

# HOMOGENEIZACIÓN DE LAS SERIES TERMOPLUVIOMÉTRICAS DE BALEARES Y SU INFLUENCIA EN LA EVALUACIÓN DE LAS TENDENCIAS

José A. GUIJARRO

*Agencia Estatal de Meteorología, Delegación Territorial en Illes Balears*  
jaguijarro@inm.es

## RESUMEN

Se han obtenido 231 series anuales de precipitación total y 66 de temperatura media de Baleares que contaban con un mínimo de 15 y 10 años de observación (respectivamente) en el periodo 1961-2007. Estas series se han homogeneizado mediante una función automática del paquete "climatol", implementado bajo el sistema estadístico R, que trabaja fragmentando las series en los puntos más significativos de las pruebas de salto en la media. Las series de precipitación han sufrido 103 cortes en 70 de las estaciones, mientras que 12 series termométricas se han fragmentado 13 veces. Se han calculado las tendencias antes y después de la homogeneización, y se ha constatado una gran disminución de la variabilidad de las tendencias originales, que han alcanzado una mayor coherencia en su distribución espacial.

**Palabras clave:** Series Climatológicas, Homogeneización, Precipitación, Temperatura, Islas Baleares.

## ABSTRACT

Annual series of 231 total precipitation and 66 average temperature values have been compiled from the Balearic Islands, with a minimum of 15 and 10 years of observation respectively in the period 1961-2007. These series have been homogenised through an automatic function of the package "climatol", implemented under the R statistical system, that works splitting the series in the most significant points detected by several average shift tests. The precipitation series have been splitted 103 times in 70 of the stations, while 12 thermometric series have been splitted 13 times. Trends have been computed before and after homogenisation, and a great decrease in the original trend variability and a greater spatial distribution consistency has been observed.

**Keywords:** Climatological Series, Homogenisation, Precipitation, Temperature, Balearic Islands.

## 1. INTRODUCCIÓN

Las series de observaciones climatológicas son de capital importancia para la monitorización del clima, dentro del actual contexto de preocupación por el cambio climático y sus impactos sobre las actividades humanas. Desafortunadamente, sus registros están frecuentemente contaminados por diversos accidentes ajenos a la evolución del clima: errores puntuales, cambios de ubicación, cambios en la instrumentación, o cambios en el entorno de los

observatorios (bruscos, como la puesta en irrigación de cultivos de secano, o graduales, como el crecimiento urbano).

Para abordar este problema se han venido desarrollando desde hace tiempo diversos métodos de homogeneización de series climáticas, que intentan discriminar en las mismas la señal propiamente climática, que es la que nos interesa, del ruido producido por los factores mencionados. La variedad de metodologías hace imposible mencionarlas todas, pero pueden encontrarse síntesis de las mismas en los trabajos de PETERSON et al. (1998) y AGUILAR et al. (2003). La mayor parte se basan en comparar las series problema con otras supuestamente homogéneas, que o bien pertenecen a observatorios cuya historia puede hacer pensar que se hallan libres de los factores perturbadores (cosa harto difícil), o bien se construyen sintéticamente a partir de otras series vecinas, con la esperanza de que sus respectivas inhomogeneidades queden diluidas en la serie compuesta.

La selección de las series que han de servir de referencia para la homogeneización de la serie problema se suele hacer eligiendo, de entre las disponibles, las que presentan un mayor coeficiente de correlación. Pero de esta forma se prescinde muchas veces de la información que podrían aportar otras estaciones climatológicas más próximas para las que no es posible calcular una correlación fiable por su corto o nulo periodo común de observación con la problema. Por ello, y para tratar de aprovechar el máximo de información climática disponible, se desarrolló un método basado en la interpolación de datos estandarizados de las estaciones vecinas a la que se trata de homogeneizar, de manera similar al utilizado por PAULHUS y KOHLER (1952) para el relleno de datos ausentes en los boletines de precipitación diaria. Este método se implementó en forma de paquete (denominado Climatol) para la aplicación estadística R (de código abierto y multiplataforma), y en sus primeras versiones (GUIJARRO, 2004), si bien permitía la comparación de cada serie con su referencia sintética para detectar las anomalías puntuales, saltos en los valores medios y tendencias no achacables al clima, únicamente podía realizar de modo automático la corrección de errores aislados y el relleno de las lagunas de las series.

La experiencia acumulada desde entonces ha llevado a la conclusión de que la homogeneización de una base de datos climatológica es una tarea bastante ardua, lo que, unido a la constatación de que debería repetirse cada dos o tres años (al irse alargando las series con nuevos datos), ha llevado a implementar una nueva función, disponible desde la versión 2.0, que automatiza también la corrección de los saltos anómalos en los promedios de las series, atribuibles a cambios súbitos en las condiciones de observación.

En esta comunicación se pretende presentar sucintamente esta metodología, así como los resultados de su aplicación a las series termoplumiométricas de Baleares, estudiando el impacto de la homogeneización sobre la evaluación de las tendencias de la temperatura y la precipitación en el archipiélago.

## **2. METODOLOGÍA**

El método no presupone la existencia de series homogéneas en la base de datos a tratar, sino que todas son susceptibles de contener inhomogeneidades. Tampoco incorpora datos del historial de las estaciones (metadatos), que ofrecen información sobre los posibles momentos en que puede haber alteraciones en las series, porque estos historiales frecuentemente están

incompletos o no están al alcance del investigador. (Siempre existirá la posibilidad de usar esa información a posteriori, para contrastar los resultados de la homogeneización).

Los datos pueden ser series simples de cada observatorio (como sería el caso si se tratara de valores anuales), o conjuntos de series (valores mensuales o estacionales). En cualquier caso se leen cronológicamente, estación por estación, junto con las coordenadas e identificadores de cada una de ellas.

A continuación se procede serie a serie, calculando todos sus datos (como si no existieran) a partir de los de las demás estaciones, mediante medias ponderadas de valores estandarizados. Además de la estandarización propiamente dicha, también se puede optar por dividir los valores por sus respectivas medias, lo que resulta más recomendable para el tratamiento de la precipitación y otras variables con un cero natural. La ponderación se realiza mediante la función  $a/(a + d^2)$ , donde  $d$  es la distancia a cada estación, y  $a$  un parámetro de forma que permite modular la influencia relativa de las estaciones más próximas respecto de las más alejadas. Cuando las series están incompletas, sus promedios no son comparables, pero se vuelven a calcular tras el relleno de las lagunas, proceso que se repite hasta alcanzar valores estables.

Después se procede a estudiar las diferencias entre cada serie estandarizada y su réplica estimada en función de las demás, que sirve como serie de comparación. Opcionalmente se pueden substituir los valores anómalos por los estimados, y también visualizar gráficos de diagnóstico de cada una de las series. Las versiones anteriores del programa terminaban aquí con la grabación de un fichero con los datos corregidos y las lagunas rellenadas, pero ahora se continúa aplicando distintas pruebas de detección de saltos en la media, y asignando puntuaciones variables a cada término de la serie según la significación estadística obtenida. Si la máxima puntuación total tras aplicar todas las pruebas supera un umbral definido por el usuario, la serie se fragmenta en el punto de dicho máximo, traspasando toda la información posterior a una serie "hija". Este proceso se repite hasta que no se fragmenta ninguna serie, pues la corrección de algunas de ellas puede influir en la evaluación de la homogeneidad de las demás, y además las series fragmentadas pueden ser objeto de nuevos cortes si presentan más de un salto en la media.

Todo el procedimiento anterior se puede repetir tantas veces como desee el usuario, con lo que es posible eliminar al principio únicamente los errores y saltos en la media más patentes, y refinar después el proceso (aceptando niveles de significación más bajos) para corregir progresivamente las demás inhomogeneidades. Como productos de todas estas operaciones se obtiene un fichero con los datos corregidos y rellenados de todas las series resultantes, y un fichero de texto con una relación de las correcciones efectuadas y los niveles de significación de las pruebas de homogeneidad. Opcionalmente, también se obtienen gráficos diagnósticos de todas las series iniciales, y representaciones de las series finales, distinguiendo los datos originales, los rellenados, y las series fragmentadas a que hayan podido dar lugar. En este caso, como el número de gráficos puede llegar a ser muy grande, también se genera un fichero LaTeX que permite incluirlos todos en un solo documento.

### 3. APLICACIÓN A SERIES DE BALEARES Y DISCUSIÓN DE LOS RESULTADOS

#### 3.1. Precipitaciones

Se han recopilado los totales anuales de precipitación de las Islas Baleares que contaban con un mínimo de 15 años de observación en el periodo 1961-2007, lo que ha dado lugar a un conjunto de 231 series (198 de Mallorca, 17 de Menorca, 12 de Ibiza, y 4 de Formentera).

Primeramente se realizó una aplicación exploratoria para estudiar las frecuencias de las máximas puntuaciones de posibles saltos en la media acumulados por las diferentes pruebas utilizadas que, como se observa en la figura 1, tienen un máximo secundario en el valor 5, con una cola que se extiende hacia la derecha hasta alcanzar valores de alrededor de 13. Al no haber una diferencia clara entre estaciones homogéneas e inhomogéneas, el umbral de aceptación será, necesariamente, arbitrario. Aquí se aplicó el programa de homogeneización automática con tres niveles de exigencia: dos iniciales con niveles de corte de 8 y 7 puntos sucesivamente, y una tercera en la que ya no se realizan pruebas de saltos en la media, sino que se limita al relleno de lagunas definitivo. De este modo el número final de estaciones ha aumentado de 231 a 334, tras 103 cortes en las series. Se han fragmentado un total de 70 series, de las que 43 se han dividido en dos, 22 en tres, 4 en 4, y una serie en 5 fragmentos. Además se corrigieron 6 valores anómalos de una de las series.

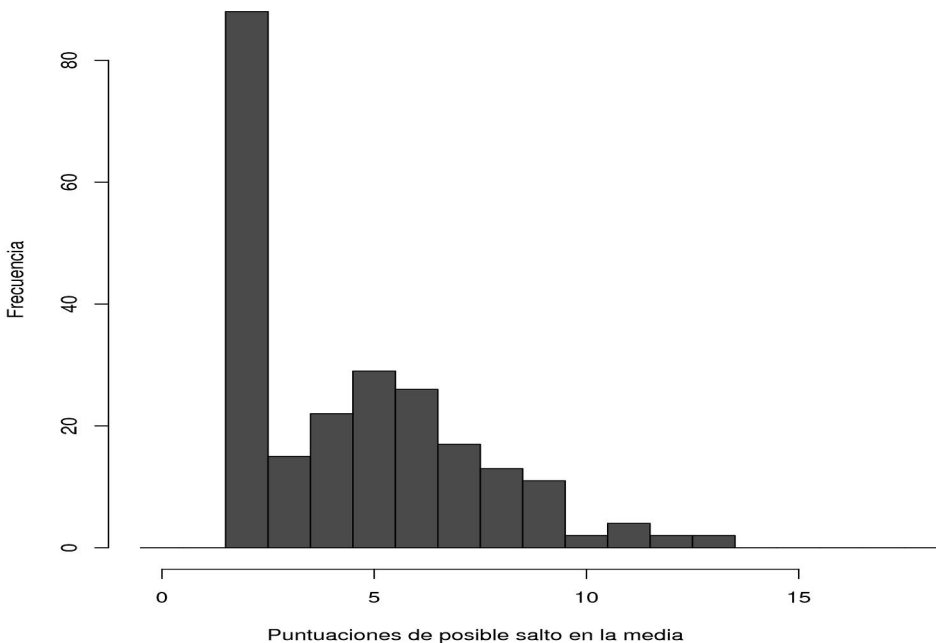


Fig. 1: Histograma de las puntuaciones máximas de posible salto en la media de cada serie de precipitación.

Anteriormente se había aplicado un proceso dedicado únicamente al relleno de lagunas, lo que nos va a permitir comparar el impacto que el proceso de homogeneización va a tener en la evaluación de las tendencias. Éstas se han calculado para los últimos 40 años (periodo 1968-2007), y se resumen, expresadas en %/siglo, en el cuadro estadístico de la tabla 1.

	Mínimo	Cuartil 1	Mediana	Media	Cuartil 3	Máximo
Sin homogeneizar	-384,3	-52,4	-40,6	-39,5	-27,7	163,4
Homogeneizadas	-83,6	-45,2	-38,0	-37,3	-29,3	40,8

Tabla 1: RESUMEN ESTADÍSTICO (VALORES MEDIOS, EXTREMOS Y CUARTILES) DE LAS TENDENCIAS (EN %/SIGLO) DE LA PRECIPITACIÓN ANUAL DE 231 SERIES DE BALEARES EN LOS ÚLTIMOS 40 AÑOS (1968-2007).

Queda patente cómo los valores extremos sufren severas correcciones, mientras que los valores centrales (la media y la mediana) resultan poco modificados. Como indicador de la reducción de la variabilidad de las tendencias, véase que el rango intercuartílico pasa de 24,7 a 15,9 %/siglo.

La figura 2 muestra las anomalías de precipitación anual de las 231 estaciones pluviométricas estudiadas, antes y después de la homogeneización (estas últimas, desplazadas 2000 mm hacia abajo para permitir la comparación). La existencia de una serie fuertemente anómala en el conjunto original domina la gráfica, de forma que aparte de su supresión en el conjunto homogeneizado, visualmente no se aprecian muchas más diferencias. La serie anómala presentaba varios errores en algunos valores mensuales de seis años, producidos por un problema informático, y los valores anuales correspondientes fueron corregidos automáticamente durante el proceso.

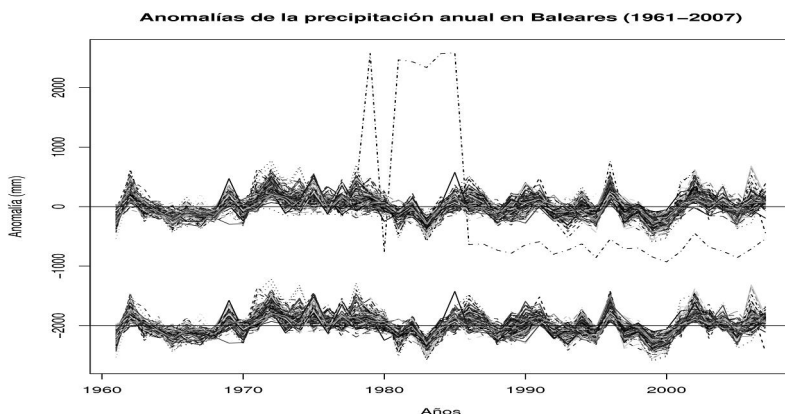


Fig. 2: Anomalías de las 231 series de precipitación anual de Baleares, antes de la homogeneización y después de la misma (desplazadas 2000 mm hacia abajo).

En los mapas de la figura 3 se muestra la distribución geográfica de las tendencias mediante segmentos verticales y horizontales, que representan las tendencias superiores e inferiores a la media respectivamente, con trazos de longitud proporcional a su desviación respecto a dicha media general. Se observa en ellos cómo las tendencias antes de la homogeneización presentaban una elevada variabilidad espacial, que le conferían un aspecto de aleatoriedad. Tras la homogeneización, la distribución espacial de las tendencias es mucho más homogénea, concentrándose las tendencias más negativas en el noroeste de Mallorca (en la Sierra de Tramontana, donde la precipitación alcanza los valores más altos por efecto orográfico).

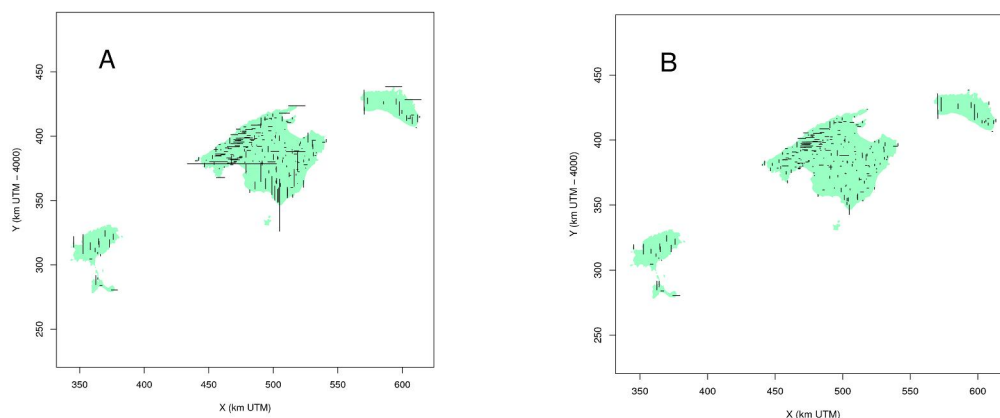


Fig. 3: Distribución espacial de las tendencias de la precipitación anual en Baleares (1968-2007), antes (A) y después (B) de la homogeneización. Las tendencias superiores a la media (las menos negativas) se han representado mediante trazos verticales, y las inferiores (más negativas) con trazos horizontales. La longitud de los trazos es proporcional a la desviación respecto a la tendencia media.

Un trabajo anterior de cartografía de las tendencias de la precipitación en Baleares (GUIJARRO, 2002) parecía apuntar a una mayor disminución de los niveles pluviométricos en Menorca y norte de Mallorca que en el resto, donde las islas de Ibiza y Formentera especialmente parecían verse menos afectadas por las tendencias negativas. Estas diferencias, aunque eran coherentes con la observada disminución de la actividad ciclónica en el Golfo de Génova (GUIJARRO, 2003; GUIJARRO et al., 2006), no aparecen en los mapas mostrados aquí. No obstante, en el artículo antes citado (GUIJARRO, 2002) se ponía de manifiesto lo cambiantes que pueden ser las tendencias según el periodo elegido para su cómputo, por lo que serán necesarios más estudios para avalar una u otra hipótesis.

### 3.2. Temperaturas

En el caso de las temperaturas se ha trabajado con series mensuales, seleccionando todas las que tenían un mínimo de 10 años de funcionamiento en el mismo periodo usado para las precipitaciones (1961-2007). No se han tratado directamente las series anuales porque la existencia de meses aislados sin datos impedía calcular las correspondientes medias anuales, con la consecuencia de que alguno de los primeros años del periodo considerado se quedaba

sin ningún dato, imposibilitando continuar con el análisis. De este modo se ha podido contar con un número inicial de 66 estaciones (49 de Mallorca, 9 de Menorca, 7 de Ibiza y 1 de Formentera).

Un primer análisis exploratorio ha permitido obtener las puntuaciones de posible salto en la media que se reflejan en la figura 4. Se observa en ella que no hay una separación clara entre estaciones homogéneas (con una puntuación baja) e inhomogéneas (valores altos), por lo que la decisión de fijar un umbral de aceptación va a ser también (como sucedía con las series de precipitación anual) arbitraria. En este caso se hizo una homogeneización conservadora, admitiendo como homogéneas las estaciones con una puntuación inferior a 8 en la primera fase y a 7 en la segunda, y reservando una tercera fase para el relleno definitivo de las lagunas de observación.

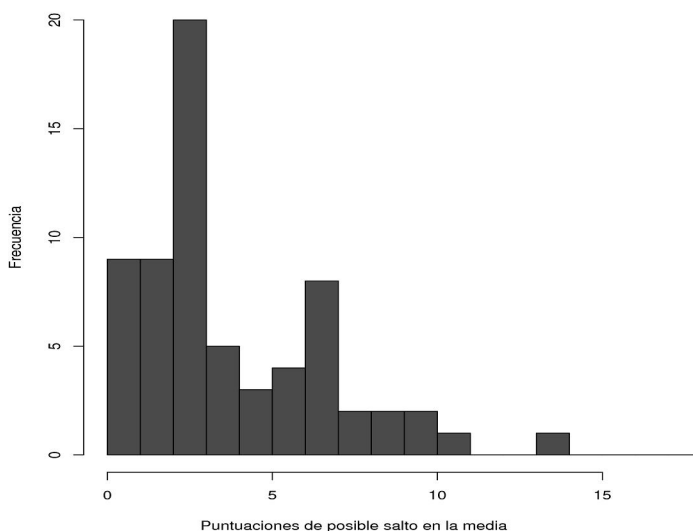


Fig. 4: Histograma de las puntuaciones máximas de posible salto en la media de cada estación.

El proceso corrigió 14 anomalías mensuales de 9 estaciones, y realizó 13 cortes a 12 de las estaciones (11 estaciones con un corte, y una estación con dos). Con los valores mensuales completos tras el proceso de homogeneización, se procedió a calcular las tendencias de los promedios anuales de los últimos 40 años (periodo 1968-2007). Una nueva aplicación del programa solicitando únicamente la corrección de errores y el relleno de los valores ausentes permite calcular esas mismas tendencias cuando no se ha corregido ninguna serie, con objeto de estudiar el impacto que la homogeneización ha tenido en la evaluación de las tendencias. En la tabla 2 se ofrece un resumen estadístico de estas tendencias (expresadas en °C/siglo) sin homogeneización y con ella.

	Mínimo	Cuartil 1	Mediana	Media	Cuartil 3	Máximo
Sin homogeneizar	0,68	1,45	1,74	2,05	2,33	4,98
Homogeneizadas	0,03	0,64	0,90	1,10	1,37	4,32

Tabla 2: RESUMEN ESTADÍSTICO (VALORES MEDIOS, EXTREMOS Y CUARTILES) DE LAS TENDENCIAS (EN C/SIGLO) DE LA TEMPERATURA MEDIA ANUAL DE 66 SERIES DE BALEARES EN LOS ÚLTIMOS 40 AÑOS (1968-2007).

Se observa cómo tanto la mediana como la media de las tendencias homogeneizadas disminuyen casi a la mitad de los valores previos a la homogeneización, mientras que la dispersión, medida por el rango intercuartílico, también disminuye (aunque mucho menos), al pasar de 0,88 a 0,73.

En la figura 5 se han representado conjuntamente las anomalías de las 66 series de temperatura media anual antes de la homogeneización (desplazadas 2 C hacia arriba) y después de la misma (desplazadas 2 C hacia abajo). La dispersión de los valores ha disminuído después de la homogeneización, pero todavía se observan series discordantes en algunos puntos (como en los primeros tres años de una de las series).

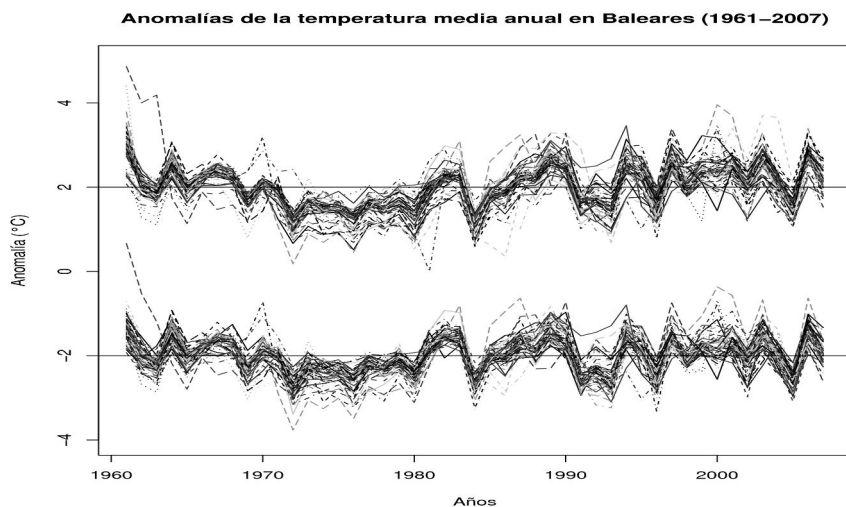


Fig. 5: Anomalías de las 66 series de temperatura media anual de Baleares, antes de la homogeneización (desplazadas 2 °C hacia arriba) y después de la misma (desplazadas 2 C hacia abajo).



La distribución espacial de las tendencias se ha representado en los mapas de la figura 6. Antes de la homogeneización (arriba) se observa una elevada variabilidad espacial, apareciendo tendencias grandes y pequeñas bastante mezcladas en todo el territorio. Después de la homogeneización (abajo) se puede ver una mayor uniformidad (aunque queda una tendencia discordante en el noroeste de Mallorca), correspondiendo las mayores tendencias a Ibiza y Formentera principalmente, y las menores a la zona más oriental de Menorca.

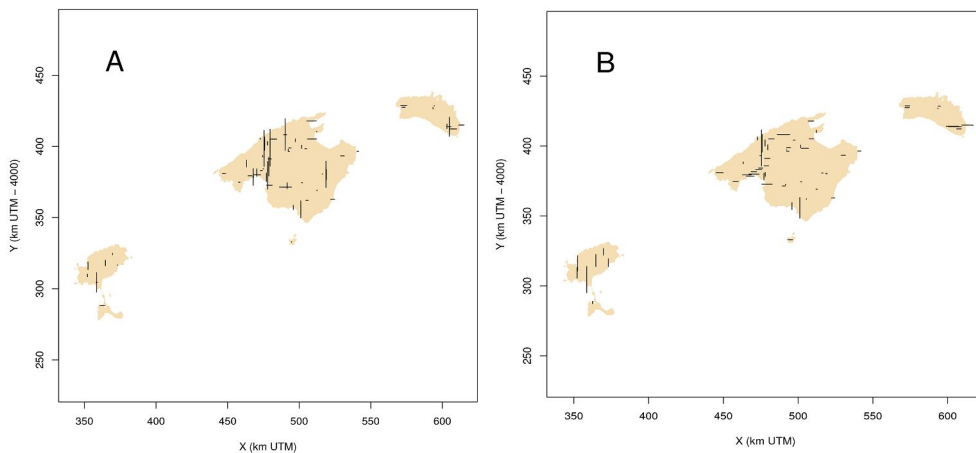


Fig. 6: Distribución espacial de las tendencias de la temperatura anual en Baleares (1968-2007), antes de la homogeneización (A), y después de la misma (B). Las tendencias superiores a la media se han representado mediante trazos verticales, y se han usado trazos horizontales para las demás. La longitud de los trazos es proporcional a la desviación respecto a la tendencia media.

#### 4. CONCLUSIONES

La nueva función facilita enormemente la tarea de homogeneización de las series de una base de datos climatológica, automatizando procesos que sería muy laborioso realizar manualmente. No obstante, la labor del investigador sigue siendo fundamental, pues los resultados de la homogeneización van a depender en gran medida de diversos parámetros opcionales, y particularmente de los umbrales de significación que se establezcan para corregir datos anómalos puntuales o saltos en la media de las series.

Las inhomogeneidades de las series climatológicas pueden tener mucha influencia en la evaluación de las tendencias, generando en ocasiones valores contradictorios en una misma zona geográfica. La homogeneización permite obtener tendencias espacialmente más coherentes, y por tanto más confiables. En el caso de las temperaturas medias de Baleares las tendencias observadas han disminuido mucho tras la aplicación de la nueva función, y serán necesarias nuevas aplicaciones de la misma con otros parámetros para confirmar o corregir estos resultados.

## 5. REFERENCIAS

- AGUILAR, E.; AUER, I.; BRUNET, M.; PETERSON, T.C. y WIERINGA, J. (2003). *Guidelines on climate metadata and homogenization*. WCDMP-No. 53, WMO-TD No. 1186. World Meteorological Organization, Geneve.
- GUIJARRO, J.A. (2002). Tendencias de la precipitación en el litoral mediterráneo español. En: GUIJARRO et al. (Eds.). *El agua y el clima*. Asociación Española de Climatología, A-3, pp. 237-246.
- GUIJARRO, J.A. (2003). El flujo geostrófico superficial en el Mediterráneo Balear durante el periodo 1948-2002. *Rev. climatol.*, 3, pp. 45-59.
- GUIJARRO, J.A. (2004). Climatol: Software libre para la depuración y homogeneización de datos climatológicos. En: GARCÍA-CODRON et al. (Eds.). *El clima, entre el Mar y la Montaña*. Asociación Española de Climatología, A-4, pp. 493-502.
- GUIJARRO, J.A.; JANSÁ, A. y CAMPINS, J. (2006). Time variability of cyclonic geostrophic circulation in the Mediterranean. *Advances in Geosciences*, 7, pp. 4549.
- PAULHUS, J.L.H. y KOHLER, M.A. (1952). Interpolation of missing precipitation records. *Month. Weath. Rev.*, 80, pp. 129-133.
- PETERSON, T.C.; EASTERLING, D.R.; KARL, T.R.; GROISMAN, P.; NICHOLLS, N.; PLUMMER, N.; TOROK, S.; AUER, I.; BOEHM, R.; GULLETT, D.; VINCINT, L.; HEINO, R.; TUOMENVIRTA, H.; MESTRE, O.; SZENTIMREY, T.; SALINGER, J.; FØRLAND, E.J.; HANSEN-BAUER, I.; ALEXANDERSSON, H.; JONES, P.D. y PARKER, D. (1998). Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology*, 18, pp. 1493-1517.