

MÉTODO DE REGIONALIZACIÓN DE PRECIPITACIÓN BASADO EN ANÁLOGOS.

EXPLICACIÓN Y VALIDACIÓN

Santos Eduardo Petisco de Lara

Área de Evaluación y Modelización del Cambio Climático

AGENCIA ESTATAL DE METEOROLOGÍA



2008

Catálogo de publicaciones oficiales
<http://www.060.es>

*Nota Técnica 3A del Área de Evaluación y
Modelización del Cambio Climático (NT AEMCC-3A)*
Agencia Estatal de Meteorología

Edita: Centro de Publicaciones
Secretaría General Técnica
Ministerio de Medio Ambiente, y Medio Rural y Marino ©

ISBN: 978-84-8320-447-4
NIPO: 310-08-061-7
Depósito Legal: M-20504-2008
Imprime: Imprenta de la Agencia Estatal de Meteorología

Índice

RESUMEN	v
1. INTRODUCCIÓN	1
2. DATOS UTILIZADOS	2
2.1. Datos de calibración y validación	2
2.1.1. Datos atmosféricos	2
2.1.2. Datos de precipitación	2
2.2. Datos de aplicación de la metodología	3
3. METODOLOGÍA DE LA ESTIMACIÓN DE PRECIPITACIÓN	4
3.1. Primera fase. Determinación de la información de base general	4
3.1.1. Obtención de tipos sinópticos	4
3.1.2. Obtención de análogos de cada tipo sinóptico	5
3.1.3. Obtención de predictores significativos para cada tipo sinóptico en cada punto	5
3.2. Segunda fase. Cálculo de los datos base del día problema y estimación de la precipitación	6
3.2.1. Determinación de los campos de viento geostrófico	6
3.2.2. Obtención del conjunto de análogos del día problema	6
3.2.3. Determinación del tipo sinóptico y de los predictores adecuados	6
3.2.4. Búsqueda del grupo definitivo de análogos a usar en la estimación	7
3.2.5. Estimación de la precipitación	7
4. VALIDACIÓN DE LAS ESTIMACIONES DE PRECIPITACIÓN	8
4.1. Validación de las medias anuales	8
4.2. Validación de las medias mensuales	10
4.3. Validación de las estimaciones de precipitación diaria	16
4.3.1. Comparación de percentiles	16
4.3.2. Índices de verificación. Errores	23
4.3.3. Índices de verificación. <i>Skill-scores</i>	26
5. CONCLUSIONES	29
BIBLIOGRAFÍA	30

Resumen

En el presente trabajo se explica y se valida el método de regionalización estadística basado en análogos, que se ha utilizado por la AEMET en la primera fase de la generación de escenarios regionalizados de cambio climático para España.

Se trata de obtener estimaciones de precipitación sobre una rejilla de alta resolución de 203 puntos que cubre la España peninsular y Baleares, basándose en datos en baja resolución de reanálisis o modelos globales.

La regionalización se realiza obteniendo un conjunto de análogos sinópticos al día de estudio, tomados en el período de referencia 1961-1990.

La estimación de precipitación se realiza calculando la media ponderada de la precipitación en el punto de estima, de un conjunto de análogos al día de estudio, obtenidos buscando la similitud sinóptica y la de un grupo de predictores calculados en la zona de estima; los predictores adecuados dependen del tipo sinóptico del día estudiado y se seleccionan entre los siguientes:

- Geopotencial, vorticidad y viento geostrófico en 1000 y 500hPa
- Intensidad del viento térmico 500/1000
- Tendencia del geopotencial en 1000hPa

La validación se realiza mediante la comparación de los datos observados en un período, con los correspondientes estimados habiendo calibrado el método en un período distinto.

En primer lugar se evalúa el comportamiento de las medias anuales y mensuales. En segundo lugar se evalúa el comportamiento de las estimaciones diarias mediante la comparación de percentiles significativos y mediante el cálculo de errores medios de estimación. Además, se analizan, mediante *skill-scores*, las mejoras del método respecto de las estimaciones basadas en los valores climatológicos mensuales.

1. INTRODUCCIÓN

Existe una demanda creciente de proyecciones regionalizadas de cambio climático consecuencia de los distintos escenarios de emisión a la atmósfera de gases de efecto invernadero, para ser utilizadas como entradas en modelos de análisis de impactos en ecosistemas y sectores económicos y sociales sensibles a las condiciones climáticas.

Todas estas proyecciones se construyen a partir de la información suministrada por modelos climáticos globales que tratan de reproducir el comportamiento del sistema climático en función de las concentraciones de gases de efecto invernadero en la atmósfera, derivadas de diversos escenarios de emisión.

No obstante, hay que tener en cuenta que la actual resolución de los modelos climáticos no es suficientemente fina como para suministrar datos climáticos en alta resolución espacial o en puntos geográficos localizados, que en realidad son los que interesan en un análisis de impactos. Se hace, pues, necesario emplear técnicas o procedimientos que permitan inferir, en base a los datos de baja resolución mejor simulados por los modelos climáticos, información localizada de parámetros de interés para el analista de impactos. Esto se trata de conseguir mediante las técnicas de mejora de resolución o regionalización o *downscaling* según la nomenclatura inglesa.

Existen dos tipos básicos de *downscaling*: el ***downscaling* estadístico** y el ***downscaling* dinámico**. El primero se basa en encontrar relaciones empíricas y/o estadísticas entre las variables mejor simuladas (presión, geopotencial, temperatura o parámetros derivados) en baja resolución por los modelos climáticos, y las variables de interés en la zona de estudio. Estas relaciones se buscan en una base de datos de referencia observados y después se aplican a los datos simulados por los modelos obteniéndose así estimaciones de las variables climáticas en la zona de interés, siempre bajo la hipótesis de que las relaciones obtenidas siguen manteniéndose válidas. El ***downscaling* dinámico** se basa en el aumento de resolución de los modelos en el área donde se sitúan las zonas de estudio. Esto puede hacerse aumentando directamente la resolución en dicha área, o bien anidando en el modelo climático de baja resolución un modelo de área limitada de alta resolución. La aplicación conjunta de ambos procedimientos (dinámico y estadístico) en un ***downscaling* dinámico-estadístico** puede aprovechar las ventajas de los dos tipos de *downscaling*. En nuestro caso la obtención de escenarios ha sido llevada a cabo mediante procedimientos de *downscaling* estadístico que serán explicados y validados en los siguientes apartados.

La precipitación es una variable de gran interés para el estudio de impactos del cambio climático. Sin embargo, su comportamiento estadístico se aleja bastante de la distribución normal, especialmente en el caso de las precipitaciones diarias, por lo que es preferible la utilización de métodos no basados en relaciones lineales como, por ejemplo, los fundamentados en la obtención previa de análogos sinópticos. Una metodología de este tipo es la que se utiliza en el presente trabajo. En cualquier caso, las estimaciones de precipitación vendrán acompañadas de mayor incertidumbre debido a la influencia de factores locales o mesoescalares no completamente controlables sinópticamente.

2. DATOS UTILIZADOS

En cualquier regionalización basada en *downscaling* estadístico se utilizan dos tipos fundamentales de datos: 1) Datos de referencia y de calibración de la metodología, y 2) Datos de aplicación de la metodología para la estimación de las variables deseadas. En lo que sigue se describen los datos e información utilizados en nuestro caso.

2.1. Datos de calibración y validación

Estos datos sirven como base de la metodología del *downscaling* que se va aplicar y para realizar su validación. Consta, a su vez, de una base de datos de geopotencial, y de su correspondiente base de datos de precipitación.

2.1.1. Datos atmosféricos

Esta es la que denominaremos **base sinóptica de referencia** y está constituida por datos de geopotencial en 1000 y 500 hPa sobre rejilla latitud-longitud y resolución $2,5^\circ \times 2,5^\circ$, obtenidos a partir del reanálisis del NCAR (*National Center for Atmospheric Research*), (KALNAY, 1996). Se obtienen los campos de geopotencial medios diarios promediando los campos de 00h, 06h, 12h, 18h y 24h (00h del día siguiente), ponderando los datos extremos del día (00h y 24h) la mitad que los restantes. De esta rejilla de geopotencial medio se obtienen las componentes *u* y *v* del viento geostrofico con la misma resolución y para el área de trabajo que abarca de 30°N a 55°N y de $27,5^\circ\text{W}$ a 15°E y un período de 30 años entre el 1 de enero de 1961 y el 31 de diciembre de 1990 con un total de 198 puntos de rejilla. Esta rejilla la denominaremos de aquí en adelante rejilla sinóptica y puede verse en la figura 1 donde también se muestran las ponderaciones asociadas a cada punto de rejilla y que se utilizarán posteriormente para el cálculo de distancias sinópticas.

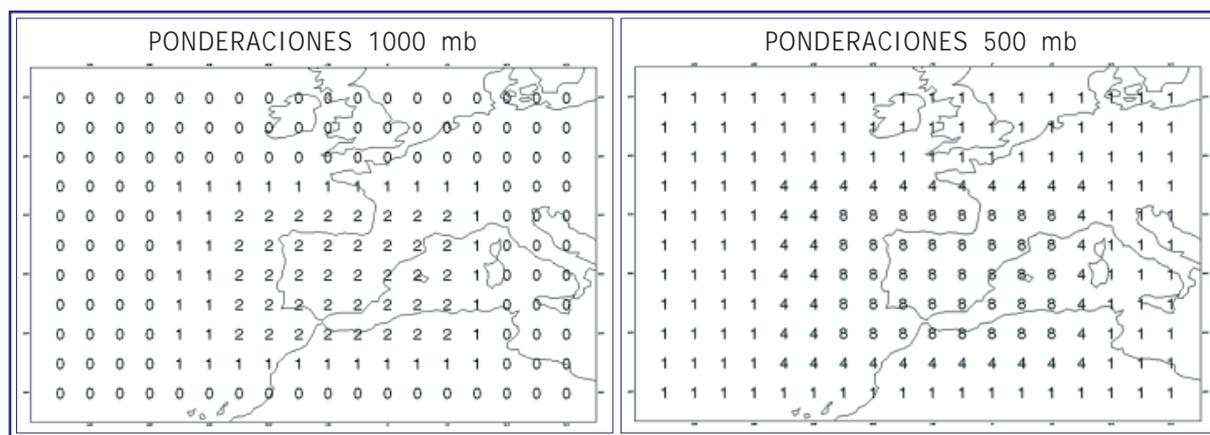


Figura 1. Rejilla sinóptica de baja resolución mostrando las ponderaciones asociadas a cada punto para el cálculo de distancias sinópticas.

2.1.2. Datos de precipitación

Comprende los campos de precipitación asociados a cada uno de los días de la base de datos atmosféricos anteriormente citados. Los datos se presentan en una rejilla de alta resolución con 203 puntos espaciados meridianamente y zonalmente aproximadamente 50 y 60km respectivamente y distribui-

dos por el territorio de la España peninsular y Baleares (figura 2). Esta rejilla fue obtenida por BORÉN y RIBALAYGUA (1997) a partir del banco de datos del Instituto Nacional de Meteorología interpolando los datos diarios de los observatorios disponibles próximos a los puntos. Para el cálculo de la precipitación en el punto se determinó la media ponderada de dichos datos de observatorios siendo las ponderaciones las inversas de las distancias observatorio-punto y tomando las distancias inferiores a 10km iguales a 10km en cualquier caso; de esta manera el dato de cada punto de la rejilla viene a ser una representación promedio de su área de influencia. Las series así obtenidas para cada punto constituyen una base de datos sin ninguna laguna.

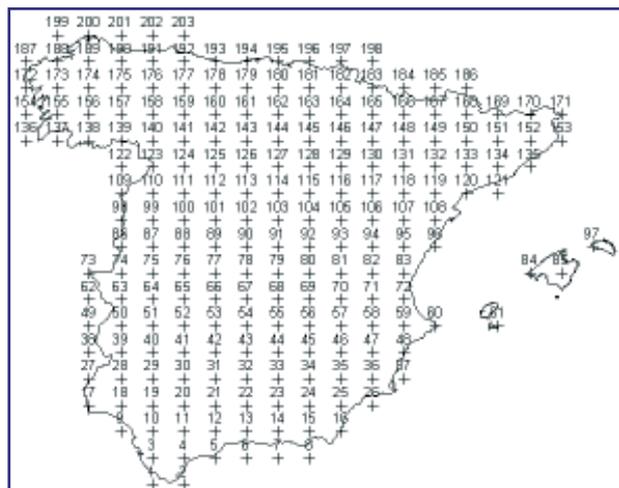


Figura 2. Rejilla de puntos de alta resolución donde se calculan las estimaciones de precipitación.

2.2. Datos de aplicación de la metodología

Se trata del conjunto de datos a los que se va aplicar la metodología de *downscaling*, desarrollada con el objeto de estimar las precipitaciones correspondientes. Estos datos son los campos de geopotencial medio diario sobre la misma rejilla de los campos sinópticos de referencia, dados por los diversos modelos globales de los que se pretende obtener la información regionalizada.

3. METODOLOGÍA DE LA ESTIMACIÓN DE PRECIPITACIÓN

El objetivo es estimar la precipitación en cada uno de los 203 puntos de la rejilla de alta resolución definida anteriormente.

El método de estimación de las precipitaciones se basa en un método de análogos: se trata de buscar un conjunto de días en la base de referencia que tengan la mayor analogía sinóptica con el día problema junto con la similitud de un conjunto de valores de diversos predictores seleccionados según el tipo sinóptico de que se trate y el punto en el que se va a estimar la precipitación. (PETISCO y otros, 2005).

3.1. Primera fase. Determinación de la información de base general

En una primera fase se determinan los parámetros (predictores) significativos para cada tipo sinóptico en cada punto de estima, los cuales quedarán definitivamente establecidos como base para cualesquiera estimaciones que posteriormente queramos realizar. En esta fase se trabaja sobre los datos de referencia tanto sinópticos como de precipitaciones asociadas y consta de las siguientes etapas:

3.1.1. Obtención de tipos sinópticos

Se busca un conjunto de tipos sinópticos que representen la circulación atmosférica en la zona de trabajo para el período de referencia considerado (1961-1990). Para ello se aplica un análisis *cluster* en dos etapas: en la primera etapa se buscan unos "días patrón" mediante el procedimiento de Lund (LUND, 1963); estos "días patrón" se utilizan como "semillas" base de una clasificación por el método de las *k-medias* (ANDERBERG, 1973) y los promedios de los grupos así obtenidos (mapas medios de 1000 y 500 hPa de los elementos de cada grupo) definirán los tipos sinópticos representativos de la circulación atmosférica en la zona de trabajo. La similaridad utilizada para construir los grupos y tipos, es el promedio de las distancias euclídeas entre los campos de valores tipificados de componentes del viento geostrófico en 1000 y 500 hPa. La tipificación se realiza, en cada punto de rejilla, restando la media correspondiente y dividiendo por la desviación tipo. Estas distancias euclídeas se calculan introduciendo ponderaciones en cada punto de rejilla siendo mayores cuanto más nos aproximemos a la zona de la Península Ibérica y Baleares y diferentes para los campos de 1000 hPa y 500 hPa como puede verse en la figura 1; este esquema de ponderaciones se mantendrá en etapas posteriores cuando se calculen distancias entre los campos sinópticos en la zona de trabajo.

La distancia euclídea entre los campos tipificados de cada componente c de dos elementos m y n , donde c representa una componente cualquiera $(u_{1000}, v_{1000}, u_{500}, v_{500})$ del viento geostrófico, vendría dada por la expresión [1].

$$d_c = \sqrt{\frac{\sum_1^n w_i (c_{mi} - c_{ni})^2}{\sum w_i}} \quad [1]$$

en donde:

c_{mi} y c_{ni} representan el valor de la componente c tipificada de viento geostrófico en el punto de rejilla i para dos situaciones sinópticas m , n cualesquiera cuya distancia se quiere calcular;

w_i es la ponderación correspondiente al punto de rejilla i ;

$n = 198$ es el número total de puntos de rejilla del campo.

Una vez calculadas las cuatro distancias d_c correspondientes a cada uno de los cuatro campos de componentes de viento geostrófico, se toma la media de ellas como distancia D_s entre los elementos considerados.

Los requisitos de similaridad utilizados para la obtención de los "días patrón" de la primera etapa del análisis son suficientemente rigurosos como para obtener finalmente un número de tipos bastante amplio que represente de una forma bastante exhaustiva los distintos rasgos de la circulación, dado el objetivo que se pretende de estimar precipitaciones.

3.1.2. Obtención de análogos de cada tipo sinóptico

Para cada tipo sinóptico se determinan los 150 elementos de la base de referencia sinóptica más análogos al tipo en cuestión. Para ello igualmente se calculan las distancias euclídeas entre los campos de componentes de viento geostrófico tipificadas. Así, la distancia euclídea entre los campos tipificados de cada componente c del viento geostrófico vendría dada por la expresión [2] que, en este caso representa la distancia entre un tipo sinóptico y una situación sinóptica de la base de referencia:

$$d_c = \sqrt{\frac{\sum_1^n w_i (c_{ti} - c_i)^2}{\sum w_i}} \quad [2]$$

en donde:

c_i representa el valor de la componente c tipificada de viento geostrófico en el punto de rejilla i para una situación sinóptica cualquiera de la base de referencia;

c_{ti} representa igualmente el valor de la componente c tipificada de viento geostrófico en el punto de rejilla i pero, en este caso, para el tipo sinóptico t_i ;

w_i es la ponderación correspondiente al punto de rejilla i ;

$n = 198$ es el número total de puntos de rejilla del campo.

Una vez calculadas las cuatro distancias d_c correspondientes a cada uno de los cuatro campos de componentes de viento geostrófico consideradas, se toma la media de todas ellas como distancia D_s entre el tipo sinóptico y el elemento considerado. Hecho esto para todos los elementos de la base de referencia sinóptica tomamos los 150 elementos más análogos, es decir, los 150 de menor distancia D_s , que constituirían el grupo de análogos asociado al tipo en cuestión y que servirán de base para la determinación de predictores.

3.1.3. Obtención de predictores significativos para cada tipo sinóptico en cada punto

Se trata ahora de obtener los predictores que puedan tener una influencia significativa sobre la precipitación para cada tipo sinóptico de los anteriormente definidos, y en cada uno de los puntos de estimación. Hay que decir que, dado que la metodología se va a aplicar a salidas de modelos climáticos, sólo consideraremos como posibles predictores aquellos relacionados con las variables mejor simuladas por los modelos.

Teniendo esto presente, consideramos como conjunto de posibles predictores p_i que pueden influir sobre la precipitación a los diez siguientes:

- El valor del geopotencial en 1 000 y 500 hPa: p_1, p_2 .
- La tendencia del geopotencial en 1000 hPa: p_3 .
- El valor de la vorticidad relativa en 1000 y 500 hPa: p_4, p_5 .
- Las componentes del viento geostrófico en 1000 y 500 hPa: p_6, p_7, p_8, p_9 .
- La intensidad del viento térmico 1000/500: p_{10} .

La introducción de parámetros de humedad como predictores (como, por ejemplo, la humedad específica en 700hPa), se traduce en una ligera mejora de las estimaciones, aunque hay que advertir que la humedad no es de las variables mejor simuladas por los modelos. En la validación que sigue no se ha tenido en cuenta, por este motivo, el uso de este predictor. No obstante, al aplicar la metodología para regionalizar las proyecciones de los modelos globales en un contexto de cambio climático, podría ser conveniente, pese a todo, introducir predictores de humedad.

Los valores de estos predictores se consideran en el punto de la rejilla sinóptica más próximo al punto de estima de la precipitación.

A continuación, consideramos, para cada tipo, su grupo de análogos sinópticos obtenidos anteriormente y tomamos cada uno de los puntos de estima. La precipitación en el punto de estima para cada uno de los análogos del tipo la tomamos como predictando, y como predictores tomamos los valores de los parámetros p_i anteriormente mencionados, calculados, para cada análogo, en el punto de la rejilla sinóptica más próximo al punto de estima. Se procede ahora a una selección de predictores significativos aplicando un método de regresión paso a paso; de esta manera tendremos **para cada tipo sinóptico y para cada uno de los puntos de estima, el conjunto de predictores, con sus correlaciones parciales, más apropiado para usar como apoyo en la estimación de la precipitación**. Esta es una información básica, **establecida para siempre**, y que se usará después en **cualquier estimación**. Es preciso tener en cuenta que el proceso anterior se realiza con el único objeto de seleccionar aquellos predictores que podrían explicar, en cierto grado, la precipitación asociada al tipo sinóptico y al punto de estima considerados; no se trata pues, de construir una ecuación de regresión con la que luego estimar precipitaciones.

La robustez de la selección de predictores ha sido comprobada obteniendo, por el mismo procedimiento, los predictores asociados a los grupos de análogos de elementos muestra pertenecientes a los diversos tipos.

3.2. Segunda fase. Cálculo de los datos base del día problema y estimación de la precipitación

Supongamos ahora que tenemos un día determinado con su campo sinóptico asociado definido por el geopotencial en 1000 y 500 hPa en nuestra rejilla sinóptica de trabajo. El objetivo es estimar la precipitación asociada a ese campo sinóptico en el conjunto de puntos de estima (en nuestro caso la rejilla de alta resolución de 203 puntos definida al principio); para ello se realizan los siguientes procesos:

3.2.1. Determinación de los campos de viento geostrófico

Para nuestro día problema, determinamos las componentes u y v del viento geostrófico en la rejilla de trabajo y en los dos niveles considerados.

3.2.2. Obtención del conjunto de análogos del día problema

Se determinan N elementos de la base sinóptica de referencia análogos sinópticamente a nuestro día problema. Para ello se calcula el promedio de las distancias euclídeas de los campos de valores tipificados de componentes del viento geostrófico, en 1000 y 500 hPa, definidos sobre la rejilla anteriormente mencionada, entre el día problema y los días de la base de referencia. La expresión que nos da la distancia euclídea entre los campos tipificados de cada componente del viento geostrófico entre el día problema y una situación sinóptica de la base de referencia es similar a la [1] anteriormente definida.

Una vez calculadas las cuatro distancias correspondientes a cada uno de los cuatro campos de componentes de viento geostrófico tipificadas consideradas, se toma la media de ellas como distancia d_s , seleccionando aquellos N análogos cuya distancia d_s al día problema esté por debajo de la distancia umbral límite considerada, no obstante, en cualquier caso, no se utiliza un número de análogos superior a 30 y siempre se emplea al menos un análogo.

3.2.3. Determinación del tipo sinóptico y de los predictores adecuados

Se determina el tipo sinóptico de referencia al que pertenecería nuestro día problema. De esta manera sabríamos los predictores significativos para la estimación de la precipitación para cada punto de estima y calcularíamos los valores concretos de esos predictores tipificados, para el día problema y para sus N análogos, en el punto de la rejilla sinóptica más próximo al punto de estima en cuestión.

3.2.4. Búsqueda del grupo definitivo de análogos a usar en la estimación

Además de la similaridad sinóptica de los campos de valores tipificados de las componentes de viento geostrófico, para reforzar la similaridad se tiene en cuenta la similitud de los valores tipificados de las variables predictoras determinadas en 3.2.3., calculados en el punto de la rejilla sinóptica (figura 1) más próximo al punto de estima (figura 2); esta similitud se obtiene calculando la distancia d_p del conjunto de valores tipificados de los predictores utilizados, es decir:

$$d_p = \sqrt{\frac{\sum_1^J f_j (p_{ij} - p_j)^2}{\sum_1^J f_j}} \quad [3]$$

en donde:

p_{ij} y p_j son respectivamente los valores del predictor j tipificado para el día problema, y para el elemento cuya similaridad con el día problema se quiere analizar;

f_j es el peso o ponderación que se da al predictor j y que se toma igual al valor de su coeficiente de correlación parcial con la precipitación (ver punto 3.1.3);

J es el número total de predictores utilizados.

De este modo se calculan los valores de d_p para cada uno de los N análogos al día problema quedándonos solamente con aquellos cuya d_p sea inferior a un umbral.

La distancia definitiva con el día problema de cada uno de los $N1$ análogos finalmente seleccionados, se calcula promediando a su vez la distancia sinóptica d_s derivada del campo de viento geostrófico tipificado, con la distancia d_p derivada de los predictores tipificados. Es decir:

$$D = \frac{d_s + d_p}{2} \quad [4]$$

En el caso en que no existan predictores significativos la distancia sólo contemplaría la analogía sinóptica.

3.2.5. Estimación de la precipitación

En primer lugar, se comprueba si la precipitación en el punto de estima es inferior a 0,1 mm para, al menos, la mitad más uno de los análogos; si es así, la precipitación estimada se considera igual a cero concluyendo la estimación.

Si lo anterior no se verifica, el valor estimado de la precipitación se calcula determinando la media ponderada de las precipitaciones de los $N1$ análogos en el punto de estima, siendo las ponderaciones las inversas de las distancias medias D entre cada análogo y el día problema.

4. VALIDACIÓN DE LAS ESTIMACIONES DE PRECIPITACIÓN

Para estudiar posibles cambios climáticos es importante analizar la evolución de la medias anuales y mensuales calculadas en períodos de varios años (30 ó 31 normalmente). Por otra parte, el método seguido proporciona estimaciones de precipitación diaria que sirven de dato base para muchos modelos de impactos que requieren dato diario. Por todo ello, se validará el comportamiento de los valores medios estimados respecto a los valores observados y además se analizarán los errores y los *skill-score* (WILKS, 1995) de las estimaciones diarias y las diferencias entre las distribuciones de los valores diarios observados y estimados.

4.1. Validación de las medias anuales

Se ha calculado la media anual de la precipitación estimada para el período 1981-1990 al objeto de compararla con la media anual de la precipitación real del mismo período. La precipitación estimada en el período 1981-1990 se ha determinado utilizando análogos sinópticos buscados en el período 1961-1980 con el fin de independizar completamente el período de validación del período de datos base; además se ha evitado utilizar análogos que caigan dentro de un entorno temporal de 10 días respecto del día de estima para evitar también posibles influencias favorables en el resultado.

En la figura 3 pueden verse los valores de la precipitación media anual observada y estimada, así como el % de variación de esta última respecto al valor observado correspondiente, para el período 1981-1990. Como puede observarse la distribución geográfica de la precipitación es bastante similar reproduciéndose con bastante fidelidad las zonas de mayor y menor pluviosidad así como los valores de la misma. Las variaciones porcentuales no son grandes en la mayoría de la zona de estudio, oscilando entre sobreestimas que no superan el 2% en la zona noroccidental y en algunos puntos del Pirineo, y subestimas en la zona levantina que llegan a un máximo del 32% en el SE debido, en parte, a que los valores observados son más bajos, influyendo en un aumento del porcentaje de variación, aunque fundamentalmente es debido a la subestima de la precipitación en otoño en la zona mediterránea como veremos seguidamente.

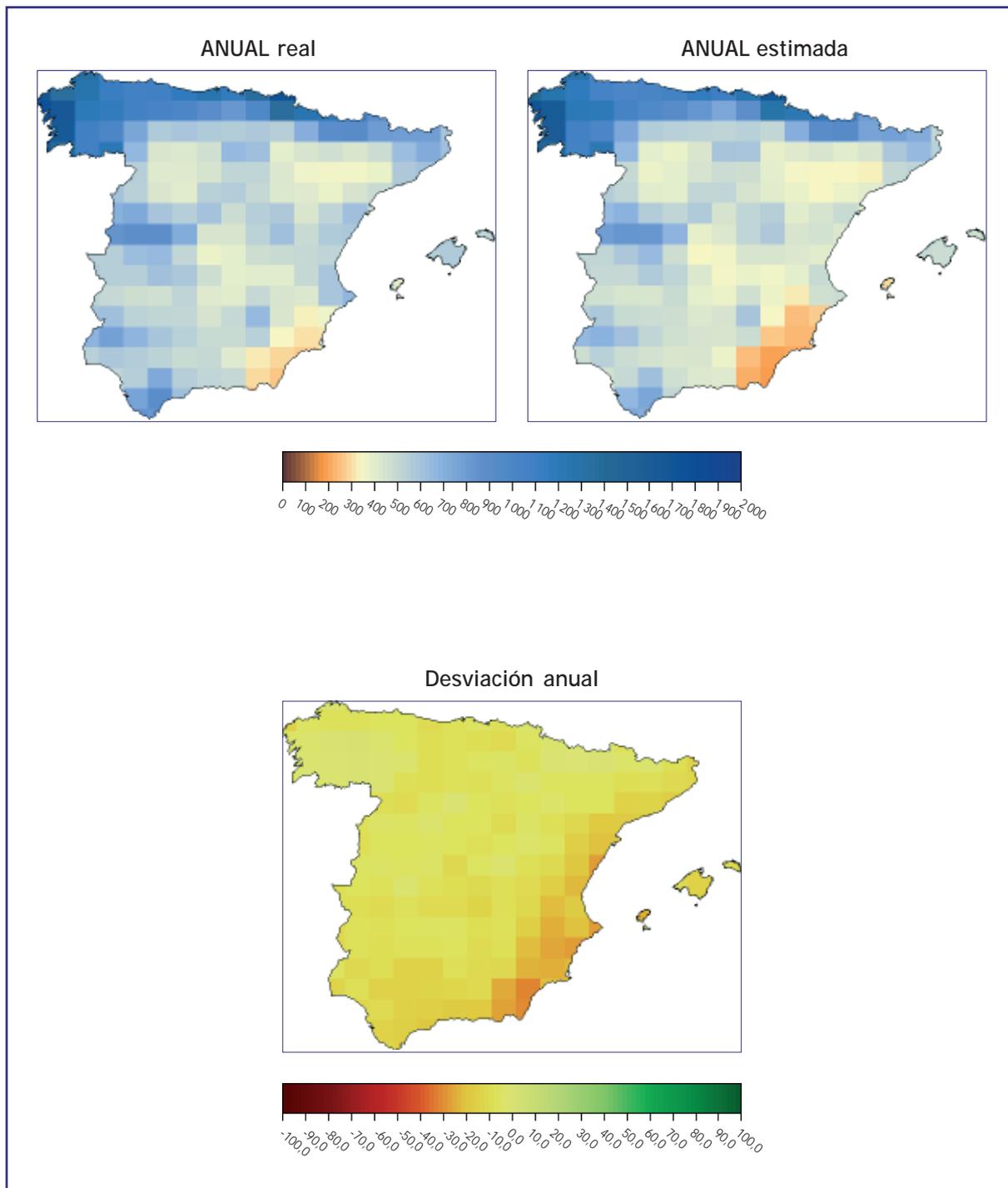


Figura 3. En la parte superior, precipitación media anual observada (real) y estimada para el período 1981-1990. En la parte inferior, desviación en % de la precipitación estimada respecto de la observada.

4.2. Validación de las medias mensuales

Se han calculado las medias mensuales de precipitación en el período 1961-1990 tanto para la precipitación observada como para la precipitación estimada; en este último caso las estimaciones de cada mes se han realizado buscando análogos fuera de los meses cuya media se está calculando para así independizar los datos utilizados en la estimación, de los datos que van a ser validados; asimismo, no se buscan análogos en un entorno de diez días del día problema para evitar también influencias que puedan sesgar el resultado. En las figuras 4-7 se muestran los mapas donde se representan los valores mensuales medios de precipitación observada y estimada para los meses de todas las estaciones del año; igualmente, en la figura 8, se representan las variaciones porcentuales de las medias mensuales de precipitación estimada respecto al valor observado correspondiente.

Analizando por estaciones podemos observar los siguientes aspectos: en los meses de invierno la precipitación estimada reproduce bastante bien la precipitación observada; puede verse en los mapas que la distribución de la precipitación está bien reproducida aunque, en general, algo subestimada, fundamentalmente en los meses de diciembre y enero, salvo en algunas zonas donde se produce una ligera sobreestima; sin embargo, en la gran mayoría de los puntos la desviación de los valores estimados respecto a los observados no supera el 20%. En los meses de primavera, la distribución de la precipitación está, en general, bien reproducida, fundamentalmente en marzo y abril aunque algo peor en mayo; hay un predominio de la sobreestima en los meses de marzo y abril mientras que en mayo vuelve a predominar la subestima; pocos son los puntos en que la desviación supera el 20%, aunque en algún caso se supera el 30% de sobreestima en abril y mayo. En verano, las zonas de mayor y menor pluviosidad están, en general, bien representadas aunque en junio las diferencias son más notables; si nos fijamos en las variaciones relativas podemos observar que estas pueden llegar a ser bastante grandes, superando incluso el 100%, debido al hecho de que las precipitaciones medias de los meses estivales son, en general, bajas con lo cual las diferencias relativas quedan excesivamente resaltadas respecto a las diferencias absolutas. Otoño es la estación en que la precipitación media resulta peor estimada, la subestima es casi absoluta en toda la estación siendo más marcada en los meses de septiembre y octubre, especialmente en la zona mediterránea donde quizás no quedarían bien reflejados en el método todos los factores influyentes; existe, sobre todo en septiembre, una subestima superior al 20% en la mayoría de los puntos llegando en algún caso puntual a valores en torno al 60%.

En resumen, podemos decir que la precipitación media mensual está razonablemente bien estimada en invierno y primavera; en verano, en términos absolutos, quedan aceptablemente reflejados la distribución geográfica y los valores bajos propios de la estación, si bien las variaciones relativas resultan altas debido precisamente a esos bajos valores de las medias mensuales; finalmente, en otoño la subestima de la precipitación media mensual es prácticamente general, las estimaciones son peores y muy especialmente en las zonas mediterráneas.

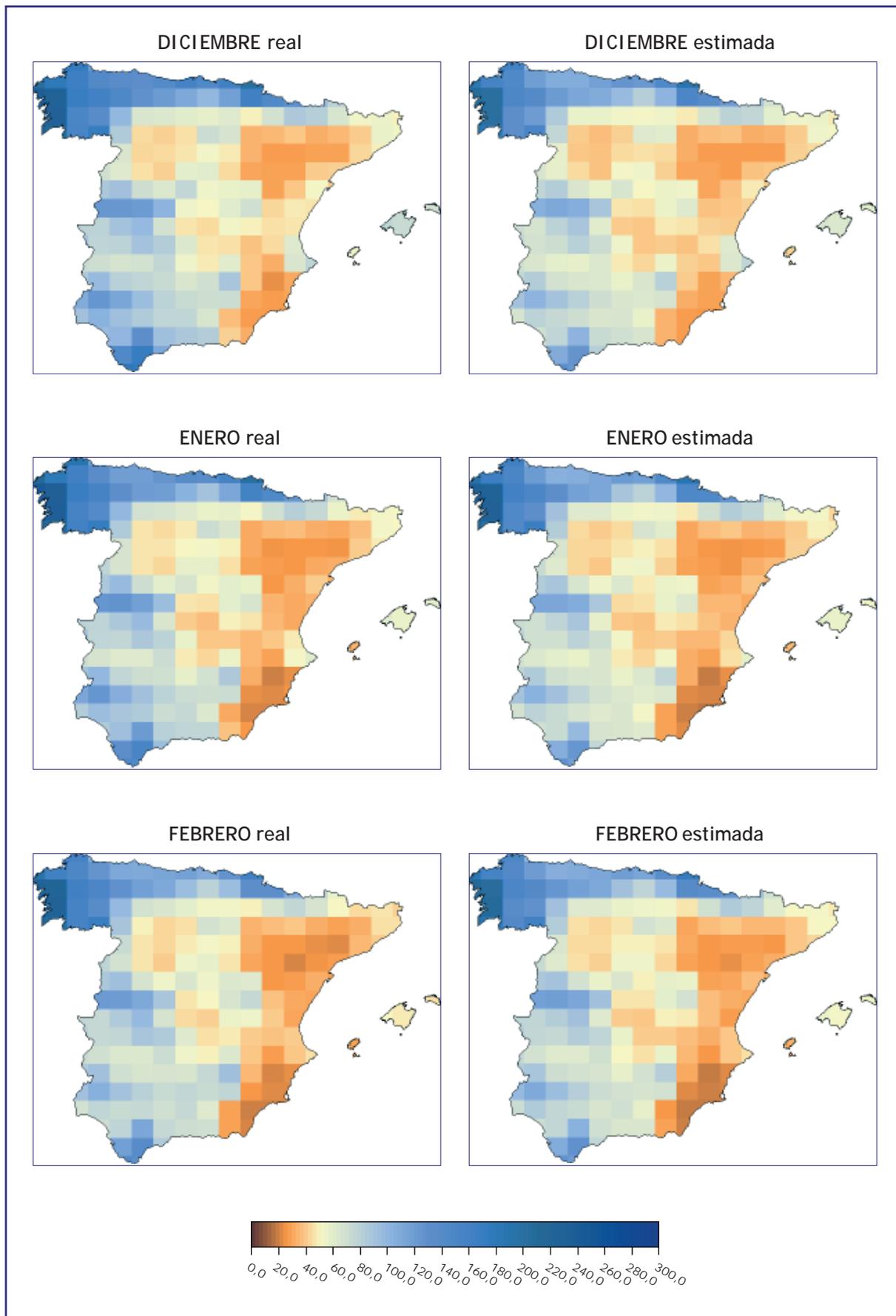


Figura 4. Precipitación media observada (izda.) y estimada (dcha.) para los meses de invierno del periodo 1961-1990.

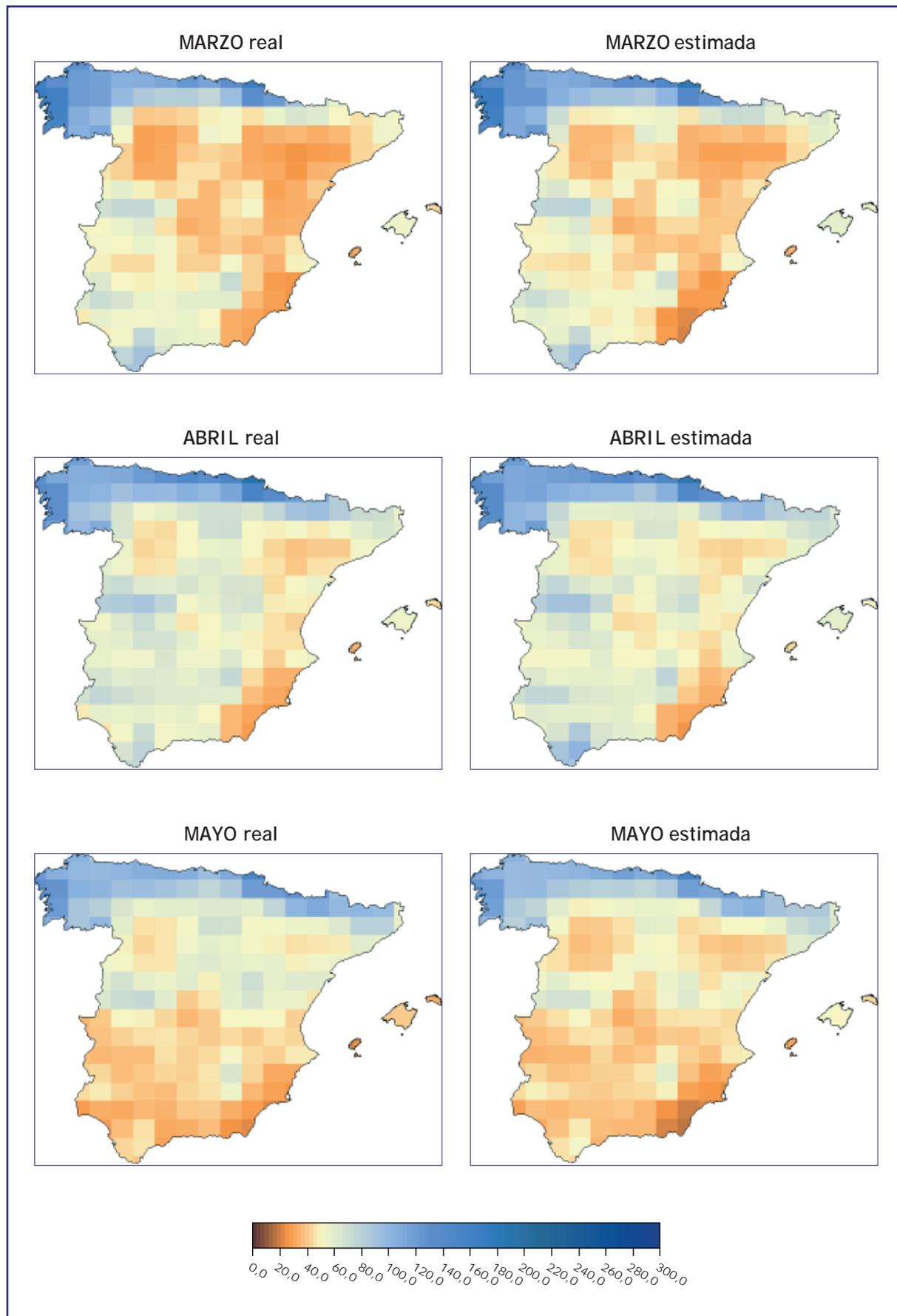


Figura 5. Precipitación media observada (izda.) y estimada (dcha.) para los meses de primavera del período 1961-1990.

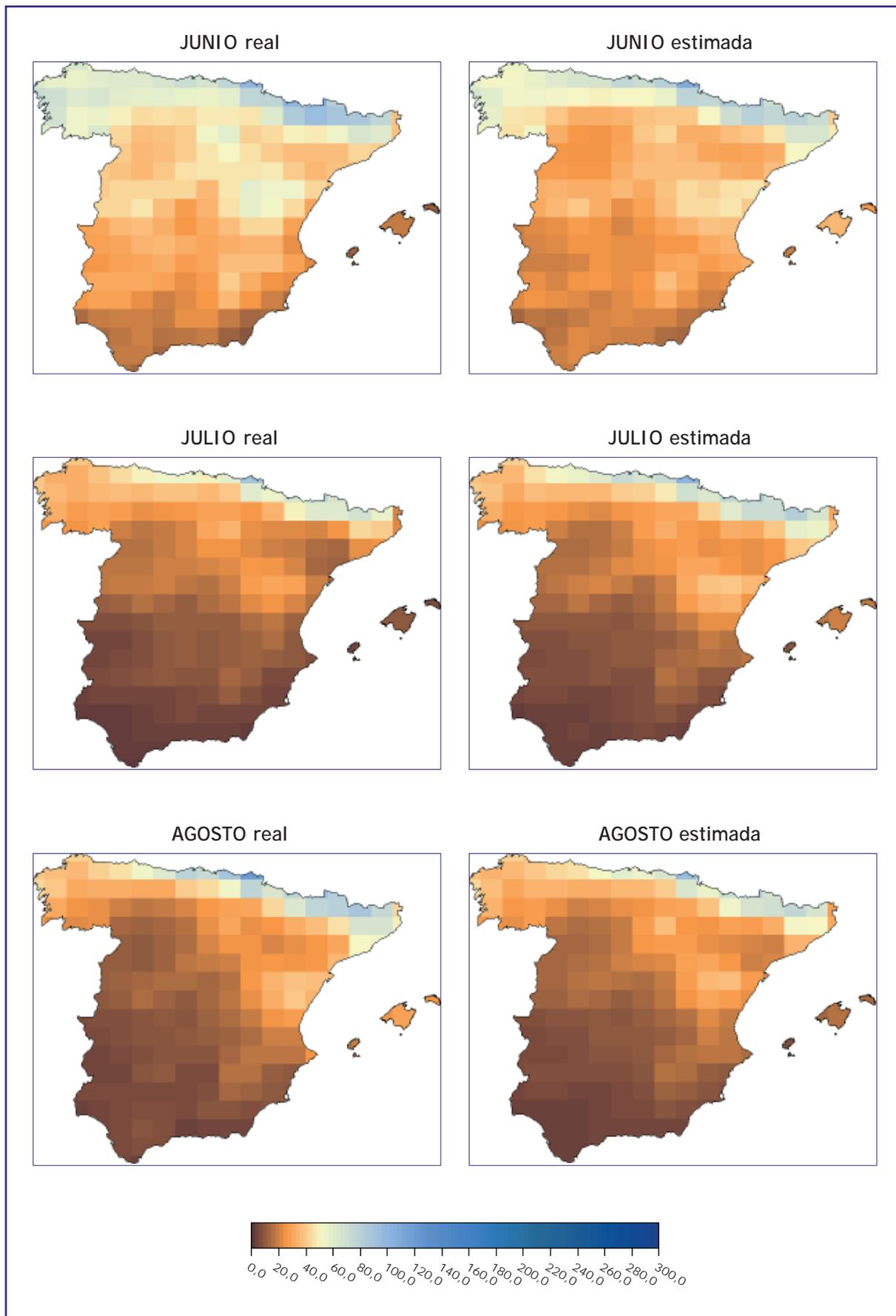


Figura 6. Precipitación media observada (izda.) y estimada (dcha.) para los meses de verano del periodo 1961-1990.

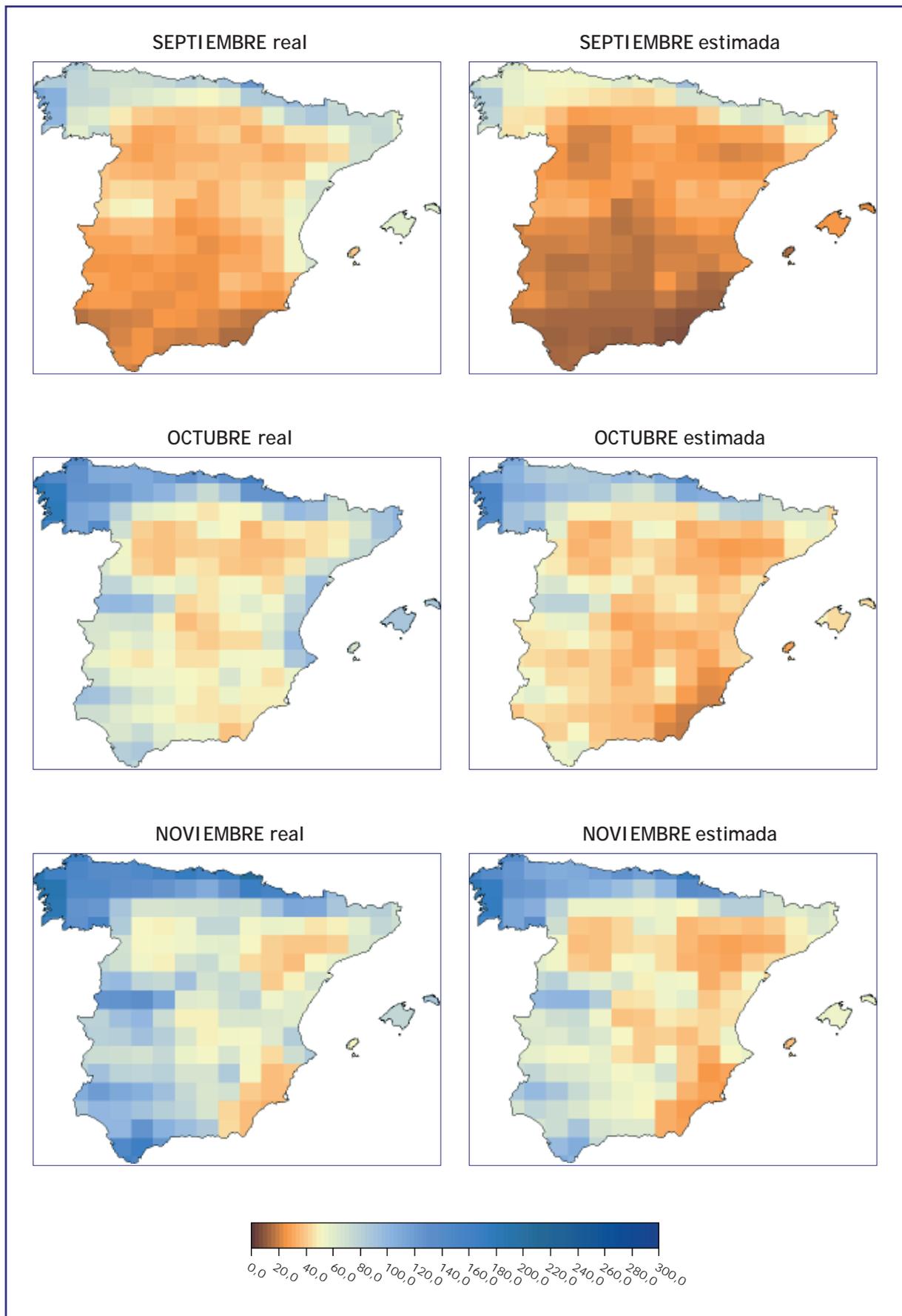


Figura 7. Precipitación media observada (izda.) y estimada (dcha.) para los meses de otoño del periodo 1961-1990.

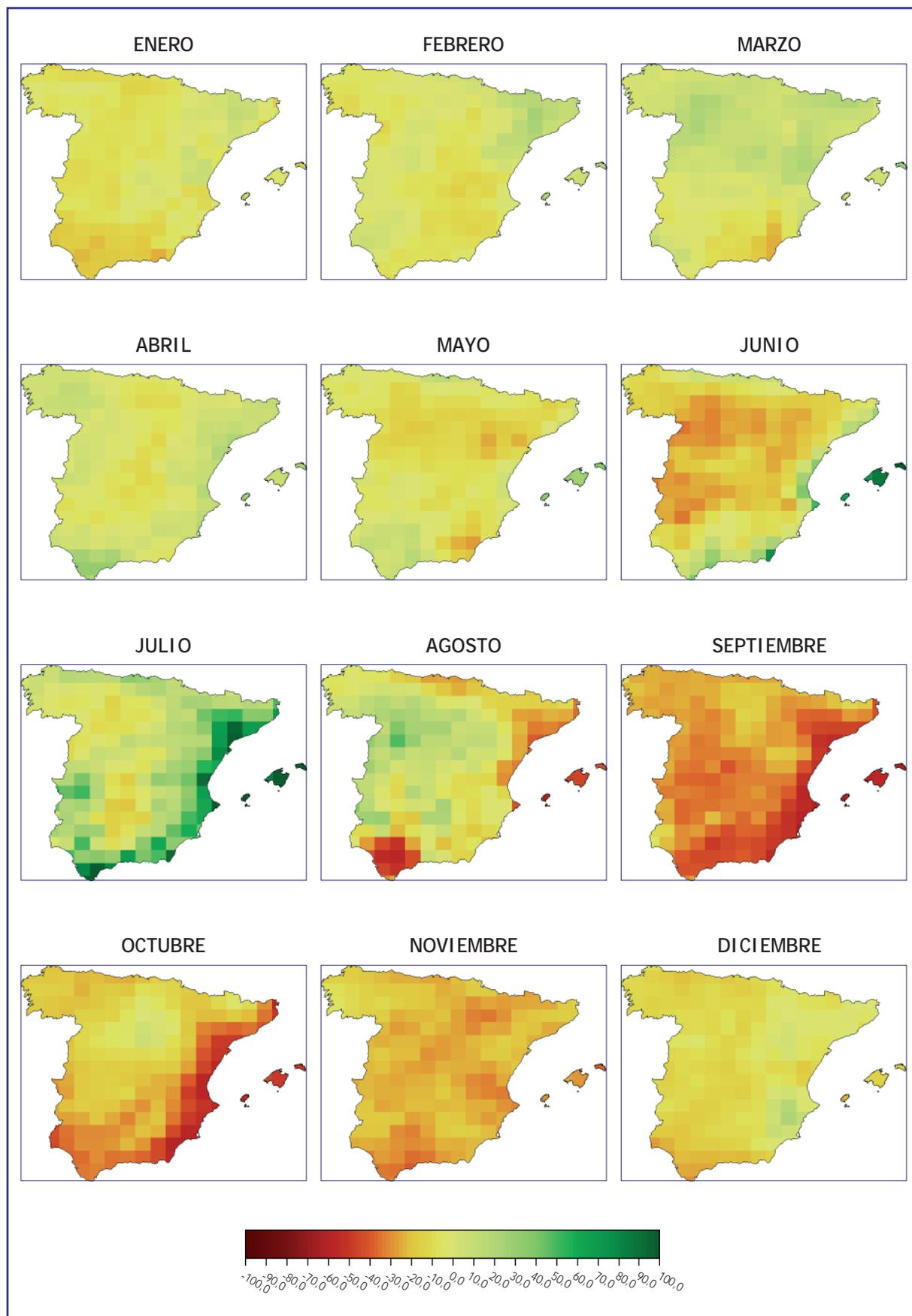


Figura 8. Desviación porcentual de la precipitación media mensual estimada respecto a la observada para el período 1961-1990.

4.3. Validación de las estimaciones de precipitación diaria

El método de *downscaling* utilizado proporciona estimaciones con resolución temporal diaria; por otra parte, en muchos casos, los modelos de impactos requieren datos con resolución diaria, por lo que es conveniente validar las estimaciones diarias realizadas por el método. Esta validación se realizará, por un lado, comparando percentiles significativos de la distribución de la precipitación diaria estimada, con los correspondientes de la distribución de la precipitación observada, y por otro lado analizando los errores y los *skill-score* de los valores diarios estimados.

4.3.1. Comparación de percentiles

Se comparan las distribuciones de los valores de la precipitación diaria observada y estimada analizando los valores de los percentiles más representativos. Las estimaciones en este caso se hacen en el período 1981-1990 utilizando análogos del período 1961-1980 para así separar el período de estima del período de datos utilizados para la misma; igualmente, no se utilizan análogos en un entorno de 10 días del día de estima para evitar sesgos favorables en los resultados. En las figuras 9 a 14 se presentan los percentiles 10, 25, 50, 75, 90 y 99 para la precipitación observada y estimada así como la diferencia entre el valor estimado y el valor observado, nótese que se ha empleado distinta escala en cada caso debido a la gran diferencia en el rango de valores para los distintos percentiles. El percentil 10 es cero para la precipitación observada y para casi todos los puntos en la estimada, sólo en algunos puntos del norte el percentil 10 llega a los 0,2mm máximo. Las diferencias en el percentil 25 son pequeñas, en la gran mayoría de los puntos no superan los +0,1mm siendo las mayores de +0,8mm en algún punto del norte; el percentil 50 sigue manifestando valores algo superiores en las precipitaciones estimadas aunque no llegan a +1mm, salvo en unos pocos puntos del norte, alcanzando como mucho valores en torno a +1,5mm; el percentil 75 también presenta diferencias pequeñas que son inferiores a +1mm en la mayoría de los puntos llegando como mucho a +1,3mm; el percentil 90, considerado generalmente como límite para valores extremos, es bastante concordante en la precipitación observada y estimada, como puede verse en los mapas, aunque se presenta generalmente con valores inferiores en la estimada respecto a los correspondientes valores en la observada, y que son mayores en puntos del norte alcanzando diferencias en torno a -2 y -3mm; por último el percentil 99, representante de los valores máximos es claramente inferior en la precipitación estimada con algunos casos en que la diferencia está en torno a los -19mm en puntos de la zona mediterránea y del norte.

En resumen, podemos decir que las precipitaciones más bajas, representadas por los percentiles 10, 25 y 50, aunque bien representadas en general, están ligeramente sobreestimadas con respecto a las correspondientes precipitaciones observadas presentando mayor número de días con precipitación en el caso de las estimadas. Esto puede ser consecuencia del método de estimación que se basa en promedios de las precipitaciones de los análogos, algunos de los cuales pueden presentar valores que, aunque sean bajos, pueden hacer subir el valor de la media. Para mayores valores de precipitación, percentiles 75 y 90, las diferencias son pequeñas en términos relativos siendo la representatividad de los valores estimados buena en general, aunque la subestima aumenta a medida que los valores van siendo más altos; llegamos así a que los valores más extremos quedan bastante subestimados como claramente manifiestan las diferencias en los valores del percentil 99, consecuencia derivada, al menos en parte, del promedio de los valores de los análogos, aunque quizás se deba fundamentalmente a que no estén del todo bien representados los factores que influyen en las precipitaciones más copiosas.

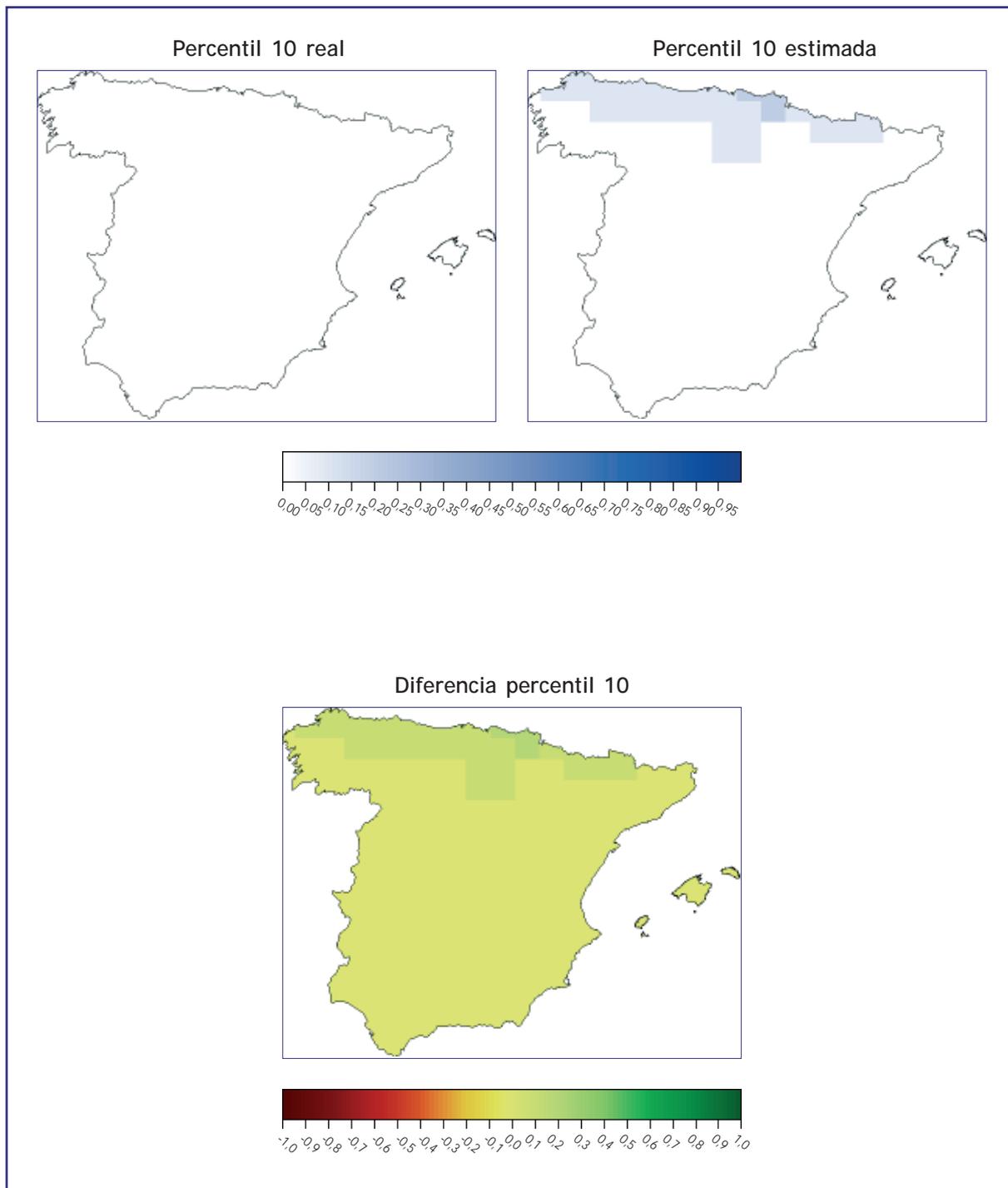


Figura 9. En la parte superior, percentil 10 de la precipitación diaria observada y estimada del periodo 1981-1990. En la parte inferior, diferencia entre el percentil de la precipitación estimada y el de la observada.

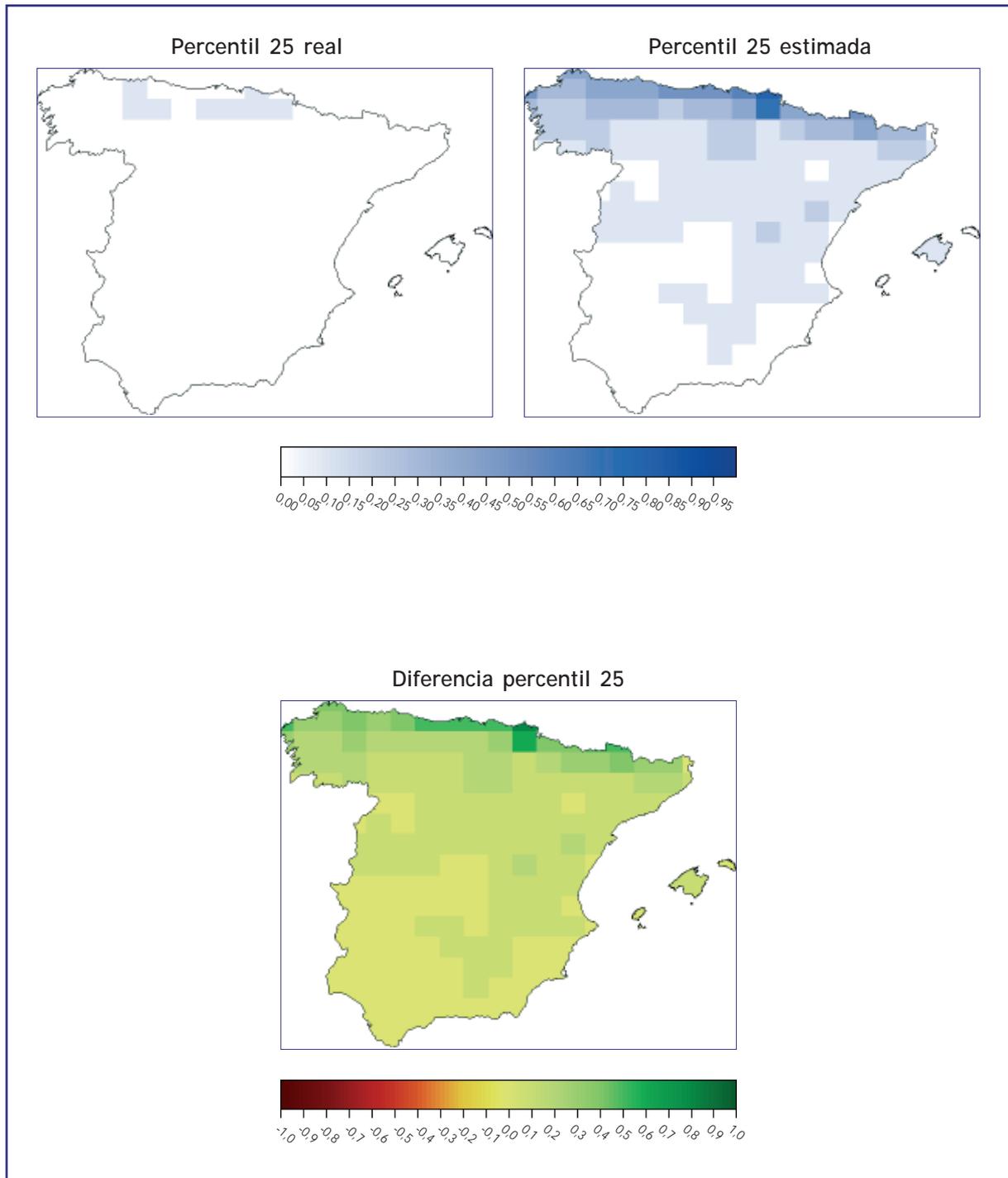


Figura 10. En la parte superior, percentil 25 de la precipitación diaria observada y estimada del periodo 1981-1990. En la parte inferior, diferencia entre el percentil de la precipitación estimada y el de la observada.

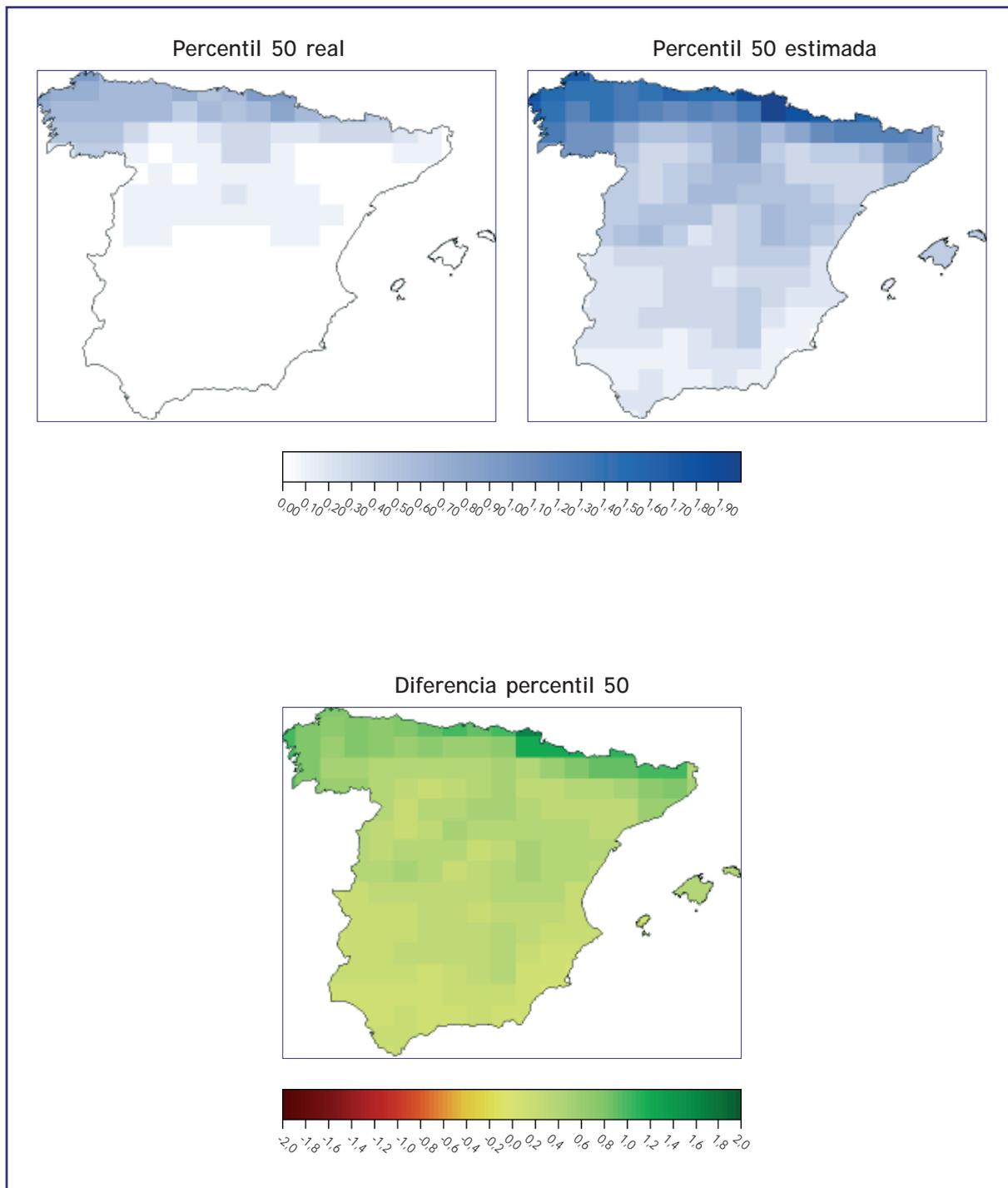


Figura 11. En la parte superior, percentil 50 de la precipitación diaria observada y estimada del periodo 1981-1990. En la parte inferior, diferencia entre el percentil de la precipitación estimada y el de la observada.

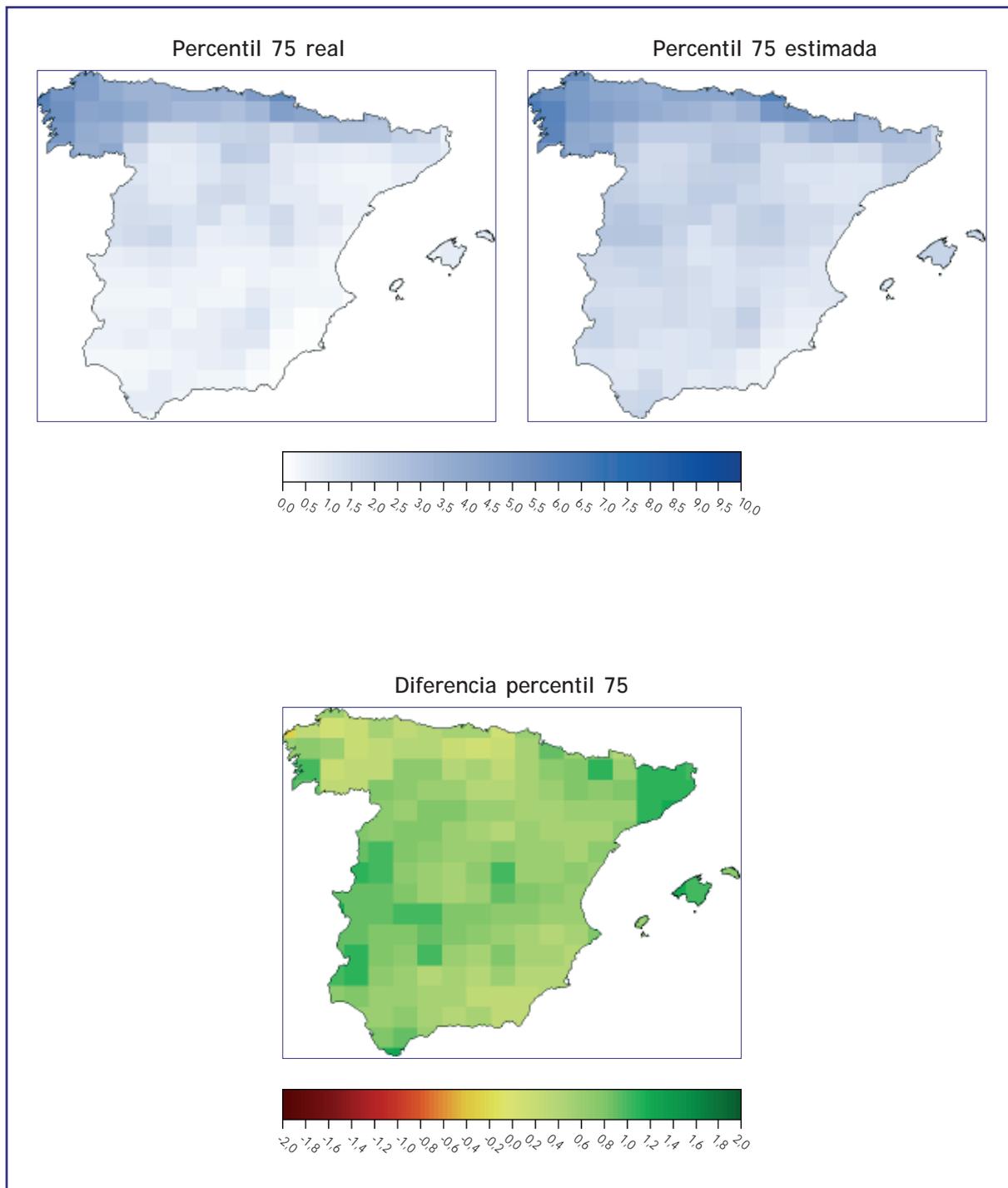


Figura 12. En la parte superior, percentil 75 de la precipitación diaria observada y estimada del periodo 1981-1990. En la parte inferior, diferencia entre el percentil de la precipitación estimada y el de la observada.

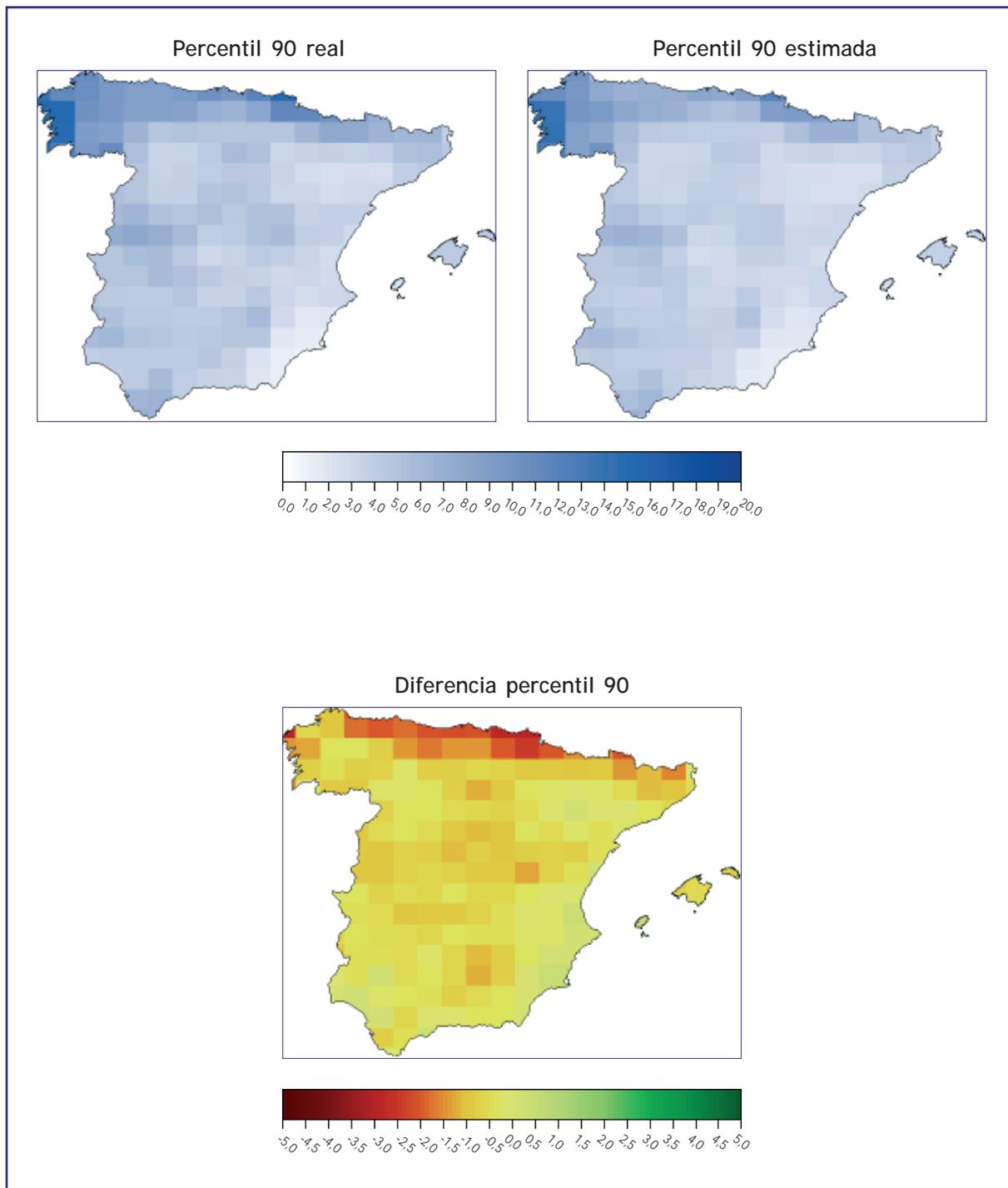


Figura 13. En la parte superior, percentil 90 de la precipitación diaria observada y estimada del periodo 1981-1990. En la parte inferior, diferencia entre el percentil de la precipitación estimada y el de la observada.

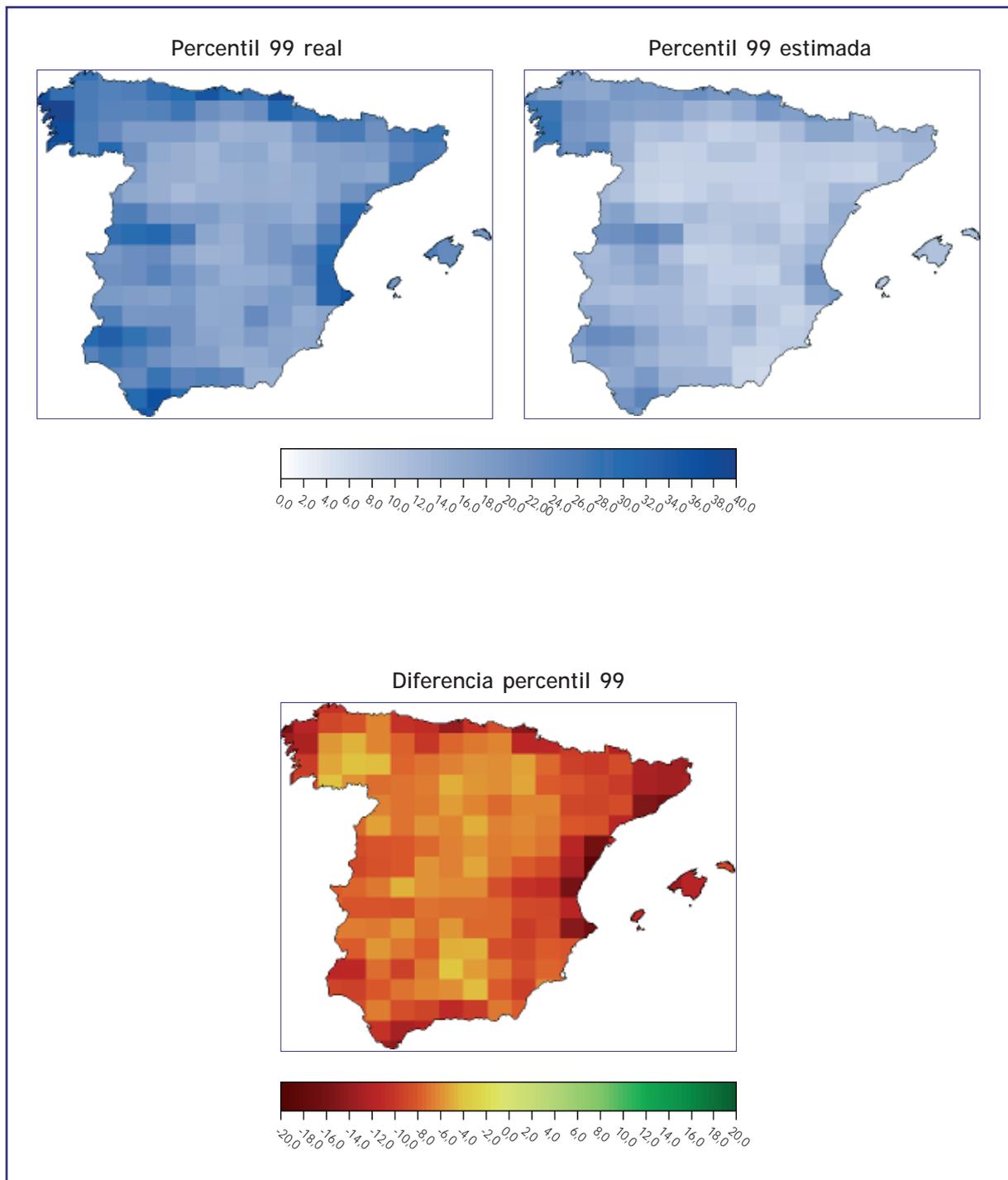


Figura 14. En la parte superior, percentil 99 de la precipitación diaria observada y estimada del periodo 1981-1990. En la parte inferior, diferencia entre el percentil de la precipitación estimada y el de la observada.

4.3.2. Índices de verificación. Errores

Se consideran las estimaciones diarias de precipitación para el período 1961-1990. Estas estimaciones se realizan buscando análogos fuera del entorno de diez días del día a estimar por la razón anteriormente mencionada.

Se calcula para cada mes, el error absoluto medio EAM ó media de los valores absolutos de las diferencias entre los valores estimados diarios y los correspondientes observados. Su expresión sería

$$EAM = \frac{\sum_{i=1}^n |x_{ie} - x_{io}|}{n} \quad [5]$$

donde n es el número de datos y x_{ie} y x_{io} son los valores estimados y observados respectivamente.

Dado su amplio uso, también calcularemos la raíz cuadrada del error cuadrático medio RECM, el cual, a su vez, es la media de los cuadrados de las desviaciones de los valores estimados respecto de los observados y cuya expresión sería

$$RECM = \sqrt{\frac{\sum_{i=1}^n |x_{ie} - x_{io}|^2}{n}} \quad [6]$$

Si bien el error EAM es más propio como medida de error, el error RECM es más sensible a los casos en que las diferencias entre valores estimados y observados son grandes, debido a que las diferencias están elevadas al cuadrado. A continuación se presentan los mapas de la distribución mensual de ambos tipos de errores, que pueden verse en las figuras 15 y 16.

Podemos observar que los errores absolutos medios tienden a ser mayores en las zonas de mayor pluviosidad, lo cual puede deberse simplemente al hecho de ser los valores diarios mayores, dando mayor margen para el error, lo que no necesariamente implicaría mayor error en términos relativos. Destacan los bajos errores en los meses estivales, debidos a que hay un predominio de días sin precipitación, los cuales están generalmente bien estimados en esta época dando lugar a un error promedio bajo, aunque pudiera haber algunos pocos días con errores más importantes. La distribución del error RECM es parecida a la de EAM si bien destacan más los errores en otoño en las zonas mediterráneas donde los valores de precipitación diaria pueden llegar a ser bastante subestimados, efecto mejor captado por el RECM.

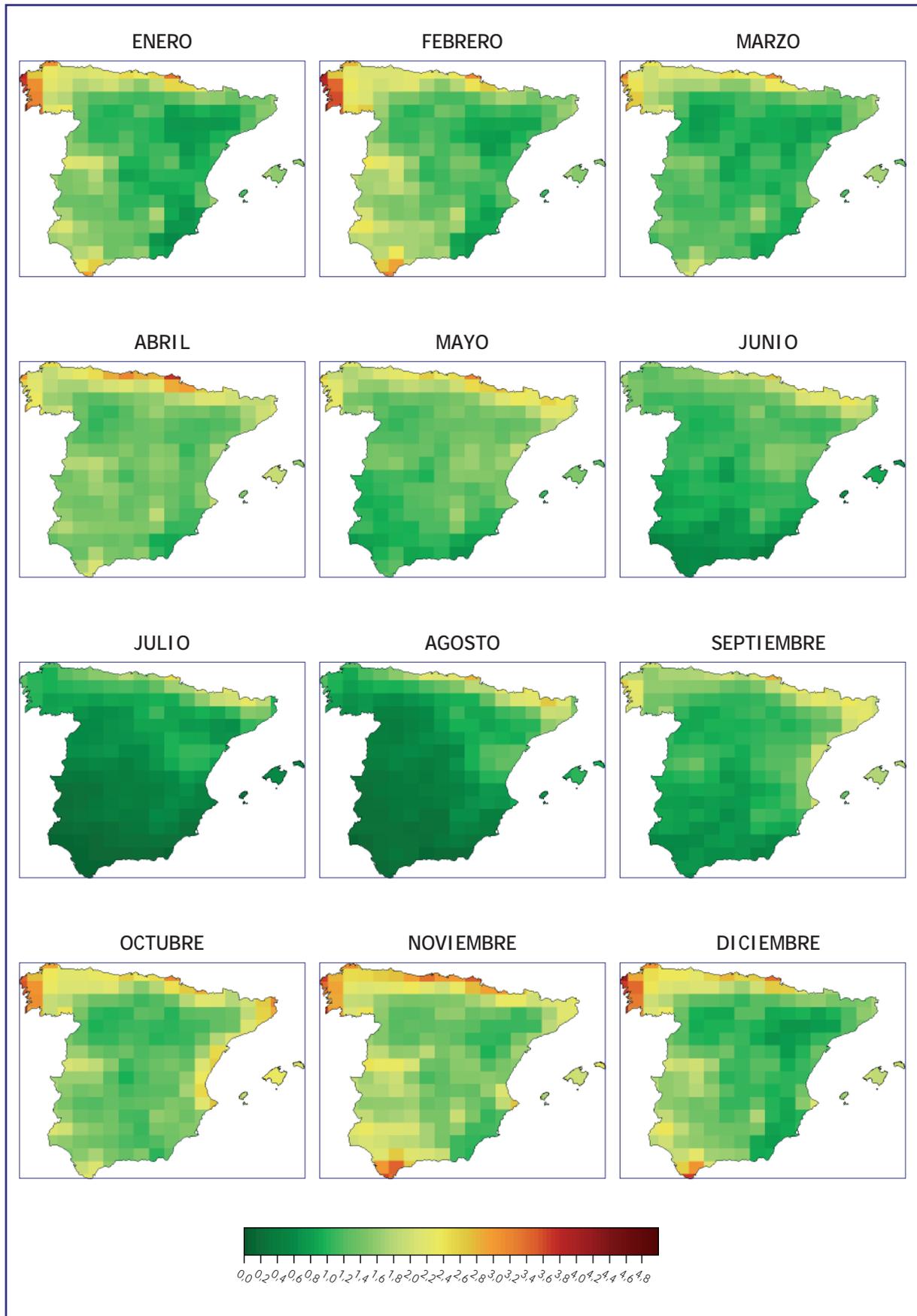


Figura 15. Error absoluto medio mensual de la precipitación diaria estimada respecto a la observada para el período 1961-1990.

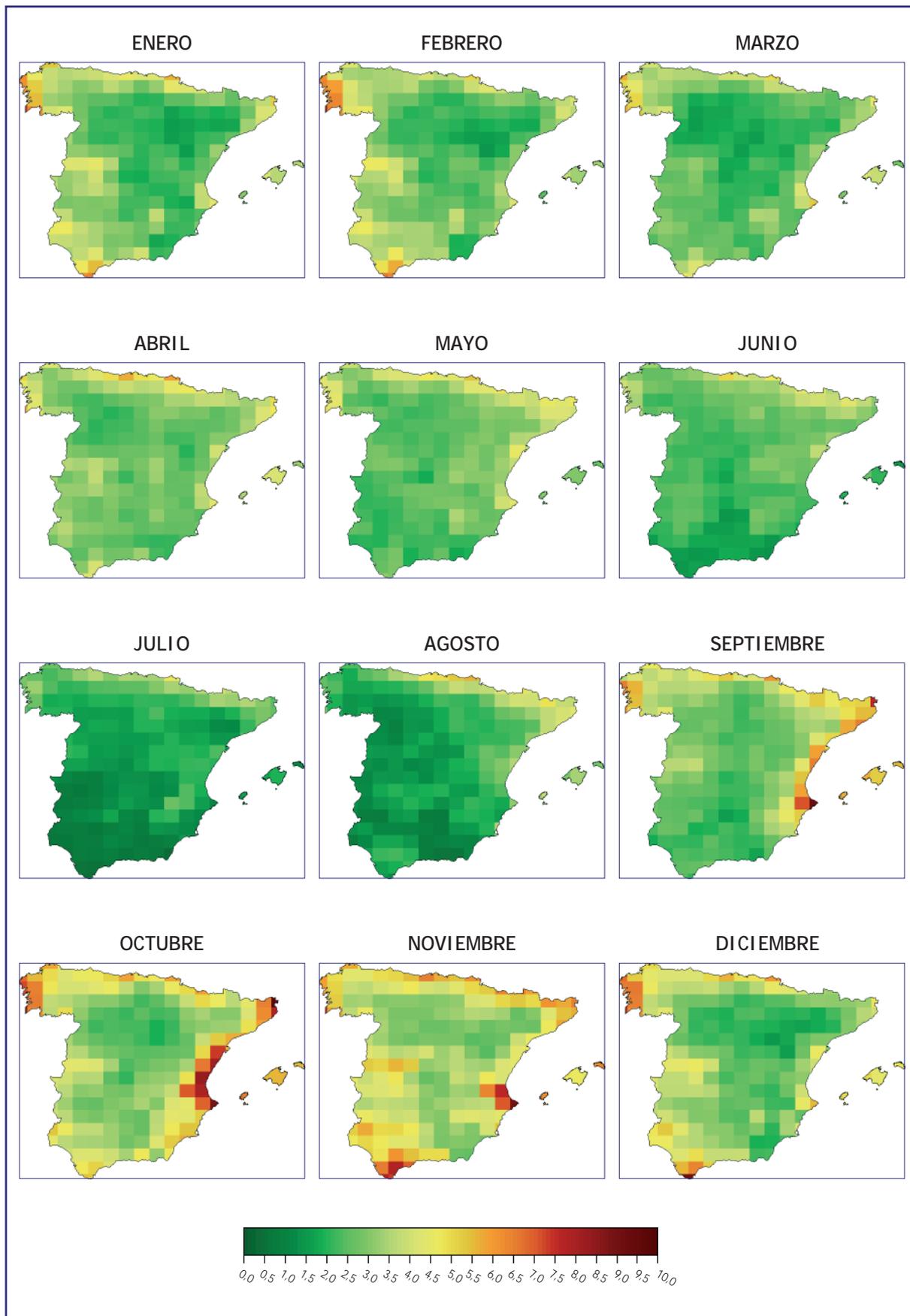


Figura 16. Raíz cuadrada del error cuadrático medio mensual de la precipitación diaria estimada respecto a la observada para el período 1961-1990.

4.3.3. Índices de verificación. *Skill-scores*

Se trata de determinar si las estimaciones obtenidas con el método son mejores que las obtenidas mediante otro procedimiento que tomamos como referencia. En este caso se consideran como referencia las estimaciones obtenidas tomando los valores climatológicos mensuales del punto de estima, correspondientes al mes del día para el que se trata de estimar. Se valoran las estimaciones diarias en el período 1961-1990, que se han realizado tomando análogos fuera del entorno de diez días del día de estima.

Las expresiones [7] y [8] representarían respectivamente, el error absoluto medio y la raíz del error cuadrático medio para el caso en que las estimaciones se realizasen tomando los valores climatológicos mensuales correspondientes. Estas expresiones representarían además, medidas de dispersión de los datos observados de cada mes siendo [7] la desviación media y (8) la desviación tipo.

$$EAM_c = \frac{\sum_{i=1}^n |x_c - x_{io}|}{n} \quad [7]$$

$$RECM_c = \sqrt{\frac{\sum_{i=1}^n |x_c - x_{io}|^2}{n}} \quad [8]$$

Definimos entonces los índices *skill-score* de la siguiente manera:

$$SK_{EAM} = 1 - \frac{EAM}{EAM_c} \quad [9]$$

$$SK_{RECM} = 1 - \frac{RECM}{RECM_c} \quad [10]$$

La expresión [9] se calcula basándose en el error absoluto medio, mientras que la expresión [10] se calcula basándose en la raíz del error cuadrático medio. Obsérvese que los valores de los *skill-score* serán cero si los errores del método igualasen a los de la referencia, lo que no supondría mejora respecto a esta, mientras que serán positivos si el método mejora a la referencia, tanto más cuanto mayor sea el valor del *skill-score*, de manera que en el límite de la estimación perfecta (error cero) el *skill-score* sería 1; por el contrario, valores negativos del *skill-score* supondrían estimaciones peores que la referencia. Otra interpretación de las expresiones [9] y [10] es que las estimaciones con el método pueden considerarse mejores cuanto más pequeños sean sus errores medios en relación a la dispersión propia de los datos observados.

En la figura 17 se representan los *skill-score* mensuales calculados basándose en el EAM. En general puede observarse una mejora considerable de las estimaciones hechas con el método respecto a las realizadas con la climatología; en los meses de verano, las mejoras son bastante más discretas, singularmente en el mes de julio, habiendo algunos puntos en la zona mediterránea y sur en que las estimaciones incluso empeoran respecto a la climatología.

En la figura 18 se representan los *skill-score* mensuales calculados basándose en el RECM. También se observa una mejora general de las estimaciones hechas con el método aunque menos marcada que el caso anterior, debido probablemente a la mayor influencia en el RECM de los errores grandes; las peores estimaciones se presentan entre mayo y septiembre, donde aunque en la mayoría de los casos se mejora o al menos se iguala a la climatología hay unos pocos puntos, predominantemente en julio, donde las estimaciones son peores.

En resumen, los resultados parecen indicar que las precipitaciones de origen atlántico son, en general, mejor estimadas que las de origen mediterráneo aunque para estas últimas también se manifiesta una clara mejora respecto a la climatología; las precipitaciones estivales, debido a su carácter, están peor representadas habiendo casos en que los resultados incluso empeoran respecto a la climatología.

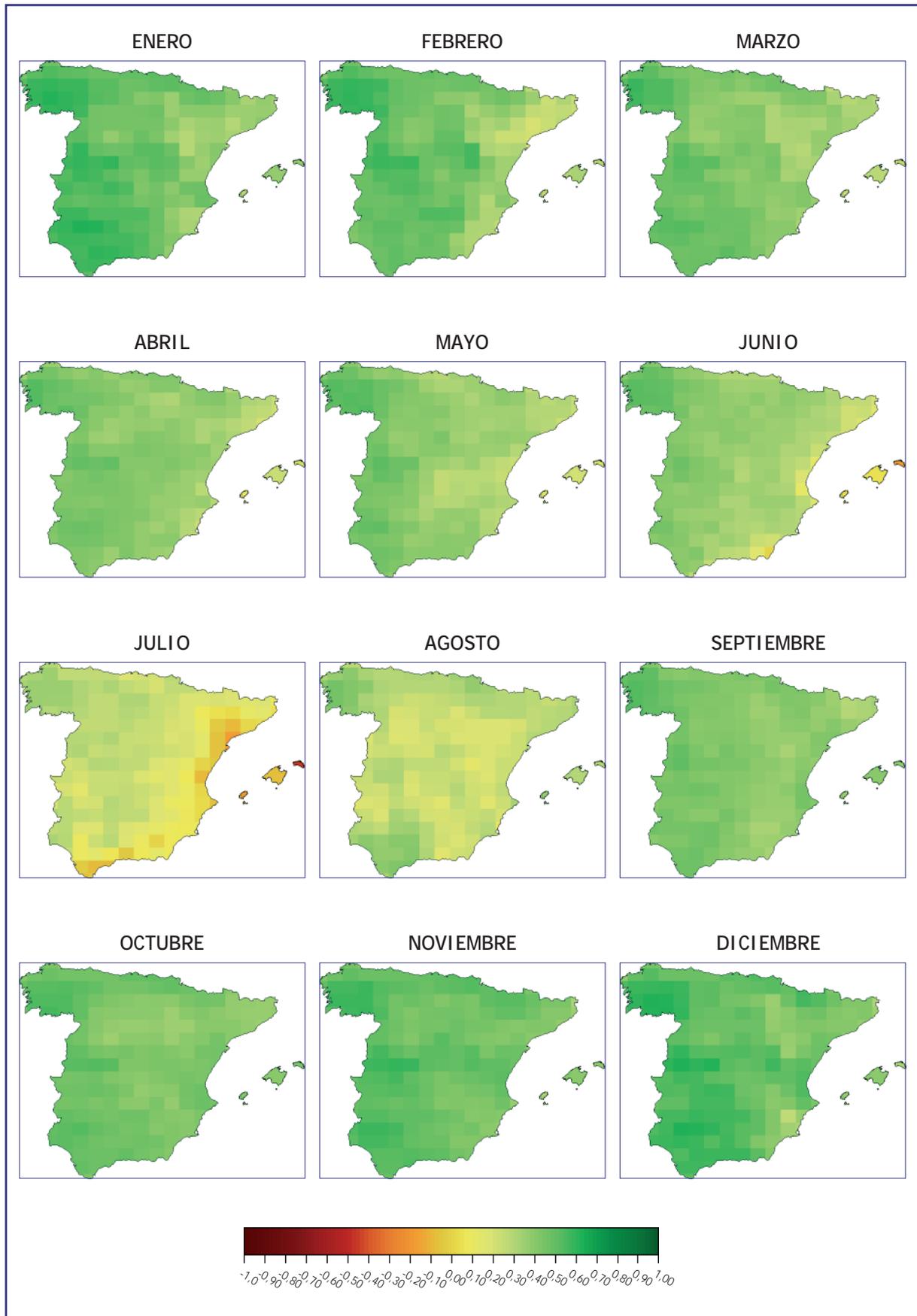


Figura 17. Skill-score mensual de la precipitación diaria estimada basado en el EAM calculado respecto a la climatología mensual para el período 1961-1990.

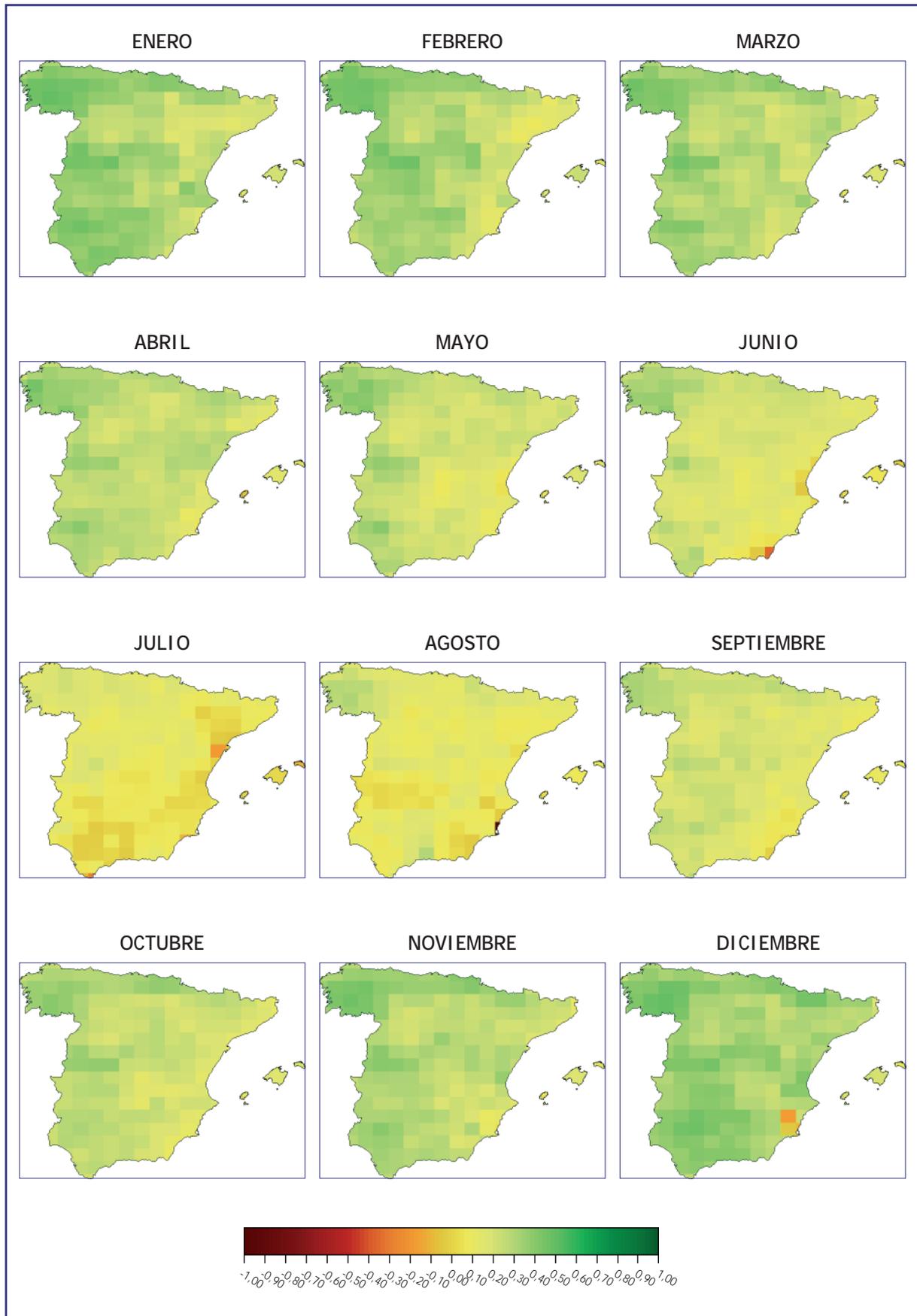


Figura 18. Skill-score mensual de la precipitación diaria estimada basado en el RECM calculado respecto a la climatología mensual para el período 1961-1990.

5. CONCLUSIONES

Se han presentado métodos de estimación para la regionalización de precipitación basándose en análogos sinópticos.

Se ha validado el método analizando el comportamiento de los valores estimados en relación con los correspondientes observados tanto para los valores medios anuales y mensuales como para los diarios donde se analiza su distribución, los errores medios y los *skill-score* indicativos de la mejora de las estimaciones basadas en el método, respecto a las estimaciones basadas en la climatología.

Las medias anuales están bien representadas tanto en su distribución geográfica como en sus valores, aunque hay una subestima en la zona mediterránea derivada básicamente de la subestima de la precipitación en otoño.

Las precipitaciones mensuales están razonablemente bien estimadas en invierno y primavera; en verano se ponen de manifiesto aceptablemente la distribución geográfica y los bajos valores propios de la estación aunque las desviaciones relativas sean, en algún caso, marcadas precisamente debido a esos valores bajos; finalmente, en otoño las medias mensuales son generalmente subestimadas, especialmente en zonas mediterráneas.

La distribución de los valores diarios de precipitación está –en general– bien representada, si bien los valores más bajos (percentiles 10, 25), aunque bien reflejados en general, están ligeramente sobre-estimados. Los valores más extremos (percentil 99) están marcadamente subestimados, debido quizás tanto a la suavización derivada del propio método como a la falta de representación de algunos de los factores que influyen en las precipitaciones más extremas.

En cuanto a los errores medios de la precipitación diaria podemos observar la tendencia a ser algo mayores en las zonas de mayor pluviosidad, lo que no implica necesariamente un mayor error en términos relativos, y la tendencia a valores bajos en los meses estivales debido al predominio de días sin precipitación, los cuales están bien reflejados aunque las precipitaciones de origen convectivo no estén del todo bien representadas.

Finalmente los *skill-score* muestran una mejora considerable de las estimaciones hechas con el método respecto a las realizadas con la climatología mensual, aunque en los meses de verano las mejoras son algo más discretas con algunos pocos puntos en que se iguala o incluso se empeora algo respecto a la climatología.

Las estimaciones se han realizado basándose exclusivamente en información de tipo sinóptico. En pruebas realizadas introduciendo parámetros de humedad, se mejoran los resultados disminuyendo algo los errores. Aunque la humedad no es un parámetro especialmente bien simulado por los modelos globales, puede ser aconsejable su introducción en la regionalización de la precipitación debido a que podría cobrar una mayor importancia en un contexto de cambio climático.

BIBLIOGRAFÍA

- ANDERBERG, M., 1973: *Cluster Analysis for Applications*, Academic Press.
- BORÉN, R., J. RIBALAYGUA, L. BALAIRÓN, 1997: *Método Analógico de Simulación de Escenarios Climáticos a Escala Comarcal*. Informe Técnico n.º4. Servicio de Variabilidad y Predicción del Clima, INM, Madrid.
- IPCC, 1999: *Guidelines on the Use of Scenario Data for Climate Impact and Adaptation Assessment*.
- KALNAY, E. y otros, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, 77, 437-471.
- LUND, I., 1962: Map-Pattern Clasification by Statistical Methods. *J. Appl. Meteorol.*, 2, 56-65.
- PETISCO, S. E., J. M. MARTÍN, D. GIL, 2005: *Método de estima de precipitación mediante downscaling*. Nota Técnica n.º11. Servicio de Variabilidad y Predicción del Clima, INM, Madrid.
- WILBY, R. L., T. WIGLEY, 2000: Precipitation Predictors for Downscaling Observed and General Circulation Model Relationships. *Int. J. Climatol.*, 20, 641-661.
- WILKS, D., 1995: *Statistical Methods in the Atmospheric Sciences*, Academic Press.
- ZORITA, E., H. VON STORCH, 1999: The Analog Method as a Simple Statistical Downscaling Technique: Comparison with More Complicated Methods. *J. Climate*, 12, 2474-2489.