

COMPARACIÓN DE MÉTODOS DE HOMOGENEIZACIÓN DE SERIES: PRIMEROS RESULTADOS DEL PROYECTO MULTITEST

José Antonio GUIJARRO¹, Peter DOMONKOS², José Antonio LÓPEZ¹,
Enric AGUILAR², Manola BRUNET²

¹Agencia Estatal de Meteorología.

²Center for Climate Change. Universitat Rovira i Virgili.

jguijarrop@aemet.es, dpeterfree@gmail.com, jlopezd@aemet.es, Enric.Aguilar@urv.cat,
manola.brunet@urv.cat

RESUMEN

El estudio de la variabilidad climática, tanto de sus valores medios como de los extremos, requiere disponer de series de calidad, libres de influencias no climáticas debidas a cambios en las condiciones de observación. Para corregir las series eliminando estas alteraciones espúrias y reteniendo únicamente la variabilidad climática se emplean diversos métodos de homogeneización, algunos de los cuales se han implementado en paquetes de software para facilitar su uso por parte de la comunidad climatológica. Si bien existe bastante literatura científica sobre la bondad de los diversos métodos existentes, incluyendo los resultados de la pasada Acción COST ES0601, la evolución y perfeccionamiento de las aproximaciones existentes, así como la aparición de otras nuevas, evidencian la necesidad de continuar realizando pruebas comparativas de la bondad de sus resultados. El proyecto MULTITEST tiene como objetivo aplicar los principales métodos de homogeneización automáticos existentes a conjuntos de datos mensuales de prueba, tanto de temperatura como de precipitación, con características que simulan redes de observación de diferentes zonas climáticas. En esta comunicación se describe el procedimiento seguido para incluir inhomogeneidades aleatorias en las series y aplicar los paquetes de homogeneización a las mismas, mostrando los resultados obtenidos hasta ahora.

Palabras clave: comparación, metodologías, homogeneización de series climáticas.

ABSTRACT

The study of climate variability, both their means and extreme values, requires the use of high quality observational series, free from non-climatic influences due to changes in the conditions of observation. Several methods of homogenization are used to correct the series eliminating these spurious changes and leave only the climate, some of which have been implemented in software packages to facilitate its use by the climatological community. While there is enough scientific literature about the goodness of the various methods, including the results of last COST Action ES0601, evolution and improvement of existing approaches and the emergence of new ones points at the need to continue making comparative tests of the goodness of their results. The MULTITEST project aims at applying the main current automatic methods of homogenization to monthly test data sets of both temperature and precipitation,

with features that simulate observation networks of different climate zones. The procedures followed to include random inhomogeneities in the series and to apply homogenization packages to them are described in this communication, showing the results obtained so far.

Key words: comparison, methodologies, homogenization of climatic series.

1. INTRODUCCIÓN

Las series observacionales de datos meteorológicos son de gran interés para estudiar la variabilidad climática, pero al estar afectadas por alteraciones en las condiciones de observación (cambios de instrumentación, traslados, etc) o del entorno de los observatorios (Aguilar *et al.*, 2003), es preciso corregirlas para aumentar la fiabilidad de las conclusiones derivadas de su estudio.

Las principales metodologías propuestas para ello fueron comparadas en un extenso trabajo desarrollado en el marco de la Acción COST ES0601 “HOME” (Venema *et al.*, 2012), pero se hacía necesario un nuevo ejercicio de intercomparación de metodologías por al menos tres motivos:

- 1) Muchos métodos han cambiado desde entonces al haber introducido mejoras.
- 2) El elevado número de redes problema en “HOME” hizo que muchos participantes proporcionaran resultados parciales, con lo que la base de comparación no fue homogénea.
- 3) Las bases de datos utilizadas en “HOME” simulaban el clima de Centroeuropa, restringiendo así la representatividad de los resultados de cara a una aplicación de los métodos de homogeneización a otras partes del globo.

El proyecto MULTITEST pretende actualizar los resultados de “HOME” realizando una comparación de los resultados obtenidos al aplicar los principales métodos existentes a gran cantidad de conjuntos de series mensuales de temperatura y precipitación. No obstante, dada la imposibilidad práctica de repetir un trabajo con tantos participantes como en “HOME”, esta comparación tiene necesariamente que restringirse a métodos ya implementados en forma de programas de ordenador que puedan ser aplicados de forma automática. Esta limitación creemos que se compensa con creces por el hecho de poder efectuar las comparaciones sobre un mismo conjunto de un gran número de series problema, y por otra parte ese automatismo resulta necesario en climatología operativa, donde se debería proceder a la homogeneización de las bases de datos climatológicos al menos una vez al año, al verse continuamente actualizadas con la adquisición de nuevos datos.

2. METODOLOGÍA

La metodología general aplicada para comparar los resultados de los distintos métodos sigue el esquema de un trabajo anterior (Guijarro, 2011) en el que, partiendo de unas redes maestras de 100 series homogéneas de 60 años de longitud (720 datos mensuales cada una), se seleccionan aleatoriamente muestras de 10 de ellas, de las que a las primeras 5 o a todas ellas, según los casos, se les aplican diversos tipos de inhomogeneidad.

Estas muestras inhomogeneizadas son presentadas como problema a resolver por los distintos métodos de homogeneización ensayados, y los resultados de éstos se comparan con la muestra homogénea original, calculando para cada serie problema el error cuadrático medio y las diferencias entre sus tendencias, valores medios y desviaciones típicas, estadísticos usados para evaluar en términos comparativos la bondad de cada método.

2.1. Redes termométricas de datos de prueba

En el antecedente de este trabajo (Guijarro, 2011) se generaron tres redes maestras basadas en las temperaturas máximas medias mensuales de la zona central de la cuenca del Duero. Tras definir un dominio espacial con 100 localizaciones distribuidas al azar, se obtuvo una primera serie aleatoria con los citados valores medios mensuales, y la serie más próxima se generó añadiendo ruido blanco a la anterior. Este proceso se repitió hasta completar las 100 series, con distinto nivel de ruido para cada una de las tres redes (y por tanto con diferentes niveles de correlación cruzada), añadiendo después variabilidad en el ciclo anual y una tendencia de $2\text{ }^{\circ}\text{C/siglo}$ (para más detalles, véase Guijarro, 2012).

2.2. Redes pluviométricas de datos de prueba

Las series de precipitación presentan una complejidad mucho mayor que las de temperatura, puesto que a una mayor variabilidad espacial añade una distribución de probabilidades sesgada, alejada de la normal y acotada inferiormente por el valor cero, que además puede aparecer frecuentemente en climas secos.

Para contemplar la aplicabilidad de los métodos a diferentes partes del globo se construyeron redes maestras sintéticas de precipitaciones mensuales simulando tres tipos de clima diferente: oceánico (Cfb según la clasificación de Köppen), mediterráneo (Csa) y monzónico (Am) (López *et al.*, 2016). En la Figura 1 se pueden ver los regímenes pluviométricos medios de estas tres redes.

2.3. Ejecución de las pruebas

La ejecución automatizada de las pruebas se implementó en un programa de R (R Development Core Team, 2013) que va seleccionando aleatoriamente las muestras, les aplica las inhomogeneidades y, para cada método ensayado, graba las series problema en archivos con el formato requerido para cada uno de ellos, arranca el correspondiente programa, lee los resultados entregados, y los compara con las series originales para calcular sus errores.

Se han diseñado hasta ahora cinco tipos de inhomogeneidades, de los que los tres primeros son condiciones experimentales para discriminar la respuesta de los distintos métodos a situaciones ideales, mientras que los dos últimos se corresponden más a lo que cabe esperar en condiciones reales:

- a) Dos saltos de $2\text{ }^{\circ}\text{C}$ en diferentes posiciones fijas en las tres primeras series, y un salto de la misma magnitud en posición aleatoria en las series cuarta y quinta. Las demás series no se alteran.
- b) Como en a), pero con una fuerte estacionalidad en los fragmentos inhomogéneos.

- c) Las cinco primeras series se alteran con combinaciones de “plataformas” (saltos de corta duración) y tendencias locales, también de 2 °C de magnitud.
- d) Todas las series pueden sufrir saltos en número, magnitud y posición aleatorios, con una frecuencia de 5 saltos por siglo.
- e) Como en d), pero con estacionalidad de magnitud aleatoria.

En las tres primeras configuraciones se dejan los últimos 10 años inalterados para que los métodos puedan homogeneizar las series hacia atrás con suficiente robustez, pero en las dos últimas este periodo se reduce a un mínimo de 5 años.

Para cada una de estas configuraciones y cada red maestra, se realizaron 100 pruebas con cada método, inicializando el generador de números aleatorios con la misma semilla para garantizar que los métodos se prueban sobre las mismas series y con las mismas alteraciones. Se obtienen así un mínimo de 500 series homogeneizadas con cada método, red y configuración ensayados.

Los métodos de homogeneización que se han probado son los siguientes (etiquetados para su identificación en las figuras donde se mostrarán los resultados):

- 1) Climatol 3.0, correcciones constantes (cl1) y variables (Cl1). En el caso de las precipitaciones, cl1 usa una normalización por proporciones respecto a sus medias, mientras que Cl1 aplica una transformación logarítmica antes de estandarizar los valores.
- 2) ACMANT 3.0 en sus versiones para temperaturas (con estacionalidad de correcciones sinusoidal, A3s, o irregular, A3i), y para precipitaciones (A3p).
- 3) MASH 3.03 (MSH).
- 4) RHTest 4.0 con serie de referencia (Rhr) o sin ella (homogeneidad absoluta, Rha; RHR y RHA cuando además se aplica un ajuste por cuantiles). Con este método el usuario ha de proporcionar la serie de referencia, y aquí se le suministró la serie homogénea más próxima a cada serie problema en las configuraciones 1 a 3, y el promedio de las primeras diez series en las configuraciones 4 y 5.
- 5) USHCN 52d (US1).
- 6) HOMER 2.6 con distintas secuencias de detección/corrección (Hom para temperaturas, y Hor, HoA, HoB para precipitaciones).

Para más información sobre estos métodos ver <http://www.climatol.eu/TT-HOM>.

Las pruebas se realizaron en un ordenador con Linux (lubuntu 14.04), y para aplicar los métodos compilados para Windows se recurrió a la utilidad *wine*, mientras que para ejecutar HOMER de modo automático hubo que recurrir a *expect*.

3. RESULTADOS

Debido a limitaciones de espacio únicamente se mostrarán los errores típicos y de tendencia de las configuraciones con inhomogeneidades más realistas (d, e), en las que todas las series pueden verse afectadas por un número variable de saltos en la media de magnitud aleatoria. Estos resultados aparecen resumidos en las figuras 2 a 4, que contienen diagramas de caja (conteniendo un mínimo de 500 datos cada una) de los errores típicos (RMSE) de las series homogeneizadas, así como de las diferencias entre sus tendencias y las de las series originales. Cada método está identificado por el acrónimo indicado anteriormente en la lista de métodos, pudiéndose comparar entre

sí y con la primera caja (Inh), que muestra los errores de las series problema (inhomogéneas). Para facilitar la comparación visual se ha elegido un eje de ordenadas de dimensiones fijas, aunque deja fuera los errores más anómalos.

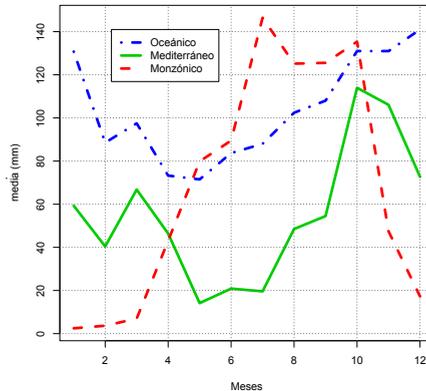


Fig. 1: Regímenes pluviométricos de las tres redes maestras de precipitación.

Las figuras 2 y 3 muestran los errores típicos de las series termométricas homogeneizadas y sus errores de tendencia respectivamente. Aunque se realizaron otras pruebas con configuraciones experimentales de inhomogeneidades, aquí se presentan únicamente los resultados de las pruebas más realistas, en las que las inhomogeneidades se pueden dar en todas las series, con una frecuencia media de 5 saltos por siglo.

4. DISCUSIÓN

Las pruebas efectuadas con las series termométricas muestran (Figura 2, fila superior) que cuando las correlaciones entre ellas son buenas (como fue el caso en las redes de la Acción COST “HOME”) todos los métodos ensayados, incluso los de homogeneidad absoluta (aunque son los peores con diferencia), proporcionan homogeneizaciones con menores errores que las series problema. Cuando las inhomogeneidades son constantes a lo largo del año (Figura 2, TA20r) los menores errores, sin distinción significativa entre ellos (las cajas se yuxtaponen), corresponden a Climatol (con corrección variable!), ACMANT y USHCN, seguidos de MASH, HOMER y RHTest. Por otra parte, la existencia de estacionalidad en las inhomogeneidades, cosa frecuente en el mundo real (Brunet *et al.*, 2011), discrimina más los resultados, dejando únicamente a ACMANT y Climatol en el grupo de los mejores, seguidos por HOMER y RHTest relativo con ajuste de cuantiles, cuando las correlaciones son altas. En caso contrario ACMANT va seguido de RHTest relativo (pero sin ajuste de cuantiles), MASH y los demás métodos.

Cuando las series tienen peores correlaciones (Figura 2, fila inferior) todos los resultados empeoran, incluso los de homogeneidad absoluta (porque estas series también son más ruidosas), que es mejor evitar. El ajuste de cuantiles de RHR también empeora los resultados, pero los demás métodos ofrecen pequeñas diferencias entre ellos. El orden medio de bondad en todas las pruebas ha sido: A3s, A3i, C11, MSH,

Rhr, US1, Hom, c11 y, ya a mucha distancia, RHR. También cabe señalar que la mayoría de métodos mejoran la homogeneidad de las series problema.

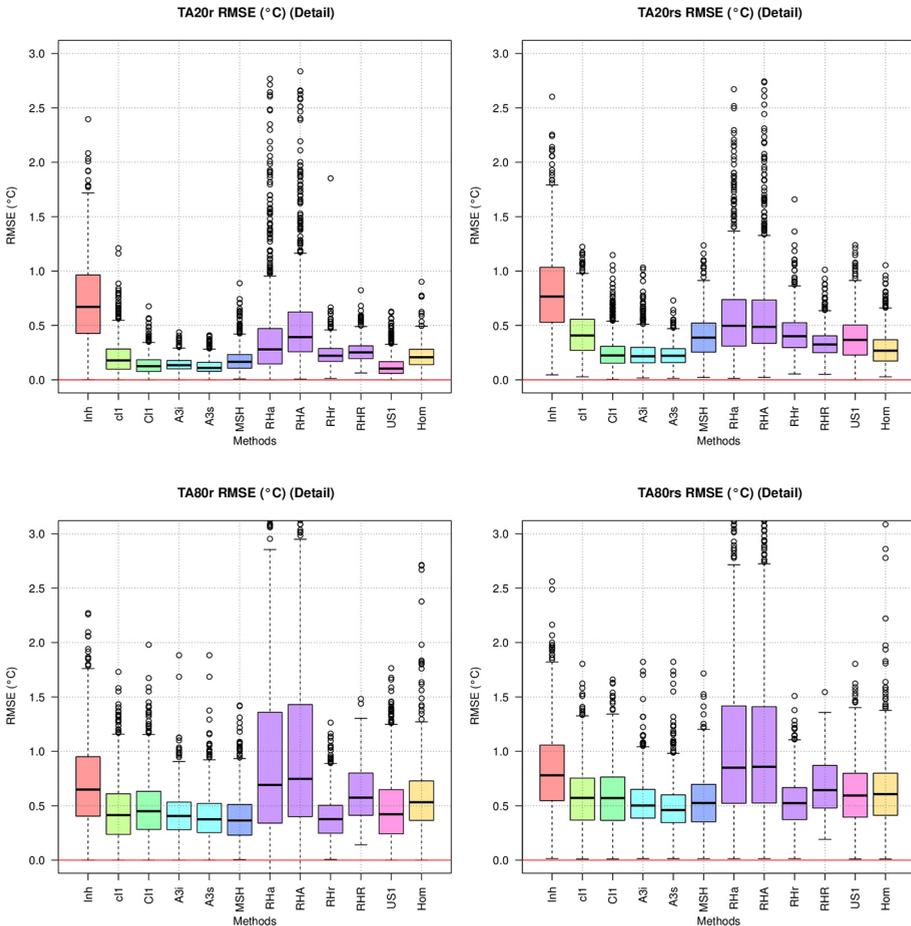


Fig. 2: Errores típicos de los métodos de homogeneización en redes termométricas. En la fila superior las correlaciones son mucho mejores que en la fila inferior. En la columna izquierda las inhomogeneidades son constantes a lo largo del año, mientras que en la derecha presentan variaciones estacionales.

En cuanto a las tendencias de las series homogeneizadas (Figura 3), todos los métodos (excepto los de homogeneidad absoluta cuando las correlaciones son bajas) disminuyen la dispersión de tendencias de las series problema. Con buenas correlaciones (fila superior) ACMANT, USHCN y MASH son los que reducen más esa dispersión, mientras que con correlaciones peores (fila inferior) las diferencias entre los métodos son menos acusadas. Es importante notar que el sesgo de las tendencias es nulo o despreciable, con excepción del de las series resultantes de HOMER, y que la existencia de estacionalidad en las inhomogeneidades (columna derecha) no cambia apreciablemente los resultados.

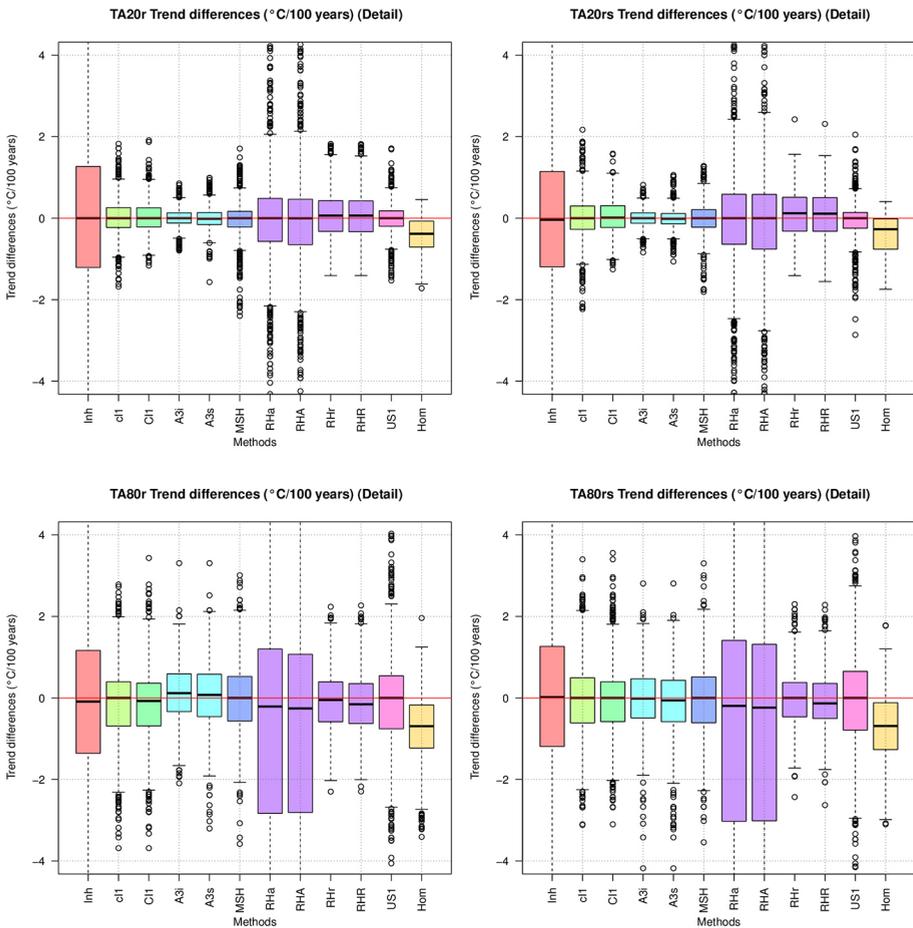


Fig. 3: Errores de los métodos de homogeneización en las tendencias de redes termométricas. En la fila superior las correlaciones son mucho mejores que en la fila inferior. En la columna izquierda las inhomogeneidades son constantes a lo largo del año, mientras que en la derecha presentan variaciones estacionales.

Como era de esperar, las series de precipitación han resultado más difíciles de homogeneizar (Figura 4). De hecho, USHCN ni siquiera se diseñó para ser usado con precipitaciones, pero, con objeto de incluirlo junto con los demás métodos, en este caso se procedió a aplicar una transformación logarítmica a los datos, del mismo modo que requiere RHTest explícitamente y usan otros programas internamente.

Con clima oceánico (Peirr) los menores errores típicos (RMSE) corresponden a ACMANT, Climatol con normalización por proporciones (pues la transformación logarítmica empeora mucho sus resultados) y HOMER. Estos mismos métodos siguen siendo los mejores con las series mediterráneas (PMcar) aunque la mejora respecto a las series inhomogéneas es ahora mucho menor. Y por último, con las series de clima monzónico ACMANT, HOMER y MASH, seguidos a corta distancia por Climatol

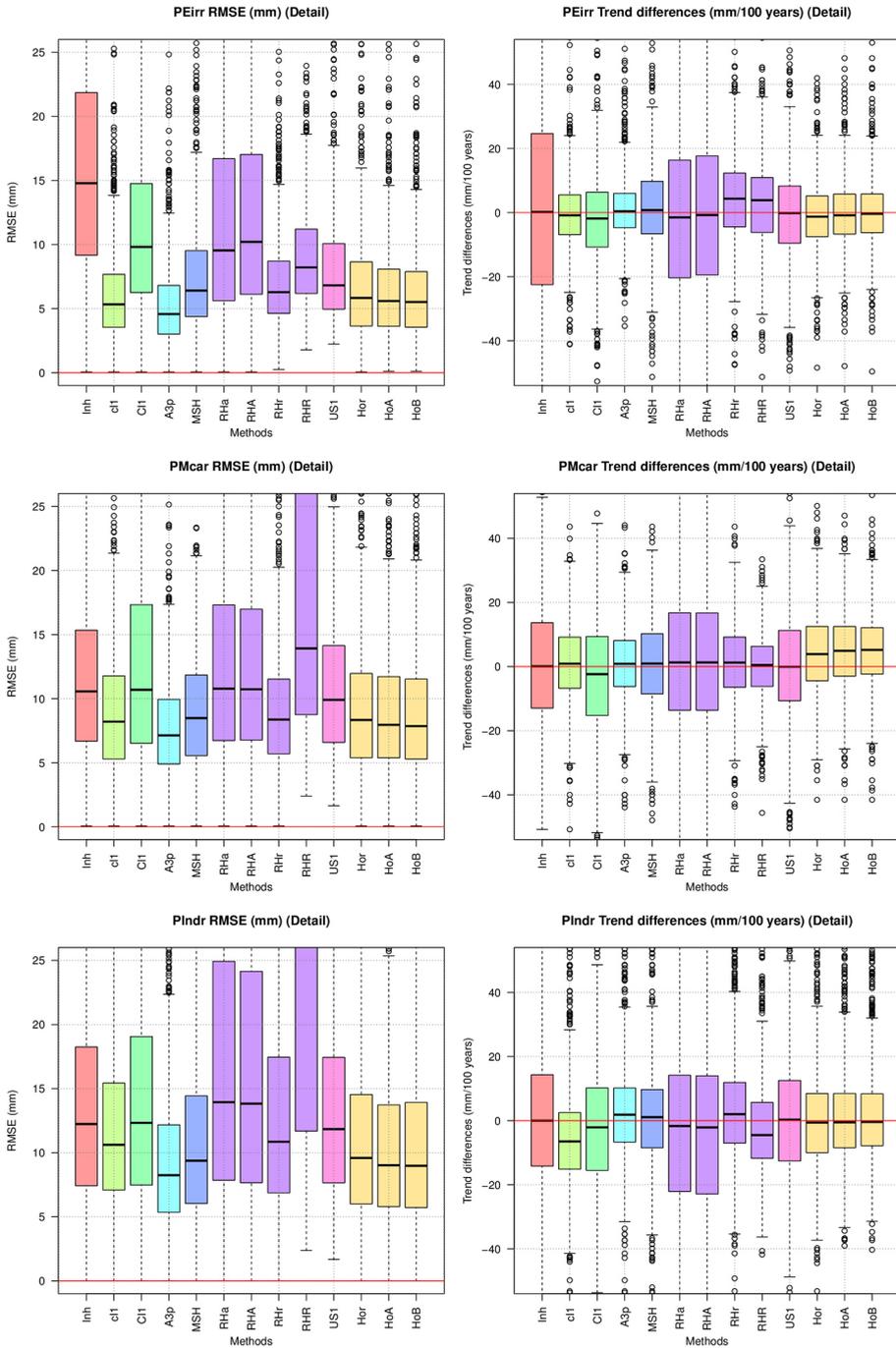


Fig. 4: Errores típicos (izda.) y de tendencia (dcha.) en las redes de precipitación de clima oceánico (arriba), mediterráneo (centro) y monzónico (abajo).

y RHTest, son los que proporcionan los mejores resultados, aunque la corrección de las inhomogeneidades resulta poco más que testimonial. Son destacables los malos resultados que proporciona el ajuste de cuantiles de RHTest, sobre todo en las series mediterráneas y monzónicas. Así pues, el orden medio de bondad en estas pruebas con precipitaciones ha sido: A3p, HoA, c11, MSH, Hor, HoB, RHr, US1, C11 y RHR (dejando aparte las versiones absolutas de RHTest).

En cuanto a las tendencias (columna derecha de la Figura 4), la mayoría de métodos siguen reduciendo la dispersión de valores de las series problema, con escasos sesgos en general, excepto los de RHTest en clima oceánico, HOMER en clima mediterráneo, y Climatol y USHCN en clima monzónico, que curiosamente presentan ese comportamiento particular en unos climas concretos y no en otros.

Finalmente cabe comentar que a algunos desarrolladores no les agrada que se realicen estas pruebas comparativas, alegando que mediante la automatización no se realiza una aplicación óptima de sus métodos. Pero esto puede aplicarse a todos ellos (excepto, quizá, a ACMANT, cuyas configuraciones son fijas), y esta creemos que es la única manera de poder comparar los resultados de los distintos métodos de modo homogéneo y estadísticamente robusto. Además, a los desarrolladores se les ofrece la posibilidad de mejorar los algoritmos que arrancan y configuran los parámetros de sus programas y, al finalizar el proyecto, tanto los resultados de las pruebas como los programas utilizados para las mismas serán accesibles públicamente en una página web, confiriendo el máximo de transparencia y reproducibilidad al proceso.

5. CONCLUSIONES

Este proyecto ha permitido probar los principales métodos de homogeneización actualmente disponibles sobre conjuntos de datos mensuales de temperatura y precipitación con características muy variadas, y sus resultados, junto con las tablas comparativas de características disponibles en <http://www.climatol.eu/tt-hom/>, serán de gran ayuda como guía para escoger las herramientas más adecuadas en cada caso.

En el conjunto de las pruebas presentadas aquí, ACMANT ha sido el que ha presentado los menores errores típicos, seguido por Climatol, MASH, RHTest, USHCN y HOMER en temperaturas, y por HOMER, Climatol, MASH, RHTest y USHCN en precipitaciones.

En próximos trabajos está previsto estudiar la influencia de una mayor longitud en las series, la existencia de datos ausentes y la presencia de variaciones estacionales en las inhomogeneidades de las series de precipitación.

AGRADECIMIENTOS

El proyecto MULTITEST está financiado por el Ministerio de Economía y Competitividad (CGL2014-52901-P). Enric Aguilar y Manola Brunet agradecen el apoyo del proyecto europeo titulado: Uncertainties in Ensembles of Regional Reanalyses (UERRA, FP7-SPACE-2013-1, No. de referencia: 607193).

REFERENCIAS

- Aguilar, E., I. Auer, M. Brunet, T. C. Peterson, J. Wieringa (2003). Guidelines on climate metadata and homogenization. WCDMP-No. 53, WMO-TD No. 1186. World Meteorological Organization, Geneva.
- Brunet M., J. Asin, J. Sigró, M. Bañón, F. García, E. Aguilar, J. Palenzuela, T. Peterson, P. Jones (2011). The minimization of the screen bias from ancient Western Mediterranean air temperature records: an exploratory statistical analysis. *Int. J. Climatol.*, 31, pp. 1879-1895.
- Guijarro, J. A. (2011). Influence of network density on homogenization performance. Seventh Seminar for Homogenization and Quality Control in Climatological Databases jointly organized with the Meeting of COST ES0601 (HOME) Action MC Meeting, Budapest, 24-27/October, WCDMP-No. 78, pp. 11-18.
- Guijarro, J. A. (2012). Después de COST/ES0601: Comparación automática de métodos de homogeneización de series. 8º congreso de la Asociación Española de Climatología, ISBN 978-84-695-4331-3, pp. 381-390.
- López, J. A., J. A. Guijarro, E. Aguilar, P. Domonkos y M. Brunet (2016). Una propuesta metodológica para la generación de redes de precipitación simuladas a partir de redes de precipitación observadas en el marco del proyecto MULTITEST. En este mismo volumen.
- R Development Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Venema, V., O. Mestre, E. Aguilar, I. Auer, J. Guijarro, P. Domonkos, G. Vertacnik, T. Szentimrey, P. Stepanek, P. Zahradnicek, J. Viarre, G. Müller-Westermeier, M. Lakatos, C. Williams, M. Menne, R. Lindau, D. Rasol, E. Rustemeier, K. Kolokythas, T. Marinova, L. Andresen, F. Acquaotta, S. Fratianni, S. Cheval, M. Klancar, M. Brunetti, C. Gruber, M. Prohom-Duran, T. Likso, P. Esteban, T. Brandsma (2012). Benchmarking homogenization algorithms for monthly data. *Climate of the Past*, 8, 89-115.