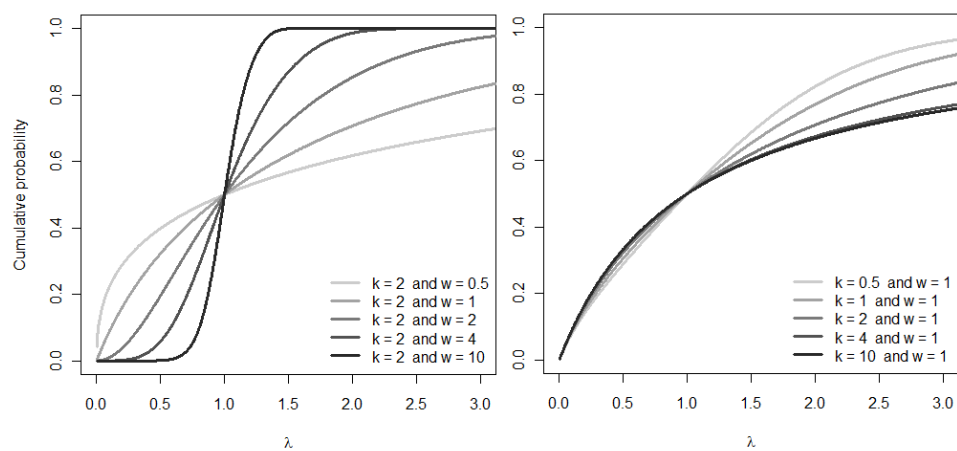


FIG. A2

**ANEXO 2. Probabilistic correction of RCM precipitation for the
Basque Country (Northern Spain)**

(Enviado a Climate Research, 2011)

*Revista con un factor de impacto de 2,25 y que
ocupa la posición 18 de un total de 70 revistas
en el tema de “meteorología y ciencias atmosféricas”*

Probabilistic correction of RCM precipitation for the Basque Country (Northern Spain)

Roberto Moncho^{*,1}, Guillem Chust² and Vicente Caselles³

¹Fundación para la Investigación del Clima, C/ Gran Vía, 22^{do}, 6ºIzda., 28041 (Madrid, Spain)

²AZTI-Tecnalia, Marine Research Division, Txatxarramendi ugarte a z/g, 48395 Sukarrieta (Vizcaya, Spain)

³Department of Earth Physics and Thermodynamics, Faculty of Physics, University of Valencia. N. 50, Dr. Moliner. 46100 Burjassot (Valencia, Spain)

*Corresponding author: E-mail: robert@temps.cat

ABSTRACT

In the context of global climate change, the study of precipitation in the Basque Country (northern Spain) is of great significance for regional planning and ecosystem conservation. It is therefore important to develop a methodology to correct the systematic errors of precipitation projected by climate models. In this study we used two modified versions of two probability models, respectively: the distribution of Gumbel and Log-Logistic. With these tools, the daily probability distribution of 7 regional climate models (RCM) was transformed: Aladin-ARPEGE, CLM-HadCM3Q0, HIRHAM-HadCM3Q0, HIRHAM-BCM, RECMO-ECHAM5-rt3, REMO-ECHAM-rt3 and PROMES-HadCM3Q0. The reference database comprises a total of 156 observatories of the Meteorological Agency. The results showed that the implemented method obtains an error of less than 5% on the simulation of the mean, and 1% on the simulation of the number of days without precipitation. Throughout the twenty-first century and under the A1B scenario, we can expect an intensification of daily and subdaily precipitation. For example, precipitation with a return period of 100 years will increase to 30% in 2001-2040 compared to 1961-2000, in western areas of the Basque Country. This intensification is interpreted as a consequence of the process of 'mediterraneanisation' that the most southern ocean climate may go through.

1. Introduction

One of the main environmental concerns of society is focused on the possible effects of global change. Water is the most important natural resource for most human activities, but especially for natural ecosystems, on which we all depend. Often, water is a scarce resource, and a clearly limiting factor in natural resources and activities such as agriculture and industry. But sometimes, the abundance or incorrect distribution of water can result in a risk of flooding that affects the urban environment and human activity (Olcina 2008). Therefore, the study of precipitation in the context of global change is critical for a land-use planning consistent with the preservation of the environment and suitable for economic activities.

In order to analyse the potential climate changes in precipitation regime, there are several numerical models that attempt to simulate the climatic conditions of the global atmospheric circulation. The Intergovernmental Panel on Climate Change (IPCC, Meehl et al 2007, Christensen et al 2007) provides different global climate scenarios based on possible developments in global economic activity. Specifically, four groups of emission scenarios of greenhouse gases (GHGs) in relation to fossil fuels are identified, according to the prevailing socio-economic model in the future (A1, A2, B1 and B2). A priori all emissions scenarios are equally likely. However, the maximum extraction of fossil fuels ("peak oil") is expected to occur before 2030 (Alekklett et al 2010, Carlson 2011). For that reason, any of the scenarios that project a combination with renewable energy seem more likely to happen.

In this regard, members of the project ENSEMBLES (Niehörster et al 2008), promoted by the European Commission, have focused their efforts on the A1B scenario, which indicates a rapid global growth combined with renewable energy. The ENSEMBLES project is developing a climate prediction ensemble based on the latest high-resolution modeling techniques, at global and regional levels. The objective of that project is to quantify and reduce uncertainty in the climate prediction, to maximize the use of results, linking agriculture, health, food security, energy, water resources, climate risk management and insurances. In this global context, Basque researchers have developed the K-EGOKITZEN project, which aims to estimate the impacts of regional climate change and some possible ways to adapt, in response to future environmental scenarios.

According to most climate projections for the twenty-first century, the north of the Iberian Peninsula lies in a borderland between the decrease (south) and the increase in rainfall (north) (Christensen et al 2007, Goubanova and Li 2007). The expansion of the Hadley cell could push the tropical subsidence zone (desert areas) up into the south Peninsula (Lu et al 2007). At the same time, it is expected that air warming will facilitate an increase in precipitable water content in northern Europe, causing a greater intensity of rainfall (Déqué et al 2007, Goubanova and Li 2007). However, the Basque Country is located between the two areas, so some kind of regionalization is needed for the climatic signals of different models in relation to the climatic characteristics unique to the area. One way to regionalize climate signal

is the calibration of models with data from observatories (Ines and Hansen 2006, Palmer et al 2008, Benestad 2010).

Therefore, the main objective of this work is to correct the Probability Distribution Function (PDF) of different Basque climate projections for this century under the A1B scenario. It aims to build a tool to be used for the analysis of potential flood risk and consequently, to quantify its potential impacts.

2. Materials

2.1. Study area

This study focuses on the Basque Country due to the special importance of climate change in areas such as urban planning and water resources, but also because of the possible effects on the rich biodiversity of this territory. In fact, the Basque country has many natural areas with some protection. In particular we can mention a Biosphere Reserve (Urdaibai), 52 Sites of Community Importance and 6 Special Protection Areas of Birds, among others. The Oka River estuary, located in the Biosphere Reserve of Urdaibai (Vizcaya), is one of the most biologically diverse and best preserved in the Basque Country, together with the Bidasoa (Borja et al 2004).

With regard to climatic aspects, its location in the north of the Iberian Peninsula causes these area to be predominantly affected by the polar jet waves or Rossby waves, but sometimes also by the subtropical jet (Alves and Verdière 1999, Peliz et al. 2002). The Basque Country, like most northern regions, is characterized by a very mountainous relief relatively near the sea. These orographic conditions cause a sharp contrast between the Cantabrian rainfall and the southern slope (Garmendia et al 1989).

In fact, annual precipitation ranges from 400 mm in the south of Alava, with Mediterranean climate, and 2500 mm in the northeast of Guipúzcoa, with oceanic climate (Capel-Molina 2000). This last climate is characterized by regular rainfall throughout the year (in some cases more than 50% of rainy days). In contrast, Mediterranean climate is characterized by dry summers and heavy rainfall concentrated in a few days (Martín-Vide 2004, Lana et al 2004), especially in autumn and spring. Throughout the year, the number of days with precipitation usually ranges from 20 to 30% on average, but it can reach values below 20% in July and in some cases more than 50% in April (Capel-Molina 2000).

2.2. Observed precipitation data

In this work we have used data from a total of 156 observatories of State Meteorological Agency of Spain (AEMET). Of these, 67 observatories have a daily availability (Fig. 1a). However, the availability of many observatories is interrupted for a few years, so the useful number of stations varies considerably over time, getting two periods with notable peaks, one between 1972 and 1980 and again between 1986 and 1995 approximately (Fig. 1b).

Observatories with monthly records show an average of 30% of gaps, and those with daily records show 15% of gaps. For the reference period 1961-2000 only 19 stations have a full series, while the total data volume of 156 stations is just over 50% compared to the maximum possible volume. Therefore it was necessary to also use a cross reference series of precipitation in the Basque Country, spatially reconstructed using a digital terrain model and a multi-linear fit (Moncho et al 2009b).

2.3. Regional Climate Models

The models used in this study are the numerical outputs of 7 RCMs (Fig. 2), available from the European project ENSEMBLES-rt3 (<http://ensemblesrt3.dmi.dk/>). These models have a daily temporal resolution and spatial resolution of 25 km × 25 km, and are forced with the SRES-A1B future scenario, corresponding to a rapid global growth with a combined contribution of different energy sources. The choice of SRES-A1B scenario by members ENSEMBLES-rt3 is justified because it is an intermediate stage between the different projections of increased greenhouse gases (Christensen et al 2007, Niehörster et al 2008).

In particular, we took the average of 20 grid points in the next 7 regional models:

- a) Aladin-ARPEGE: National Center for Meteorological Studies (CNRM), France.
- b) CLM-HadCM3Q0: Swiss Institute of Technology, Department of Geography (ETHZ), Switzerland.
- c) HIRHAM-HadCM3Q0: Norwegian Meteorological Institute (METNO).
- d) HIRHAM-BCM: Norwegian Meteorological Institute (METNO).
- f) RECMO-ECHAM5-rt3: Meteorological Institute Netherlands (KNMI).
- g) REMO-ECHAM-rt3; Max-Planck Institute (MPI), Germany.
- h) PROMES-HadCM3Q0: University of Castilla-La Mancha (UCLM), Spain.

3. Methodology

3.1. Probability distribution function

The cumulative probability $\pi(p \geq P)$ is defined as the probability that an observatory records a rainfall p equal to or greater than P . Logically, the probability of exceeding each threshold of precipitation, P , can be estimated empirically with the normalized sum of the number of times that happens. The return period or the expected time T (in days) between two precipitation events with p equal to or greater than P is given by the inverse of the cumulative probability:

$$T(P) \equiv \frac{1}{\pi(p \geq P)} \quad (1)$$

In order to model the empirical cumulative distribution function (ECDF) of each month, in this study we used a generalized version of Gumbel distribution and another of Log-Logistic (Moncho et al 2011):

$$\pi(\lambda_0 \geq \lambda) = \exp[-\exp(\lambda^w + k)] \quad (2)$$

$$\pi(\lambda_0 \geq \lambda) \approx \frac{1}{1 + \lambda^{w+\lambda^w e^{-k}}} \quad (3)$$

where w y k are adjustable parameters while λ is the relative precipitation, which in turn depends on two adjustable values, one representing the most probable precipitation P_o , called the location parameter, and another referred to the standardizing value scale parameter, P_1 :

$$\lambda \equiv \frac{P - P_o}{P_1} \quad (4)$$

The adjustments to the theoretical curves to the empirical probability are done by statistical inference, and in particular with the Profile Log-Likelihood (Akaike 1975, Raue et al 2009). When the monthly parameters are extracted from the probability distribution of a time series, it is possible to construct a series of "return period of each day", i.e. the return period associated with each value of daily precipitation. The resulting series will now be called Return Period Series or RPS. Therefore, we define the total standardization process as compared to the partial standardization, which is defined as the process of generating an RPS with the parameters extracted from a portion of the total length of the series (for example, the reference period).

3.2. Probabilistic transformation of climate models

Probabilistic transformation consists of transferring the shape of each ECDF to the precipitation simulated by the models. In order to do this, firstly time-series of models are shed of their climatic information by using a partial standardization. The selected reference period is obviously the common period regarding the observatory simulated in each case. The parameters adjusted to the time series of the observatory are added to the RPS obtained for each model.

In this case, the climate signal remains unaltered, because it is transmitted in the RPS. In a series of return periods it may be seen if there is any change in the future on the occurrence frequency of an event type with a priori low probability. And that is independent of whether the RPS is estimated by partial or complete standardization, because the time distribution of the RPS may be very different in a climate period compared to another.

However, if we try to transfer the ECDF shape of an observatory with little data, then we can find differences in the results because there are no guarantees as to the climate representativeness if the time series are very short. It is therefore important to complete the observed series with some longer series, as far as possible.

In order to complete the short series we used probabilistic transformation of the most similar and longest series (control series): the RPS of each control series was obtained by partial standardization in the common period of time for the studied series. After having partially standardized the control series, the fit parameters of the studied time series were applied. This forces the control series to seem more like the series studied in the common reference period. Finally, the resulting transformation of the control series served to expand the extension of the time series for each observatory. The choice of the control series for each observation was made on the basis of those RPS with the highest Pearson correlation and the lowest mean normalized absolute error (MNAE), in relation to the series that it was intended to simulate. Furthermore, all the fits that had a MNAE lower than 20% were rejected.

3.3. Analyses of mean and extreme precipitation

Once the PDF of each climate mode was transformed, we analysed the mean and the extrem precipitation. In particular, the rainfall expected for a return period of 100 years was projected for both the past (1961-2000) and the beginning of this century (2001-2040). Therefore, the expected relative change of precipitation according to seven climate models was estimated for each observatory.

In order to represent a map of the relative change in precipitation, two steps were taken. First we estimated a linear regression among a series of simulated observatories predictors (daily records) and the relative changes predicted for precipitation. The selected predictors that were: coordinates, altitude, direction and mean annual precipitation. From the previous regression, we estimated the values of precipitation change and its error for the observatories with monthly record. Secondly, an interpolation was done by using Thin Plate Splines (TPS) with two dimensions (Green and Silverman 1994, Furrer et al 2010), with all the observatories, both direct simulations and those estimated by multiple linear regression. Therefore, a study of the quality of interpolation by cross-validation was undertaken for the two steps at a time. In order to do this, we used the Pearson correlation and Relative Mean Absolute Error (RMAE).

Finally, we estimated the sub-daily rainfall by using the method of index n (Moncho et al 2009a). For any statistical model, it is possible to estimate the expected rainfall with a function of return period, $f(T)$. Therefore, the expected precipitation P and P_o , for a duration of t and t_o minutes respectively, can be related by:

$$\frac{P}{P_o} \approx f(T, T_o) \cdot \left(\frac{t}{t_o} \right)^{1-n} \quad (5)$$

where n is the index of precipitation. This index was estimated for all observatories and simulations from the precipitation of at least 3 days. Finally we studied the possible changes in the PDF along the XXI century.

All treatments of the data and the results were carried out by using statistical packages based on R language (R Development Core Team, 2010).

4. Results

4.1. Filling and probabilistic transformation

Firstly, a process of expansion and gap-filling was performed by using partial standardization of different control series. All the observatories of at least 40 years (11 of 67 observatories) were selected and transformed according to the probability parameters of the shorter series, by using the models of equations 2 and 3. Of the 56 observatories with less than 40 years of data, only 34 of them were successfully completed, because the rest did not pass the threshold fit (we obtained a relative error greater than 20%).

As far as the correctly extended stations are concerned, the process extended their control series between 3 and 64% (mean of 21%) to complete the 40-year reference period (1961-2000). It was found that the Pearson correlation (R) between each observed series and the most similar one ranged from a low of 0.61 to a maximum of 0.90 (with mean 0.74). The mean normalized absolute error (MNAE) of the ECDF of the simulated series was between 0.02 and 0.12 (with mean 0.05) compared to the corresponding observed series.

Therefore, a total of 45 observatories with at least 40 years of data were obtained. The time series of these observatories were standardized again by using the two probability models (equation 2 and 3). For each month of the year and each observatory, the theoretical adjustment with less normalized absolute error was selected. In 72% of the cases, the selected model was the generalized version of the log-logistic distribution (Fig. 3).

With the parameters obtained from standardization, the precipitation series of the seven RCMs were transformed to simulate each observatory. It was noted that the correction of the annual probability distribution is transmitted properly to the climate models (Fig. 4). That includes the frequency of days without precipitation, which is implicit in the lower end of the probability curves (inverse of the minimum return period).

4.2. Validation of probabilistic transformation

If the monthly average and the daily deviation of each simulated series are compared with the corresponding observed series, then it can be noted that the relative value of mean absolute error (RMAE) is generally between 1 and 15% depending on season and RCM (Fig. 5). The average value of RMAE for all observatories simulated falls below 5% in all cases and in all months, except for the Aladin-ARPEGE for the summer months, which has an average error around 10%. Similar result is obtained for the relative bias (RBIAS) which is also less than 5% in most cases, except at Aladin-ARPEGE model. The number of days without precipitation simulation has an even lower relative error, around 1%, and a negligible mean bias (Fig. 6).

In addition to low relative error, all the simulated series have a high correlation with reference observatories. For example, the correlation of monthly average is around $R^2 = 0.99$ for all models, with a slope ranging between 1.05 and 1.14

depending on the model (Table 1 and Fig 7). The correlation is slightly lower for the simulation of the daily deviation and number of days without precipitation, with R^2 values ranging between 0.96 and 0.98, but the slope is closer to 1, fluctuating between 0.95 and 1.02.

4.3. Analysis of changes in mean precipitation and extreme

For the spatial interpolation of precipitation changes, we performed a multi-linear regression between the values of the change and predictors of observatories (coordinates, altitude, direction and annual rainfall), which obtained a correlation of $R^2 = 0.87$ and a relative mean absolute error of $RMAE = 0.2$. For the confidence interval for the change of precipitation a multiple regression with the same predictors was done as well. Cross validation showed that the correlation was lower than in the previous case, around $R^2 = 0.52$, but error was $RMAE = 0.12$.

According to the seven RCMs studied, rainfall may have a slight decrease in the Basque_Country, especially in the spring. In particular, it provides a possible decrease of up to 15% in the south of the Basque Country, for the quarter April-June of 2001-2040 versus 1961-2000 period. In autumn, the average rainfall may decrease up to 10% in northern. However, the p-value of these projections is considerably less than 0.95 (Fig. 8). For the period 2041-2080, the projected decrease in precipitation in the spring in the south ranges between 10 and 30% with 95% confidence.

To analyze possible changes in extreme precipitation, the study focused on the return period of 100 years. For the time period of 2041-2080, the results showed an increase in the precipitation of around 30% in most observatories and models (Fig. 9). For the time period of 2001-2040 the variation is not so clear, although some western observatories in the Basque Country agree to an increase in extreme precipitation for most models studied. In particular, there was an area in the west of the Basque Country with a projected change of up to 30%, with an interval of more than 99% (Fig. 10). However, the probability of change is expected to be smaller, since the latter's relative error is 15% in the interpolation.

As far as the index n is concerned, we studied the empirical probability distribution for all precipitation and for precipitation that exceeded 50 mm in one day (Fig. 11). Regarding all precipitation, no significant changes were identified. This is because the type of precipitation is so diverse, masking any change in the index n , at least for the periods studied in the Basque Country. However, for precipitation higher than 50 mm in one day, we do observe changes in the probability of index n between 0.6 and 0.8. This implies a shift in the average $+0.03 \pm 0.01$ (from 0.57 to 0.60). Ie not only would the maximum daily precipitation increase (Fig. 10 and 11), but also a greater concentration of rains on the subdaily time scale could be expected.

5. Discussion

5.1. Choice of methodology

Due to the low spatial resolution of RCMs, the topography of a small region like the Basque Country is very smooth. For example, the highest peak of the grid is around 800m, compared with almost 1600m of Aitxuri (Guipúzcoa). For that reason, there is a weak correspondence between the grid points and observatories. In terms of absolute values, the simulations are unrealistic for the local level, so it is necessary to apply some type of mathematical regionalization (Piani et al 2010). Meanwhile, it is expected for the climate change signals to be reflected somehow in the change of precipitation probability distribution.

In order to correct the probability distributions of regional climate models, we chose to use theoretical curves adjusted to the observatories, because if we use the unadjusted empirical curves they can incorporate additional irregularities in the probability curve of the corrected models. In addition, the theoretical curves can easily be extended to estimate the extreme rainfall, whereas the empirical curves cannot. However, there are some possible problems in the use of theoretical adjustments. One of them is the problem of seasonality, while the other is the stationarity.

The first problem concerns the difference in the probability of precipitation depending on the season and even the month or fortnight. That means a probability distribution for each season or each month (or lower scales) needs to be adjusted so that the time series of return periods (RPS) is completely non-seasonal (no autocorrelation). The problem of seasonal separation of probability distributions is the decision of where we do the cuts. In this work we have taken the official month for three reasons:

- 1) 12 sections are considered sufficient for the resulting RPS to be non-seasonal (Elshamy et al 2009).
- 2) The borders of the cuts are smooth enough, ie there is not much difference between the probability of one month and the next one (maybe except July).
- 3) If you choose a non-calendar month (30 day moving) there would be 365 different probability distributions for each observation and model, which causes a long computation time. In addition, and taking into account what has been previously stated, it seems unnecessary.

The second problem concerns the fact that the empirical relationships in the variables of the past may be non-stationary in the future. This issue does not affect the probabilistic transformation method very much for two reasons:

- 1) When a time series is partially standardized (future with past), we get an RPS which contains all the signal of climate change (because this way the change of frequency of certain return periods can be analysed). When the twelve probability distributions of each observatory are added to the RPS of an RCM,

then it results an automatical (possible) change in the future probability distributions in relation to the past. Therefore, in reality, the future is not being forced with a probability of the past, since the essential information of the probability is given by the RPS.

- 2) Although, for example, September might seem more like August in the future, this possible change is also contained in the RPS, as it would distinctly change from being non-seasonal in the past into presenting some seasonal differences in the future.

In short, it appears that the physical relationships between the probability of precipitation and future scenarios of climate change are safeguarded by the standardization in the form of RPS.

Another important aspect of the standardization process is to start from precipitation series with sufficient detail and accuracy of values for achieving more flexible processing. For example, the numerical output of a climate model precipitation has many values below 0.1 mm so it is more detailed in the probability distribution than a meteorological observatory (which in general has a resolution of 0.1 mm).

5.2. Probabilistic transformation

The probability distributions used in this study were selected because they present better results than others for different climates (Moncho et al 2011). In fact, for the Basque Country, the two distributions have a high goodness-of-fit for the daily rainfall, with a mean error generally less than 6% (Fig. 3). But in summer the error of fit is somewhat higher, due to the high number of days with low precipitation, which makes the relative error shoot up.

In any case, the simulated time-series adequately resemble the observed series, e.g. the MNAE is generally less than 5% (Fig. 5 and 6). But that is no guarantee that the probability distributions are similar enough to say that they are statistically indistinguishable. One way to measure this is the Kolmogorov-Smirnov (KS, Marsaglia et al 2003, and Sekhon 2010). The results show that the p-value of most observatories is above 0.05, and it therefore follows that the simulated probability distributions are generally indistinguishable from those observed. Nevertheless, in the summer months nearly half of the observatories have significance less than 0.05 (Fig. 12). This can be explained by the large number of days without precipitation. That is, the KS test gives much consideration to values with a high frequency of occurrence, so small differences in the number of dry days cause the p-value to be close to zero.

One of the most sensitive aspects of the RPS methodology is that it keeps almost intact its wet and dry spells. That is, a probabilistic transformation can change the scale of the precipitation and the number of dry days, but not the disposition of dry days alternating with rainy days. Therefore, it is important that the models can reproduce the observed conditions of wet and dry spells.

In this regard, Figure 13 shows that the regional models (RCMs) adequately reproduce the climatic characteristics of wet spells for the 1961-2000 reference period. Note that in general the projection for 2001-2040 shows no remarkable changes compared to the reference period. However, there is a slight reduction in the length of wet spells in all quarters except winter. This reduction together with the slight increase in the peaks in autumn projected for some observatories implies a slight increase in the index n , as shown in Figure 11.

The reduction of length of wet spells is greater in summer and autumn. In contrast, total precipitation of wet spells decreases more clearly in summer than in autumn, but not in all the observatories. This is because in the summer months the rainfall intensity remains approximately constant (Fig. 12).

As for dry spells, the average length is properly simulated for the reference period (Fig. 14). However, the probability distributions of the length of dry spell have certain differences compared to the observations, but obviously it does not affect the seasonal cycle. For the period of 2001-2040, numerical outputs of RCMs projected a clear increase in the length of dry spells in spring and summer of more than 0.5 days per period without rainfall. Note that the increased length of dry spells is consistent with the slight decrease in the length of wet spells projected for 2001-2040 (Fig. 13).

5.3. Analyses of precipitation change

The increase in heavy precipitation in the Basque Country can be related to the possible mediterraneanisation of climate. The increase in the number of dry days (Fig. 14) and the highest concentration of precipitation in a short time (Fig. 11 and 13) support this hypothesis. The possible mediterraneanisation of Basque climate may be driven by the expansion of the Hadley cell (Lu et al 2007). The global warming projected under the A1B scenario causes a dilation of the cell and a consequent rise north of the subtropical subsidence area. The physical implications of this phenomenon are basically two: 1) Less precipitation in spring due to the increased presence of anticyclones. 2) Increased intensity of autumn precipitation in the Mediterranean area due to increased summer warming and hence more water available in the atmosphere.

6. Conclusions

Both the modified Gumbel distribution and the generalized version of the log-logistic distribution have a high goodness-of-fit for the empirical probability of daily precipitation in the Basque Country. In particular, the latter theoretical distribution showed better results, with a mean error of 4% compared to 5% obtained by Gumbel distribution.

The probabilistic transformation of RCM adequately corrects the probability distribution and makes it indistinguishable from the observatories in the period. In fact, the errors of the simulated series are generally less than 5% for both average

and standard deviation. The error in the number of days without precipitation is less than 1% in most cases.

Moreover, this transformation does not directly affect the climate signal of the models, but just brings it in the line with the frequency of precipitation probability of the observatories. Therefore, the methodology is useful to quantify possible changes in local climate, but not only in reference to the amount of precipitation, but also in the way it is distributed in time (dry spells, wet spells, index n , etc.).

The results of this study show that, during the twenty-first century and under the A1B scenario, we can expect changes in the temporal distribution of rainfall, favouring an intensification. In fact, although the duration of consecutive days with rain may be shortened, the rainfall would be concentrated in narrower time windows, thus increasing the intensity of the daily maximum in autumn. This idea is reinforced because the index n projected for the rest of the century shows a slight increase, from 0.57 to 0.60.

With regard to highest precipitation, most models project a clear increase. For example, precipitation with a return period of 100 years will increase to 30% in 2001-2040 compared to 1961-2000, in some areas of West Basque Country. For the period 2041-2080, the increase may even exceed 40% in most areas. In addition to the intensification of precipitation, which is clearer in the autumn, climate models project an increase in the length of dry spells and a slight decrease in summer precipitation. Therefore, these changes indicate a possible mediterraneanisation of the Basque climate.

Acknowledgments. This work is the last part of the doctoral thesis written in the Department of Earth Physics of the University of Valencia. This work is supported by the Department of Environment, Regional Planning, Agriculture and Fisheries of the Basque Government (K-Egokitzen II project, Etorrek Funding Program). Likewise, we acknowledge the State Meteorological Agency of Spain (AEMET) and Hydrographics Confederations of Ebro (CHE) and Júcar (CHJ) for providing the data for this study. In particular, we thank to José Ángel Nuñez, head of the Department of Climatology AEMET delegation in Valencia, and Margarita Martín, AEMET delegate in the Basque Country, for helpful comments. Finally, it is fair to acknowledge the support of Maddalen Mendizabal (Tecnalia) and especially for raising the issue of probability of daily precipitation.

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Cont* 19: 716-723
- Aleklett K, Hook M, Jakobsson K, Lardelli M, Snowden S, Soderbergh B (2010) The Peak of the Oil Age - Analyzing the world oil production Reference Scenario in World Energy Outlook 2008. *Energy Policy* 38: 1398-1414
- Alves M, Verdière AC (1999) Instability dynamics of a subtropical jet and applications to the Azores Front-Current System-eddy driven mean flow. *J Phys Oceanogr* 29: 837-864

- Benestad RE (2010) Downscaling precipitation extremes. Correction of analog models through PDF predictions. *Theor Appl Climatol* 100: 1–21
- Borja A, Aguirrezabalaga F, Martínez J, Sola JC, García-Arberas L, Gorostiaga JM (2004): Benthic communities, biogeography and resources Management”. In: Borja, A., Collins, M. (Eds.). *Oceanography and Marine Environment of the Basque Country*, Elsevier Oceanography Series 70: 455–492
- Capel-Molina JJ (2000) *El Clima de la Península Ibérica*. Ed. Ariel, Barcelona
- Carlson WB (2011) The Modeling of World Oil Production Using Sigmoidal Functions. *Energy Sources, Part B: Economics, Planning, and Policy* 6: 178–186
- Christensen JH, Hewitson B, Busuioc A, Chen A, Gao X, Held I, Jones R, Kolli RK, Kwon W-T, Laprise R, Magaña Rueda V, Mearns L, Menéndez CG, Räisänen J, Rinke A, Sarr A, Whetton YP (2007) Regional Climate Projections”. In: *Climate Change (2007): The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL (eds.)]. Cambridge University Press, Cambridge, United Kingdom y New York, NY, USA
- Déqué M, Rowell DP, Lüthi D, Giorgi F, Christensen JH, Rockel B, Jacob D, Kjellström E, De Castro M, Van Den Hurk B (2007). An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change* 81: 53–70
- Elshamy ME, Seierstad IA, Sorteberg A (2009). Impacts of climate change on Blue Nile flows using bias-corrected GCM scenarios. *Hydrol Earth Syst Sci* 13: 551–565
- Furrer R, Nychka D, Sain S (2010) fields: Tools for spatial data. R package version 6.3. <http://CRAN.R-project.org/package=fields>
- Garmendia MI, Pérez C, Rodríguez C, Garmendia J (1989) Factores determinantes de la precipitación anual en la vertiente cantábrica. *Acta Samaltina de Ciencias* 67: 113–117, 3 Ref
- Goubanova K, Li L (2007) Extremes in temperature and precipitation around the Mediterranean basin in an ensemble of future climate scenario simulations”. *Global and Planetary Change* 57: 27–42
- Green PJ, Silverman BW (1994) *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall, London
- Ines AVM, Hansen JW (2006) Bias correction of daily GCM rainfall for crop simulation studies. *Agr For Met* 138: 44–53
- Lana X, Martínez MD, Serra C, Burgueño A (2004) Spatial and temporal variability of the daily rainfall regime in Catalonia (northeastern Spain), 1950–2000. *Int J Clim* 24: 613–641

- Lu J, Vecchi GA, Reichler T (2007) Expansion of the hadley cell under global warming. *Geophys Res Let*, 34: L06805
- Martín-Vide, J (2004) Spatial distribution of a daily precipitation concentration index in peninsular Spain. *Int J Clim* 24: 959-971
- Marsaglia G, Tsang WW, Wang J (2003) Evaluating Kolmogorov's distribution. *J Stat Software* 8/18
- Meehl GA, Stocker TF, Collins WD, Friedlingstein P, Gates AT, Gregory JM, Kitoh A, Knutti R, Murphy JM, Noda A, Raper SCB, Watterson IG, Weaver AJ, Zhao Z-C (2007) Global Climate Projections. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate*
- Moncho R, Chust G, Caselles V (2011) An alternative model for precipitation probability distribution: application to Spain. *Clim Res* (in review)
- Moncho R, Belda F, Caselles V (2009a) Climatic study of the exponent n of the IDF curves of the Iberian Peninsula. *Tethys* 6: 3-14
- Moncho R, Chust G, Caselles V (2009b) Análisis de la precipitación del País Vasco en el período 1961-2000 mediante reconstrucción espacial. *Nimbus*, nº 23-24: 149-170
- Niehörster F, Fast I, Huebener H, Cubasch U (2008) The stream one ENSEMBLES projections of future climate change. *Ensembles Technical Report* nº 3, 2008
- Olcina J (2008) Riesgo de inundaciones y ordenación del territorio en España. *Documents d'Anàlisi Geogràfica* 51: 178-180
- Palmer TN, Doblas-Reyes FJ, Weisheimer A, Rodwell MJ (2008) Toward seamless prediction: Calibration of climate change projections using seasonal forecasts. *Bull Am Met Soc* 89: 459-470
- Peliz A, Rosa TL, Santos AMP, Pissarra JL (2002) Fronts, jets, and counterflows in the Western Iberian upwelling system. *J Mar Syst* 35: 61-77
- Piani C, Haerter JO, Coppola E (2010) Statistical bias correction for daily precipitation in regional climate models over Europe. *Theor Appl Climatol*, 99:187-192
- R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Raue A, Kreutz C, Maiwald T, Bachmann J, Schilling M, Klingmüller U, Timmer J (2009) Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* 25: 1923-9
- Sekhon JS (2010) Matching: Multivariate and Propensity Score Matching with Balance Optimization. R package version 4.7-11. URL <http://CRAN.R-project.org/package=Matching>.

TABLES

Table 1. Correlation (R^2) and slope of the climatic values of the average rainfall, daily deviation and dry days in the series simulated by the models, compared with baseline observatories.

RCM	Correlation			Slope		
	Mean rainfall	Standard deviation	Dry days	Mean rainfall	Standard deviation	Dry days
Aladin-ARPEGE	0.99	0.98	0.97	1.07	0.97	0.99
CLM-HadCM3Q0	0.99	0.96	0.97	1.07	0.95	1.01
HIRHAM-HadCM3Q0	0.99	0.98	0.98	1.14	1.00	0.98
HIRHAM-BCM	0.99	0.98	0.97	1.05	0.97	0.99
RACMO-ECHAM5-rt3	0.99	0.97	0.97	1.08	0.99	1.02
REMO-ECHAM5-rt3	0.99	0.97	0.97	1.08	1.01	0.96
PROMES-HadCM3Q0	0.99	0.97	0.96	1.11	0.97	0.96

Figure captions

Figure 1. a) Location of the observatories used in this study (AEMET). b) Time series of observatories.

Figure 2. a) Example of digital terrain model used in the regional climate models of ENSEMBLES-rt3. b) Election of interest grid points for the Basque Country, which distinguishes between sea points (dark grey) and ground points (grey). c) Time series of spatial average of 7 regional climate models.

Figure 3. Comparison of Mean Normalized Absolute Error (MNAE) of the two probabilistic models used for the standardization of the 45 observatories with at least 40 years of data.

Figure 4. Example of transformation of the return period curve of daily rainfall (for the year) of seven RCMs. Top left shows the changing curves (before and after processing) of the seven series that attempt to simulate the precipitation of Lasarte (Guipúzcoa). Top right shows the final result to simulate the precipitation of Aranzazu (Guipúzcoa), whereas below the same thing happens to Salinas de Añana (Alava), Punta Galea (Vizcaya), from left to right. The symbol Δd refers to the difference in the frequency of days without precipitation.

Figure 5. Relative mean absolute error (RMAE) and relative bias (RBIAS) of the average and standard deviations of simulated time series. Each bar represents the set of all observatories according to each of the seven RCMs.

Figure 6. Absolute and relative mean error (MAE) and bias (BIAS) of the days without precipitation of the simulated series. Each bar represents the set of all observatories according to each of the seven RCM.

Figure 7. Comparison between observed and simulated values of monthly averages of all observatories.

Figure 8. Projection of the relative change of average precipitation for the quarters of the year and period of 2001-2040 vs 1961-2000. The black lines represent the confidence intervals estimated from the statistics of the seven RCMs.

Figure 9. Projection of the maximum daily rainfall for a return period of 100 years according to the theoretical curves fitted to the simulated observatories. Above shows the projection of the absolute value for the three periods (1961-2000, 2001-2040 and 2041-2080). Below shows the relative change for 2001-2040 and 2041-2080 and the error of the simulation compared with observed time series (1961-2000).

Figure 10. a) Maximum rainfall in one day for 1961-2000 simulated with a return period of 100 years. The values are estimated according to the theoretical curves fitted to the simulated observatories, and they are interpolated by using a multiple spatial adjustment. b) Relative change of precipitation described in a), for the

period 2001-2040 versus 1961-2000. The black lines represent the confidence intervals estimated from the statistics of the seven RCMs.

Figure 11. a) Empirical probability distribution of the index n to the set of all the observatories in the period 1961-2000, and the range corresponding to the seven RCM simulations. b) Projection of change for 2001-2040 and 2041-2080. c) and d) Respectively, the same as a) and b) but only the rainfall exceeded 50 mm in one day.

Figura 12. P-value of Kolmogorov-Smirnov test of observed versus simulated precipitation for the 45 observatories with at least 40 years.

Figure 13. Comparison of observed and simulated wet spells for 1961-2000 and projected for 2001-2040. Each point of the boxplot represents the climatic value of a real or simulated observatory. Top left shows the average length of spells (days), top right, it shows the average daily peak (in millimeters). Bottom right shows the average total rainfall (in millimeters) and lower left compares the mean index n of wet spell.

Figure 14. Evolution of dry spells for the period of 2001-2040 versus the period 1961-2000, both observed and simulated: a) average length of dry spells, and b) cumulative probability of the length of dry spells.

Figure 1

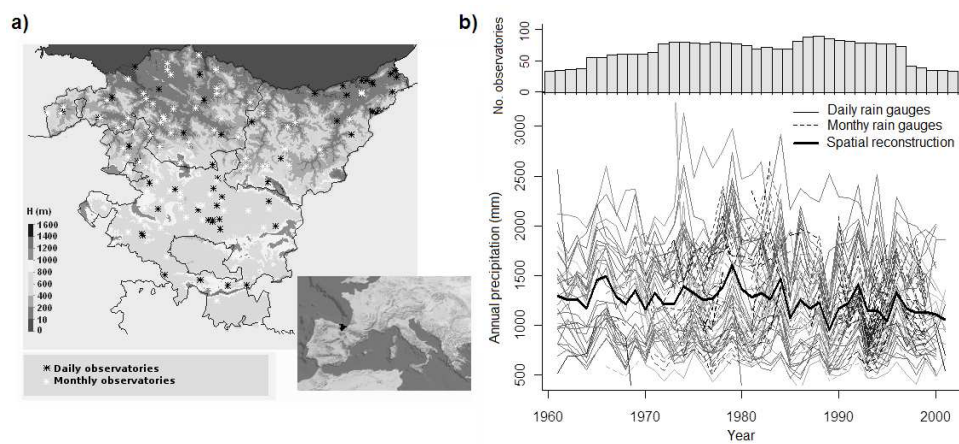


Figure 2

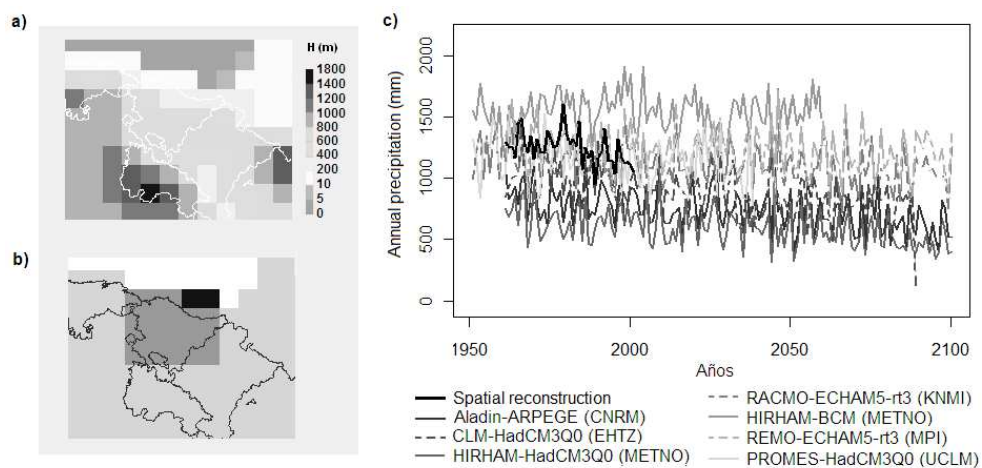


Figure 3

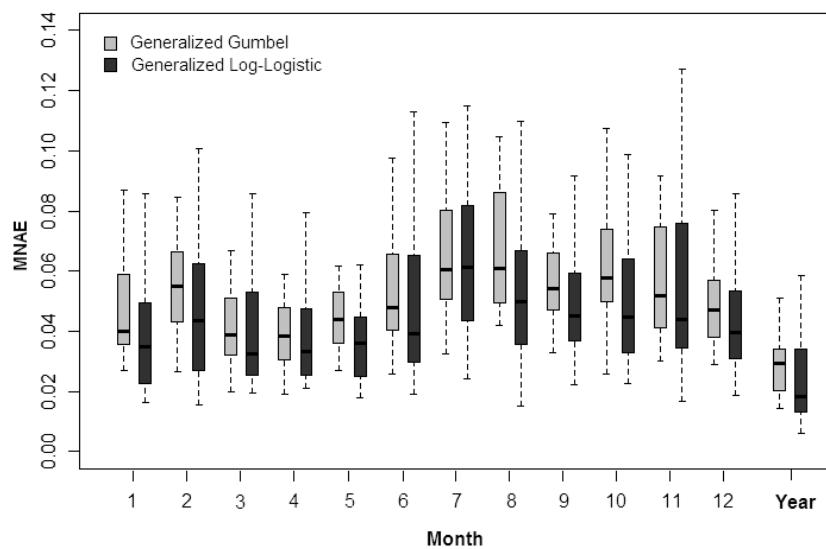


Figure 4

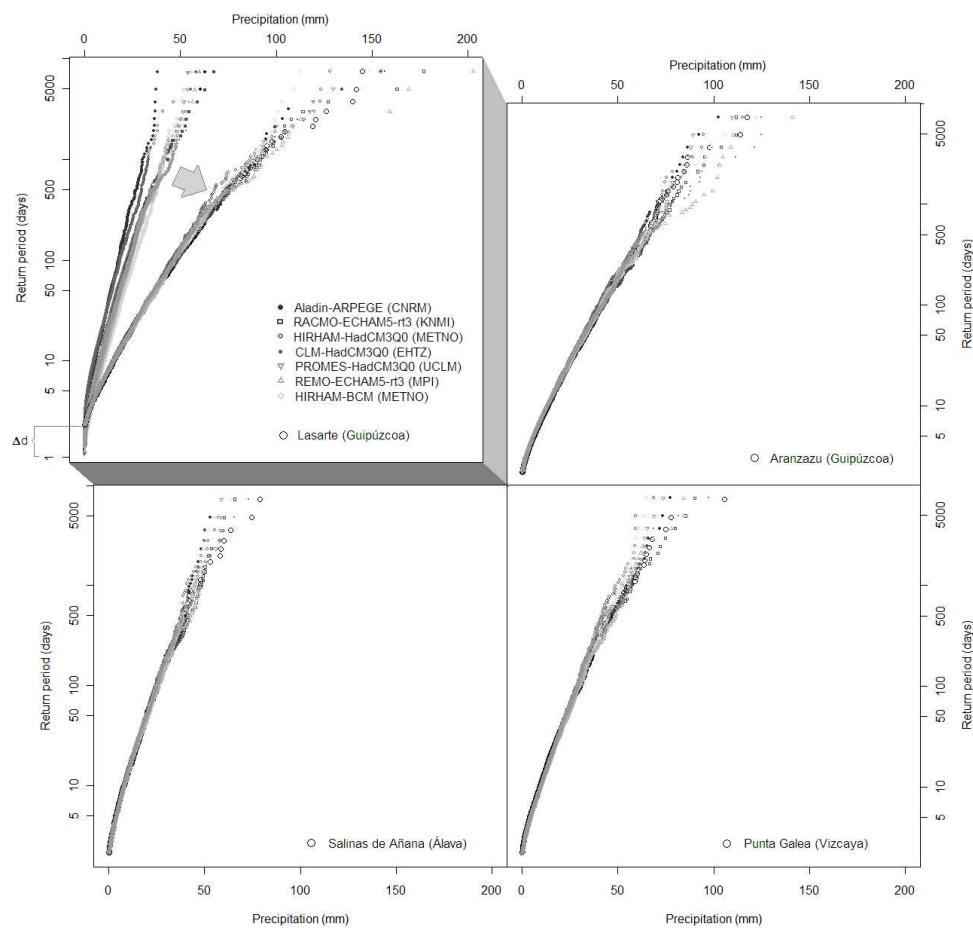


Figure 5

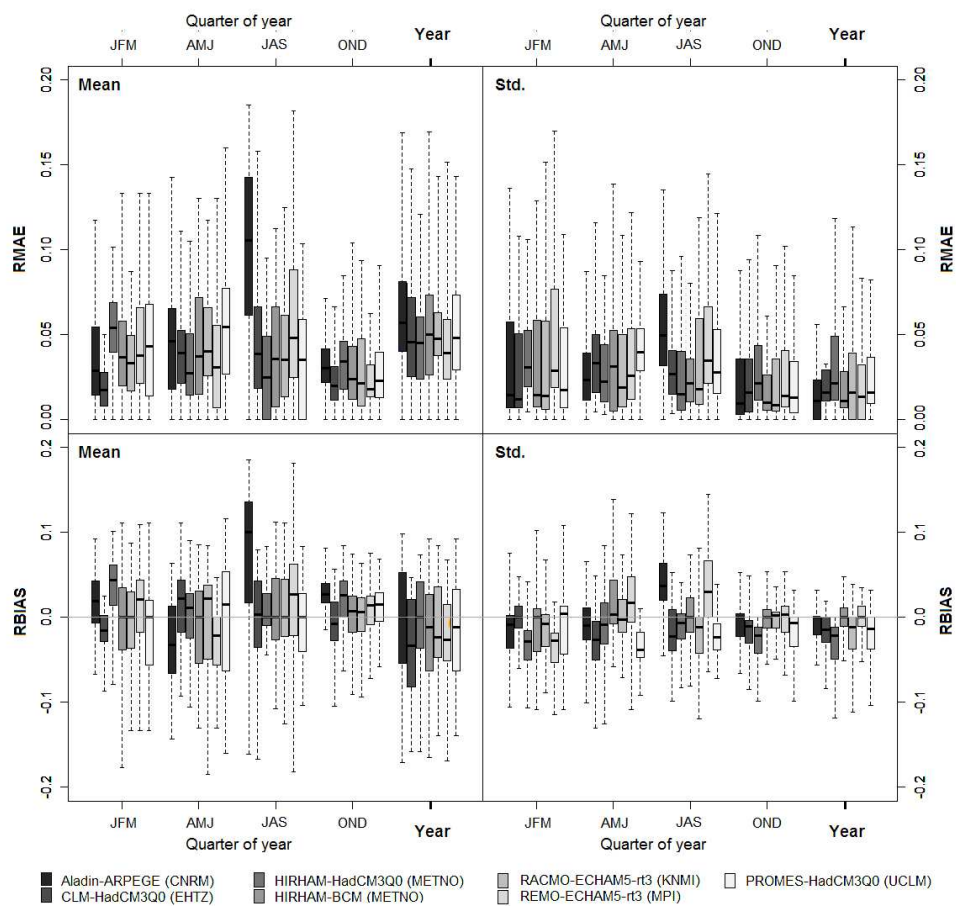


Figure 6

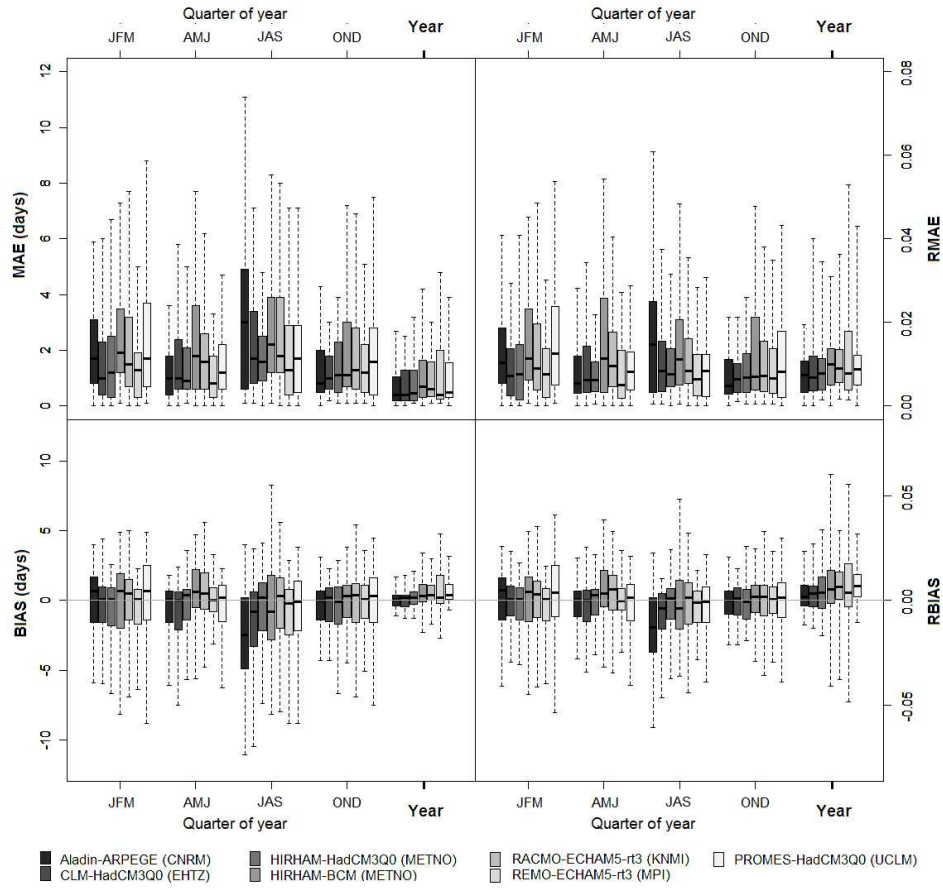


Figure 7

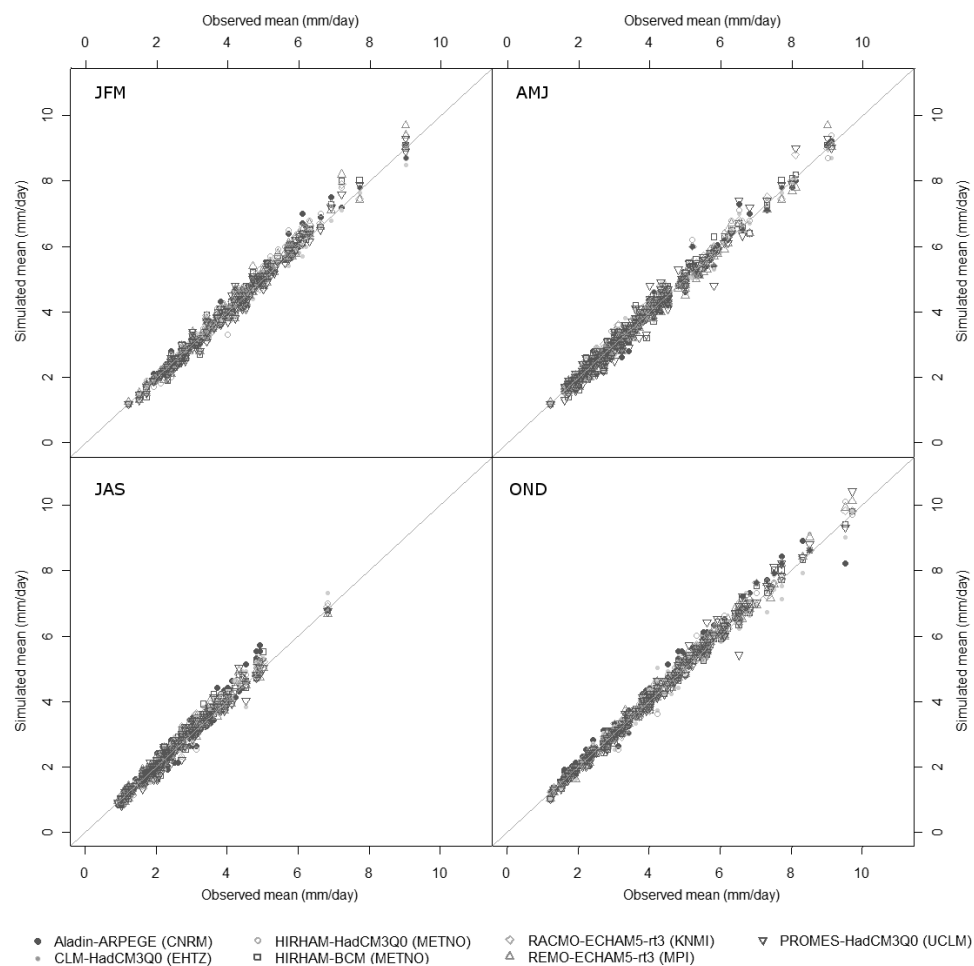


Figure 8

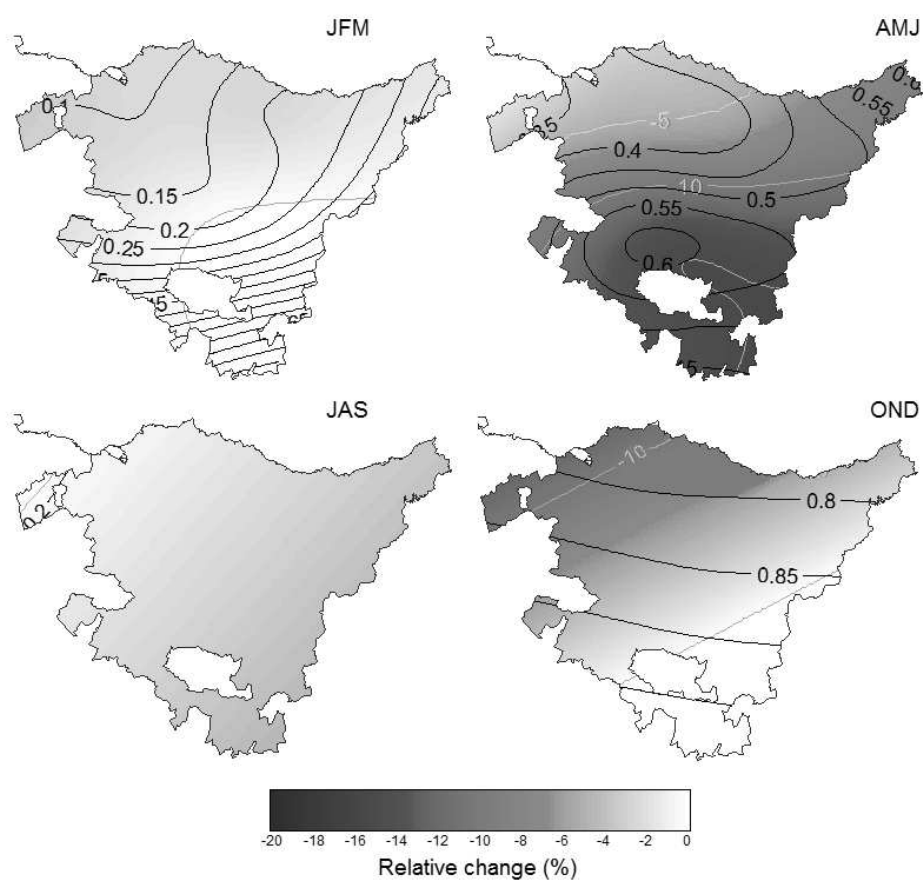


Figure 9

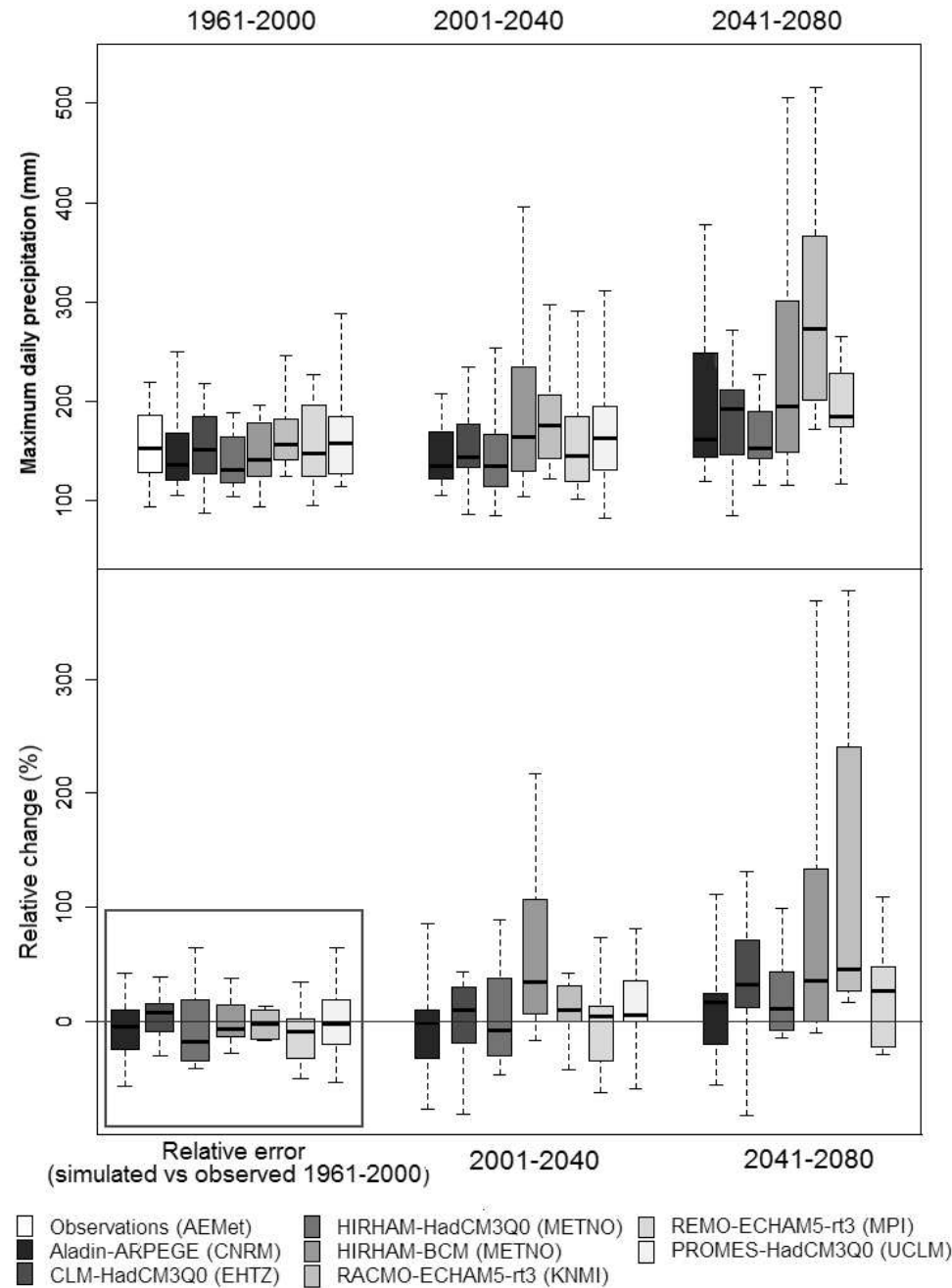


Figure 10

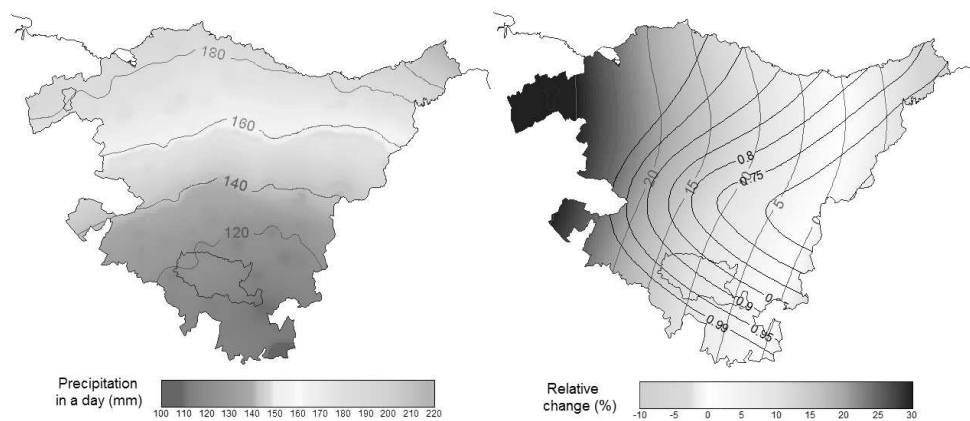


Figure 11

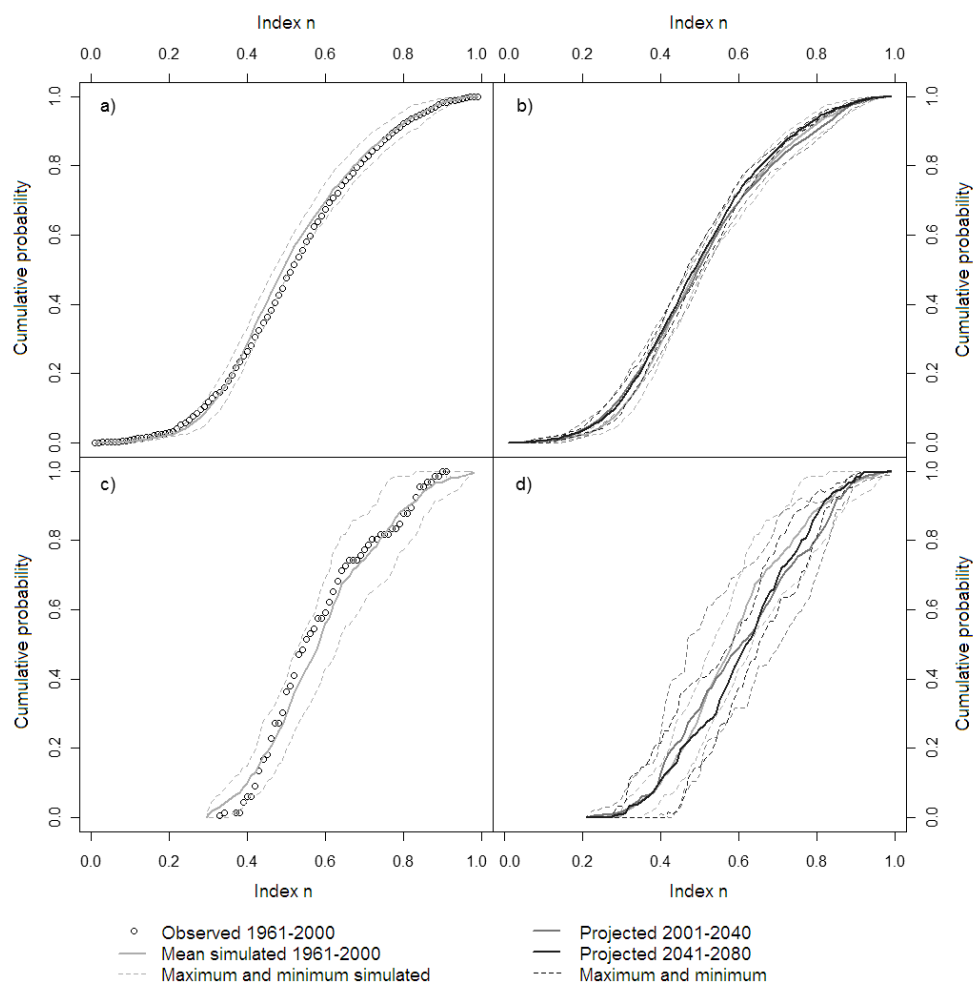


Figure 12

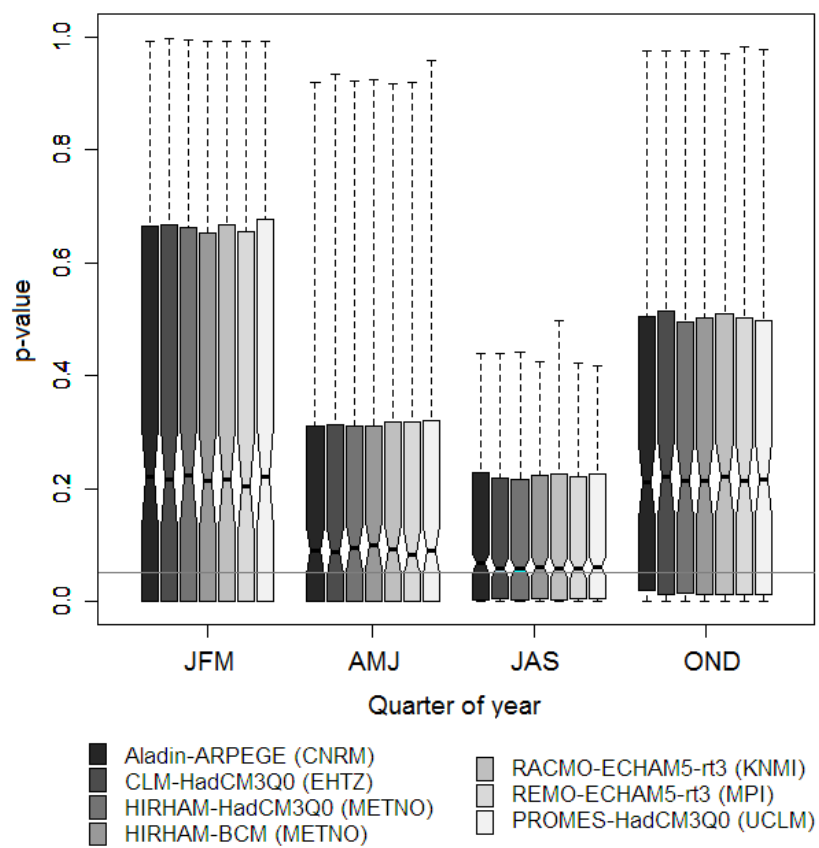


Figure 13

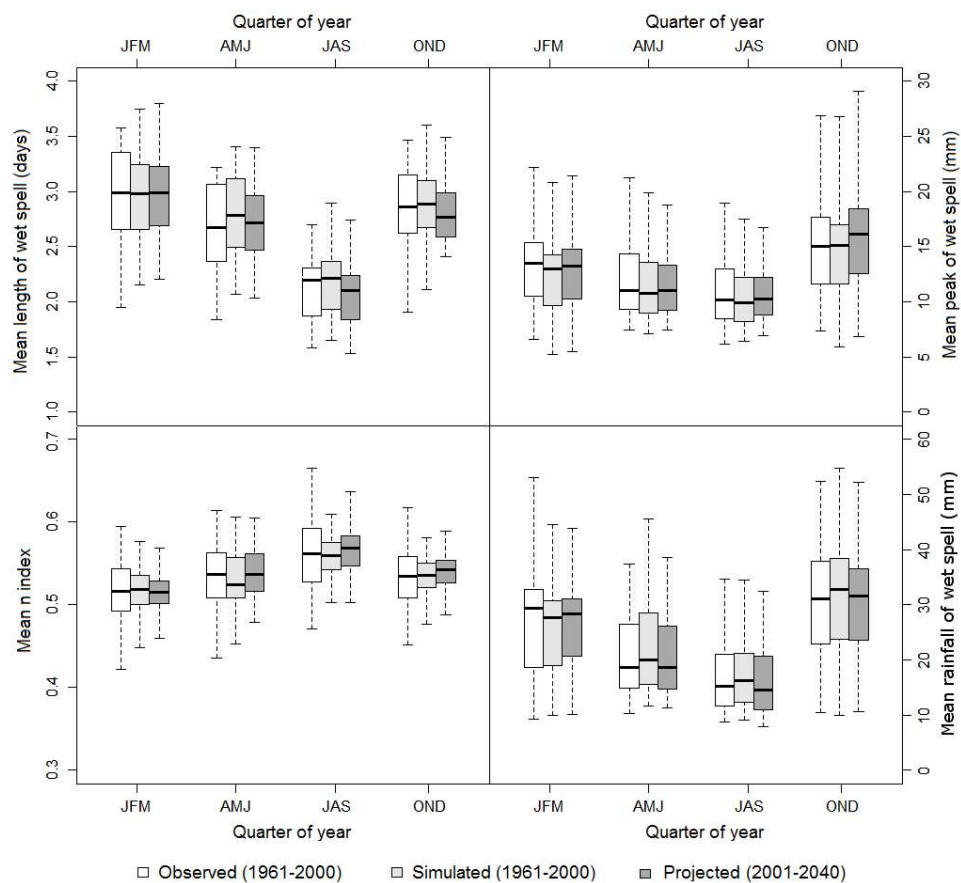


Figure 14

