

# INFORME SOBRE EL CONTROL DE CALIDAD DE DATOS CLIMÁTICOS Y DISEÑO DE PROCESOS AUTOMATIZADOS PARA LA ESPACIALIZACIÓN DE VALORES DE TEMPERATURA MÍNIMA Y MÁXIMA PARA EL PROYECTO CLIMA II.

## Introducción

Este informe se presentan los resultados de la primera etapa de control de calidad de datos, en la que se proponía la realización de una serie de tareas para asegurar la calidad de los datos originales obtenidos en los observatorios, tanto observados como modelados para periodos futuros.

En relación con los datos observados, se propuso desarrollar los trabajos en tres fases:

- 1.1 Comprobación y test de la correlación entre los datos observados de los periodos 1961-2000 y 2000-2017 y los datos extraídos de las capas ráster para dichos periodos.
- 1.2 Comparación de los datos observados con bases de datos climáticas de referencia con calidad y homogeneidad contrastadas (AEMET y CRU).
- 1.3 Desarrollo de contrastes de hipótesis entre los periodos 1961-2000 y 2000-2017 de los datos observados.

Durante el desarrollo de los trabajos se han introducido algunos cambios metodológicos que se explicitan en los apartados que siguen. Los cambios más notables han sido los siguientes: En primer lugar, debido a la importancia de la **comprobación de los aumentos observados en la variable media mensual de las temperaturas máximas diarias**, el primer paso ha sido aplazado para llevarlo a cabo posteriormente. En segundo lugar, los datos de la CRU no se han utilizado debido a su baja resolución espacial, siendo sustituidos por dos conjuntos de datos en rejilla con mayor resolución que se detallan en el documento (MOTEDAS). Finalmente, el tercer paso, tras la comprobación de la existencia de inhomogeneidades en las series, ha sido sustituido por un análisis de las tendencias.

## Selección de fuentes de datos, estaciones y series para el control de calidad

Previamente a las comprobaciones, ha sido necesario hacer una selección de los datos de referencia. Se ha optado por empezar las comprobaciones a escala mensual, sobre la variable de temperatura media de las máximas diarias. Se han utilizado dos fuentes de datos.

## Fuentes de datos utilizadas

- Estaciones de la AEMET (1961-2017).

Se trata de datos de la La Red Nacional de Observación en Superficie (RENOMS), conocida también como red mesoescalar de la AEMET, que han sido descargados mediante un script Python desarrollado *ad hoc* sobre la API REST AEMET OpenData (Anexo I). Este script guarda los datos en un archivo CSV, que posteriormente es importado en una base de datos creada en PostgreSQL para realizar todos los análisis. Los datos corresponden a un conjunto de series de 47 observatorios meteorológicos en Andalucía desde 1961 hasta 2017.

- MOTEDAS (1951-2010).

La base de datos MOTEDAS (Gonzalez-Hidalgo et al., 2015) es un conjunto de datos en rejilla de temperaturas mensuales para toda la península con una resolución de 0.1º (aprox. 10 km) que procede de una selección de 1358 observatorios de la AEMET que han sido sometidos a diferentes procesos de detección y eliminación de datos anómalos e inhomogeneidades. Son por tanto series homogéneas que garantizan que los errores achacables a cambios en los procesos de registro de los datos –cambios de ubicación, cambios de sensores, cambios en el entorno, etc...) han sido corregidos. Estos datos cubren el periodo 1951-2010.

## Selección de estaciones y series

Se han seleccionado cuatro estaciones con series largas (con el mínimo número de lagunas) y representativas de diferentes características climáticas en la región que garantizaban la mejor calidad para el estudio. Para ello, en primer lugar se ha realizado un análisis de las series para hacer la selección sobre aquellas que presentan un menor número de lagunas. Se ha elegido el periodo 1971-2017 por ser aquel en el que obtenían las series con menor porcentaje de lagunas (figura 1).



```
select * from test.mensuales_completas_tmax_lagunas_7117
where porcentaje_lagunas <=10
order by porcentaje_lagunas asc;
```

	codigo character varying(254)	nombre character varying(254)	fecha_inicio date	fecha_fin date	n_meses bigint	porcentaje_lagunas numeric
1	5796	Moron de la Frontera, Base Aerea	1971-01-31	2017-12-	526	7
2	6155A	Málaga, Aeropuerto	1971-01-31	2017-12-	521	8
3	5783	Sevilla, Aeropuerto	1971-01-31	2017-11-	518	8
4	5514	Granada, Base Aérea	1971-01-31	2017-12-	518	8
5	5960	Jerez de la Frontera, Aeropuerto	1971-02-28	2017-12-	516	9
6	6258	Lanjarón	1971-01-31	2017-12-	509	10
7	5530E	Granada, Aeropuerto	1972-08-01	2017-12-	507	10

Figura 1. Estaciones con un porcentaje menor o igual al 10% en el periodo 1971-2017.

Del grupo de estaciones que se muestran en la figura 1, se han seleccionado las de Granada Base Aérea (5514), Málaga Aeropuerto (6155A), Sevilla Aeropuerto (5783), Jerez de la Frontera Aeropuerto (5960), con el fin de disponer de series representativas de diferentes regímenes térmicos de la región.

El proceso de relleno de lagunas ha seguido dos pasos. El primer paso ha consistido en el cálculo de los valores mensuales agregados a partir de los valores diarios en aquellos meses en los que el número de datos diarios disponible es superior a 25. Una vez rellenas las lagunas

mediante este procedimiento, el porcentaje de lagunas en las estaciones seleccionadas se redujo sensiblemente (figura 2).

	<b>codigo</b> <b>character varying(254)</b>	<b>fecha_ini</b> <b>date</b>	<b>fecha_fin</b> <b>date</b>	<b>n_meses</b> <b>bigint</b>	<b>porcentaje_lagunas</b> <b>numeric</b>	<b>numero_lagunas</b> <b>bigint</b>
<b>1</b>	5783	1971-01	2017-11	561	1	3
<b>2</b>	5960	1971-01	2017-12	563	0	1
<b>3</b>	6155A	1971-01	2017-12	562	0	2
<b>4</b>	5514	1971-01	2017-12	556	1	8

Figura 2. Número y porcentaje de lagunas en las estaciones seleccionadas tras el primer relleno.

Las lagunas restantes (14 valores mensuales en total) fueron rellenadas mediante un análisis

<b>Estación</b>	<b>Nombre</b>	<b>Estación mejor correlacionada</b>	<b>Correlación de pearson</b>	<b>R2</b>	<b>Ecuación</b>
<b>5514</b>	<b>Granada Base Aérea</b>	<b>5530E</b>	<b>0,997</b>	<b>0,995</b>	<b>y = 0,9911x - 0,7028</b>
<b>6155A</b>	<b>Málaga Aeropuerto</b>	<b>6171</b>	<b>0,996</b>	<b>0,993</b>	<b>y = 1,0335x + 0,0948</b>
<b>5960</b>	<b>Jerez Aeropuerto</b>	<b>5973</b>	<b>0,982</b>	<b>0,965</b>	<b>y = 1,3927x - 5,8624</b>
<b>5783</b>	<b>Sevilla Aeropuerto</b>	<b>5796</b>	<b>0,998</b>	<b>0,995</b>	<b>y = 0,9923x - 0,5848</b>

de regresión lineal simple con la estación mejor correlacionada del entorno para cada una de las estaciones (tabla 1):

Tabla 1. Parámetros de la relación entre las estaciones y la mejor correlacionada del entorno.

En cuanto a las series procedentes de los conjuntos de datos en rejilla (MOTEDAS), se obtuvieron extrayendo los valores en las celdillas correspondientes a cada una de las estaciones seleccionadas.

### *Análisis de los datos originales y análisis de homogeneidad de las series de la AEMET*

#### *Datos diarios*

Los datos diarios a partir de los cuales se construyeron las series mensuales y la posterior espacialización y que son objeto de este control de calidad proceden de la base de datos CLIMA, del Subsistema de Información de Climatología Ambiental de la Consejería de Medio Ambiente y Ordenación del Territorio de la Junta de Andalucía.

Los datos proceden de un total de 218 estaciones que, además de la cobertura de estaciones de la AEMET, incluyen redes como la de agricultura, medio ambiente, SIVA y otras.

La metodología de relleno de lagunas que ha sido utilizada por las personas encargadas del estudio se basa en un proceso iterativo, en el que para cada estación es calculada la regresión lineal con la estación más cercana, sin ningún tipo de barrera o estrato, solo con el criterio de cercanía. Obtienen así una regresión lineal por cada mes del periodo completo 1961-2017, y se rellenan las lagunas que faltan según la expresión obtenida. Si son días faltantes en las dos estaciones, entonces se recurre a la siguiente estación más cercana y vuelve a realizarse el mismo proceso.

En un primer análisis de los datos diarios, y comparándolos con los datos de la AEMET, encontramos que el porcentaje de datos con desviaciones superiores a  $\pm 0.2^\circ$  es del 26%. De la

totalidad de los registros (unos 70000), un 24% son interpolados (cuando no deberían serlo, ya que existe el dato en la base de datos de CLIMA). De estos registros interpolados, un 21% tienen una desviación superior a  $+2^{\circ}\text{C}$ . Por otra parte, de los registros no interpolados (un 76%), un 9% presentan diferencias superiores a  $+0.2^{\circ}\text{C}$ .

También se han encontrado algunos valores anómalos, que son en algunos casos claramente erróneos (figura 11), aunque sólo afectan a algunas estaciones, entre las que no se encuentran las cuatro seleccionadas para el análisis.

Salida de datos							
Comentar							
Mensajes							
Historial							
	objectid integer	codigo character varying(10)	agno integer	mes integer	dia integer	t_max numeric	origen character varying(10)
1	899052	5071E	1971	7	2	0.0000000000000000	I00
2	899051	5071E	1971	7	1	0.0000000000000000	I00
3	4335667	6364	1975	7	21	0.0000000000000000	I00
4	4335664	6364	1975	7	18	0.0000000000000000	I00
5	4067201	6308	1981	7	10	0.0000000000000000	I00
6	4335666	6364	1975	7	20	0.0000000000000000	I00
7	4335665	6364	1975	7	19	0.0000000000000000	I00
8	2816986	5873A	1978	7	31	0.0000000000000000	I00
9	2815509	5873A	1974	7	15	0.0000000000000000	I00
10	2633604	5836A	1989	7	2	4.0000000000000000	I00
11	2796167	5870A	1978	7	31	7.7000000000000000	I02
12	838786	5044E	1977	7	1	10.0000000000000000	I00
13	919871	5085	1971	7	2	12.0000000000000000	I04

Figura 3. Consulta de valores de temperatura máxima diaria del mes de julio ordenados de menor a mayor.

#### Datos mensuales y anuales en las cuatro estaciones seleccionadas

En las figuras 4 a 11 se muestran los gráficos comparados de las series originales y las series obtenidas de la AEMET para la temperatura media de las máximas mensuales de junio y anual.

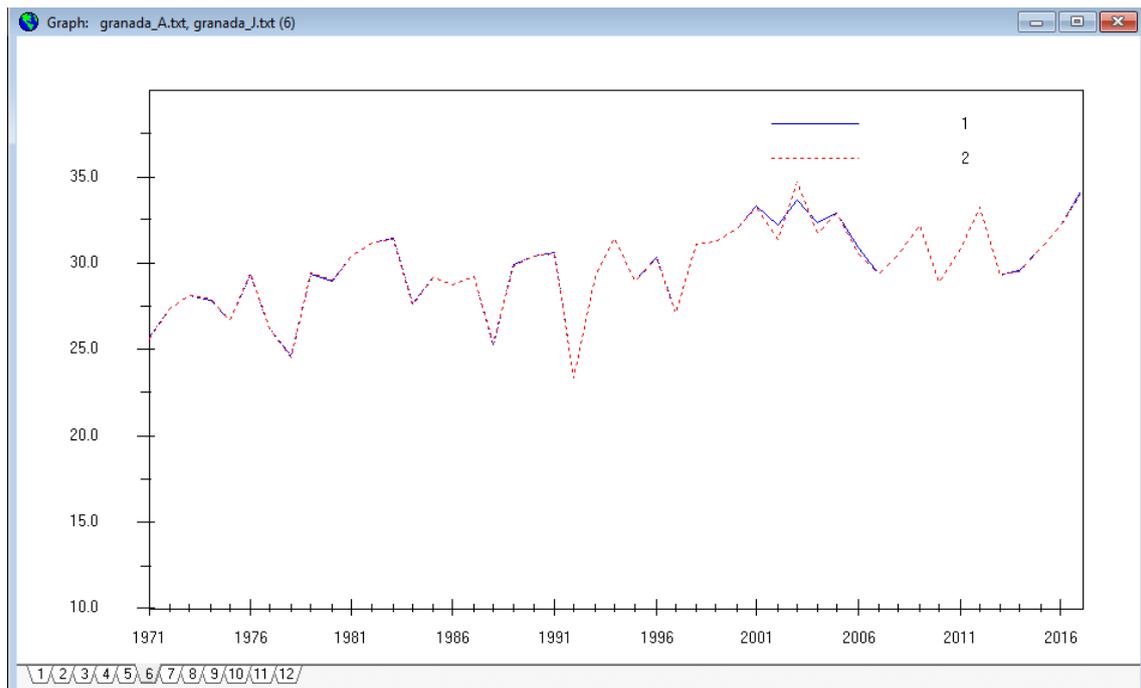


Figura 4. Series de temperatura media de las máximas mensuales de junio de AEMET (1) y CLIMA (2) en la estación de Granada Base Aérea (5514).

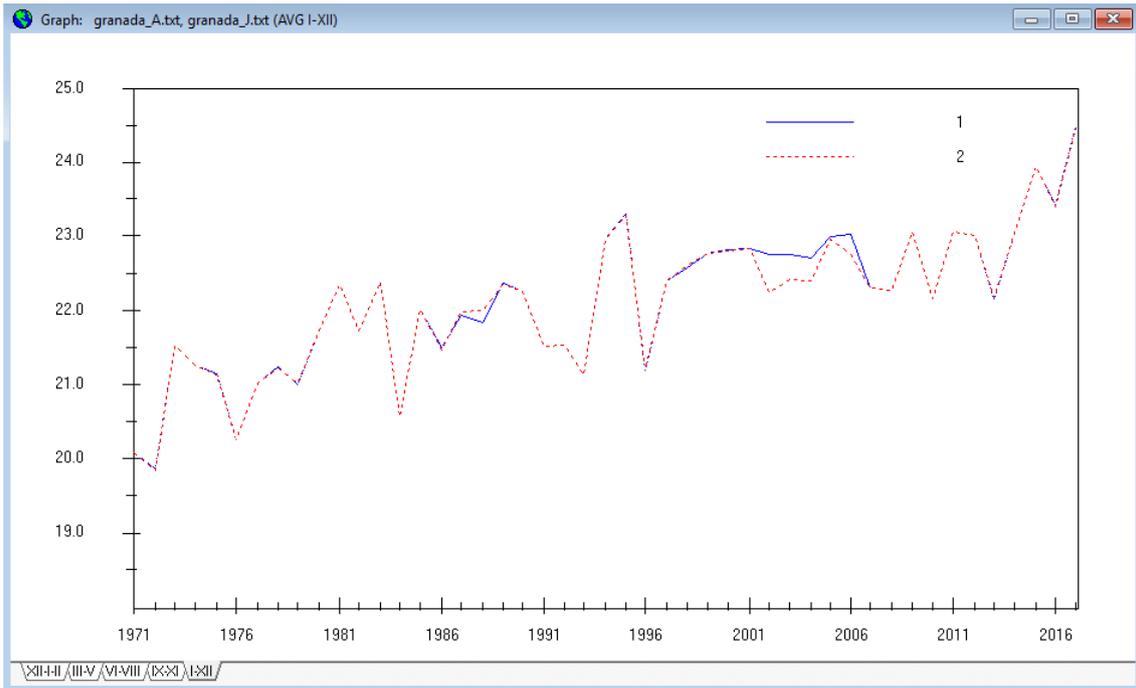


Figura 5. Series de temperatura media anual de las máximas mensuales de AEMET (1) y CLIMA (2) en la estación de Granada Base Aérea (5514).

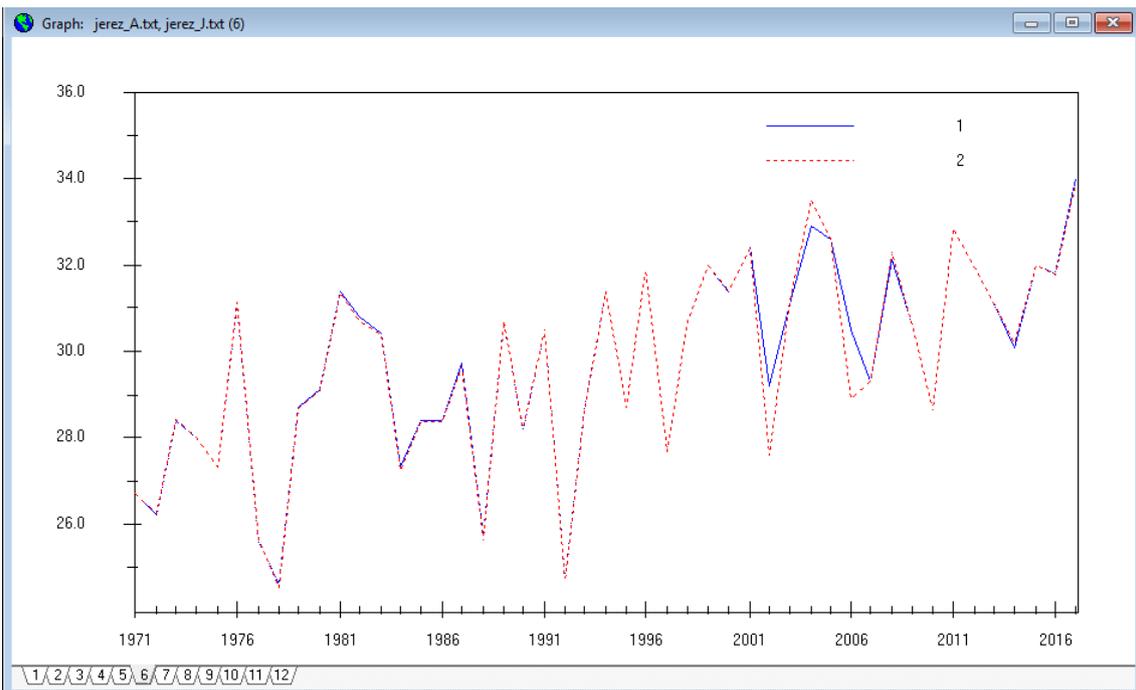


Figura 6. Series de temperatura media de las máximas mensuales de junio de AEMET (1) y CLIMA (2) en la estación de Jerez Aeropuerto (5690).

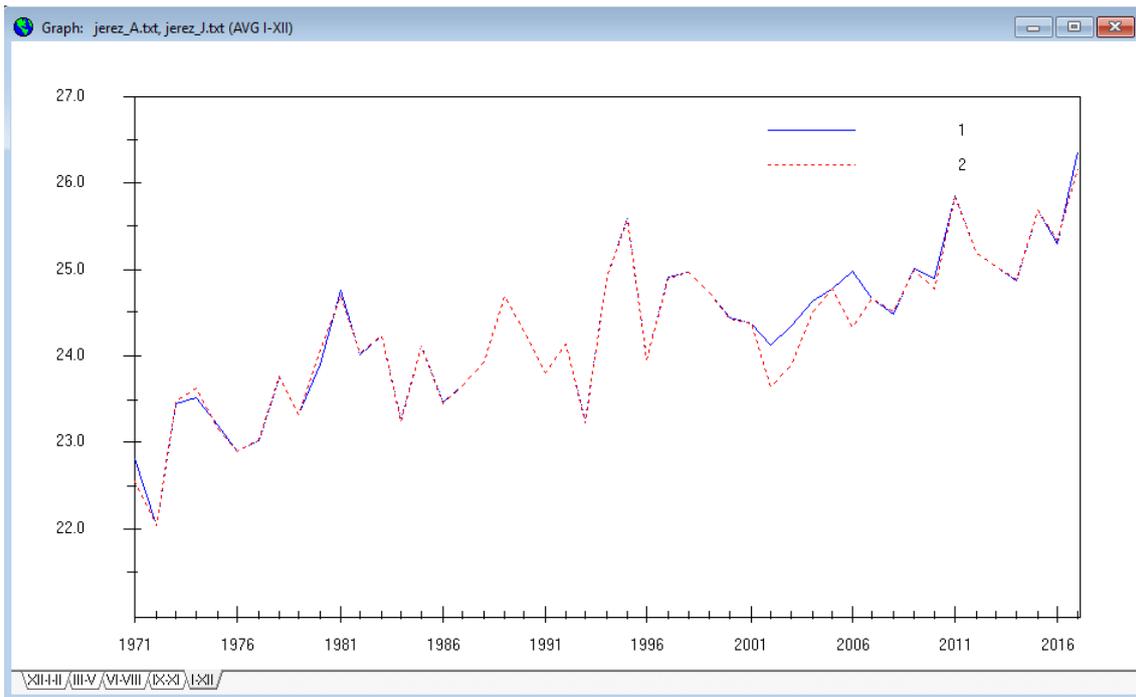


Figura 7. Series de temperatura media anual de las máximas mensuales de AEMET (1) y CLIMA (2) en la estación de Jerez Aeropuerto (5690).

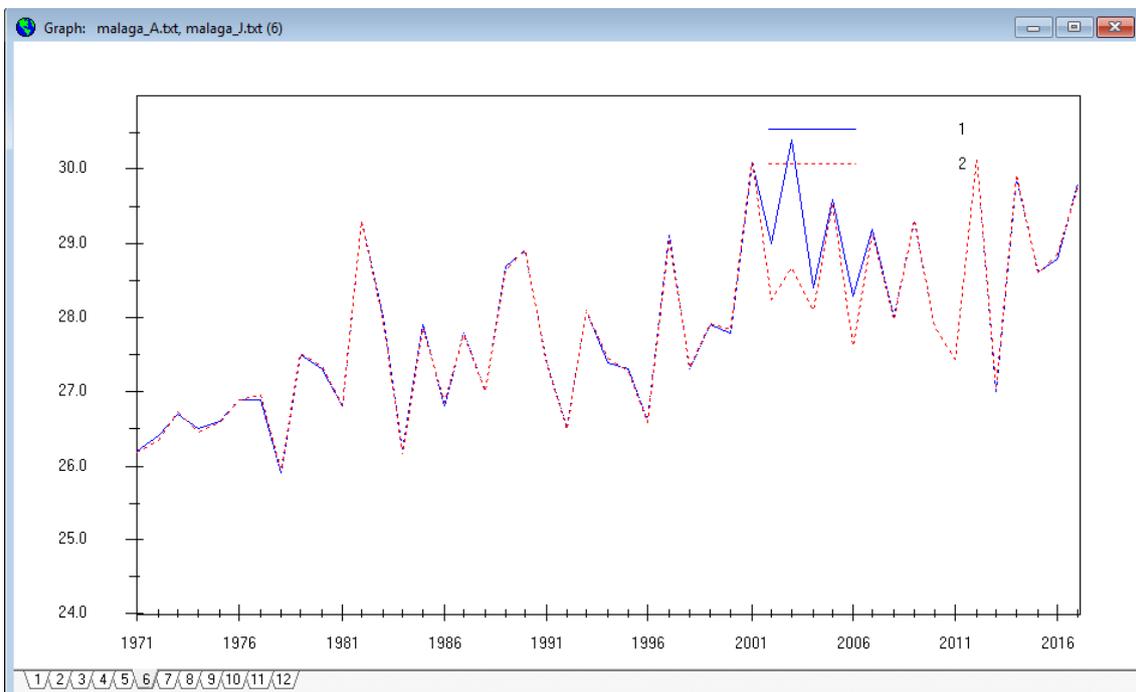


Figura 8. Series de temperatura media de las máximas mensuales de junio de AEMET (1) y CLIMA (2) en la estación de Málaga Aeropuerto (6155A).

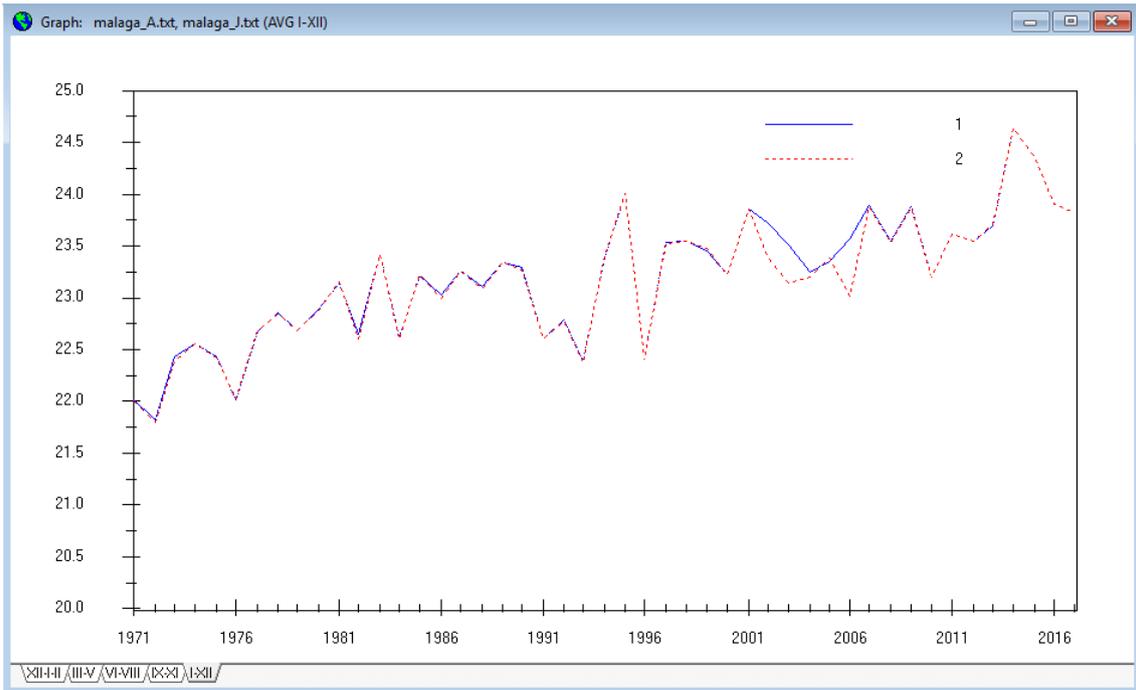


Figura 9. Series de temperatura media anual de las máximas mensuales de AEMET (1) y CLIMA (2) en la estación de Málaga Aeropuerto (6155A).

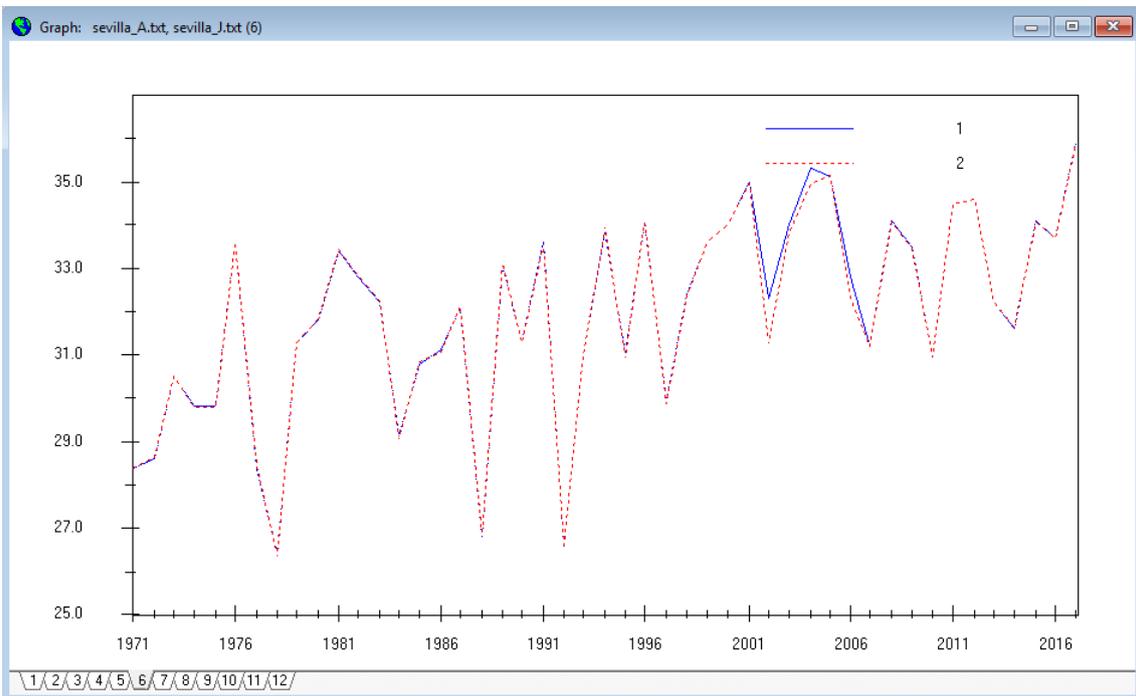


Figura 10. Series de temperatura media de las máximas mensuales de junio de AEMET (1) y CLIMA (2) en la estación de Sevilla Aeropuerto (5783).

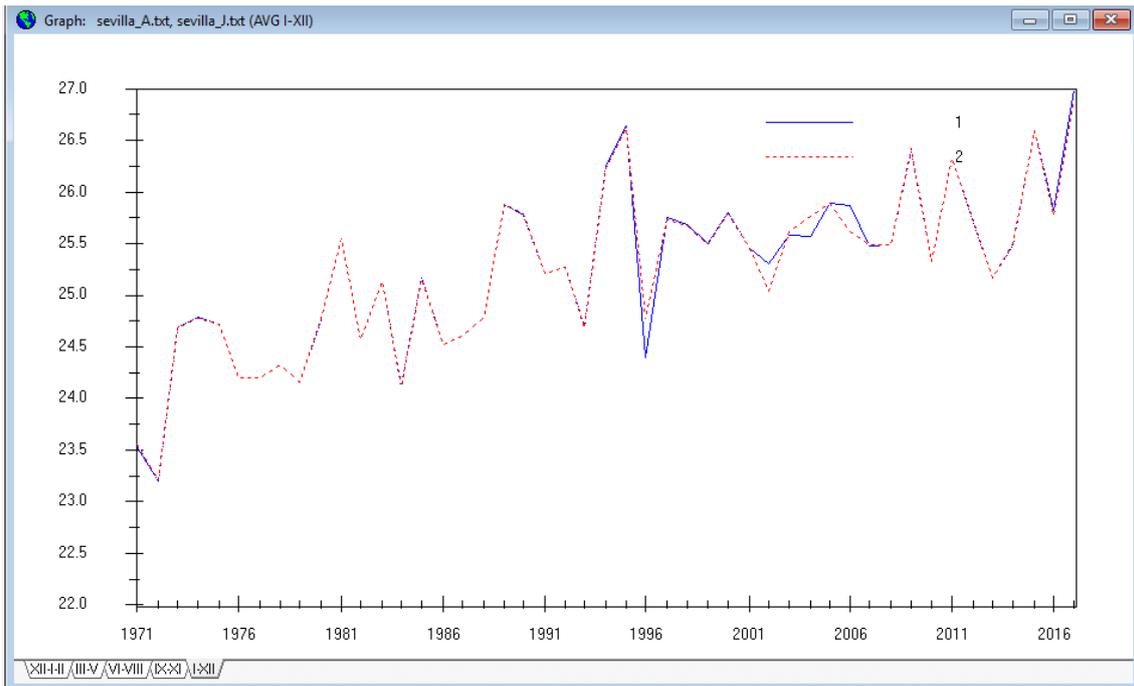


Figura 11. Series de temperatura media anual de las máximas mensuales de AEMET (1) y CLIMA (2) en la estación de Sevilla Aeropuerto (5783).

Como se puede deducir de los gráficos, existe una gran similitud entre las series que sólo presentan pequeñas diferencias, generalmente en la década 2000-2010, debido probablemente a que en esos años existen muchos valores interpolados en la serie de CLIMA.

Este hecho, unido a los problemas relacionados con los datos anómalos encontrados en los datos diarios, nos ha decidido a **optar por la utilización de las series extraídas de la AEMET para los posteriores análisis**, ya que estas respetan en mayor medida los datos originales.

### *Análisis de homogeneidad*

Los análisis de homogeneidad se han realizado tomando como series de referencia para cada una de las estaciones los datos extraídos de la base de datos MOTEDAS para las celdillas correspondientes.

En los gráficos 12 a 15 se muestran medias anuales de la temperatura máxima de las series MOTEDAS y AEMET para las cuatro estaciones analizadas. En todos los gráficos, con la excepción de Málaga, se observa cómo se produce un cambio en la relación entre las dos series que suele situarse aproximadamente a mediados y finales de los años 90. Esto puede ser una señal de posibles inhomogeneidades hacia final del periodo, que se explicaría por la progresiva implantación de las redes de estaciones automáticas.

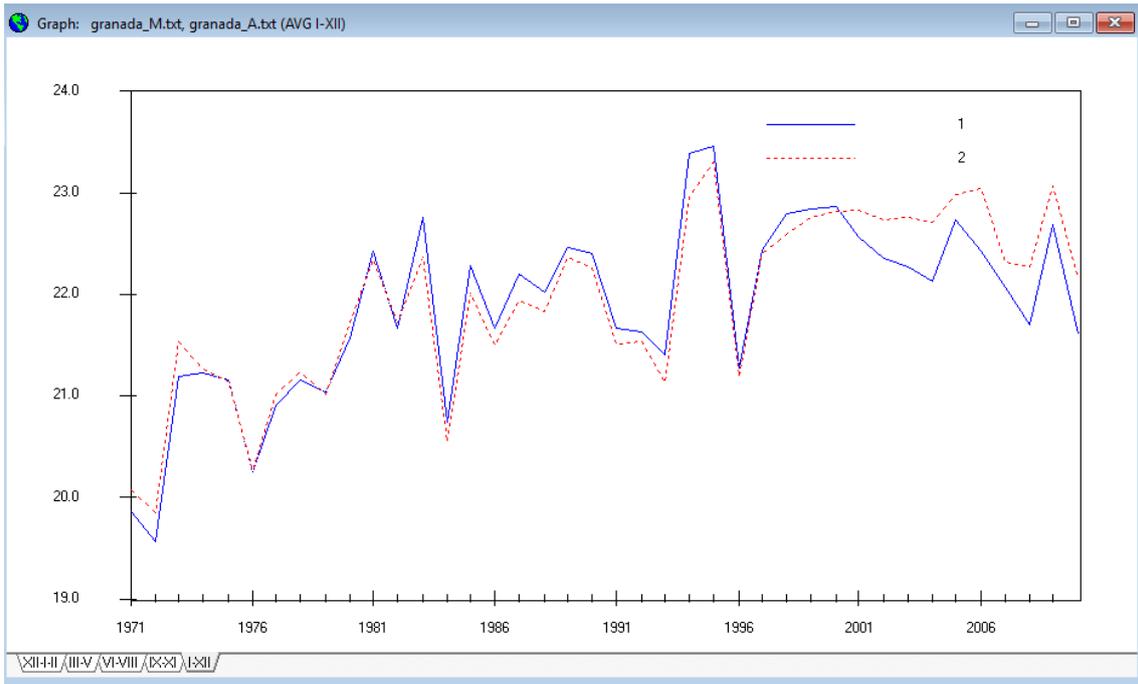


Figura 12. Gráfico de comparación de valores medios anuales de tmax en Granada (1-MOTEDAS; 2-AEMET).

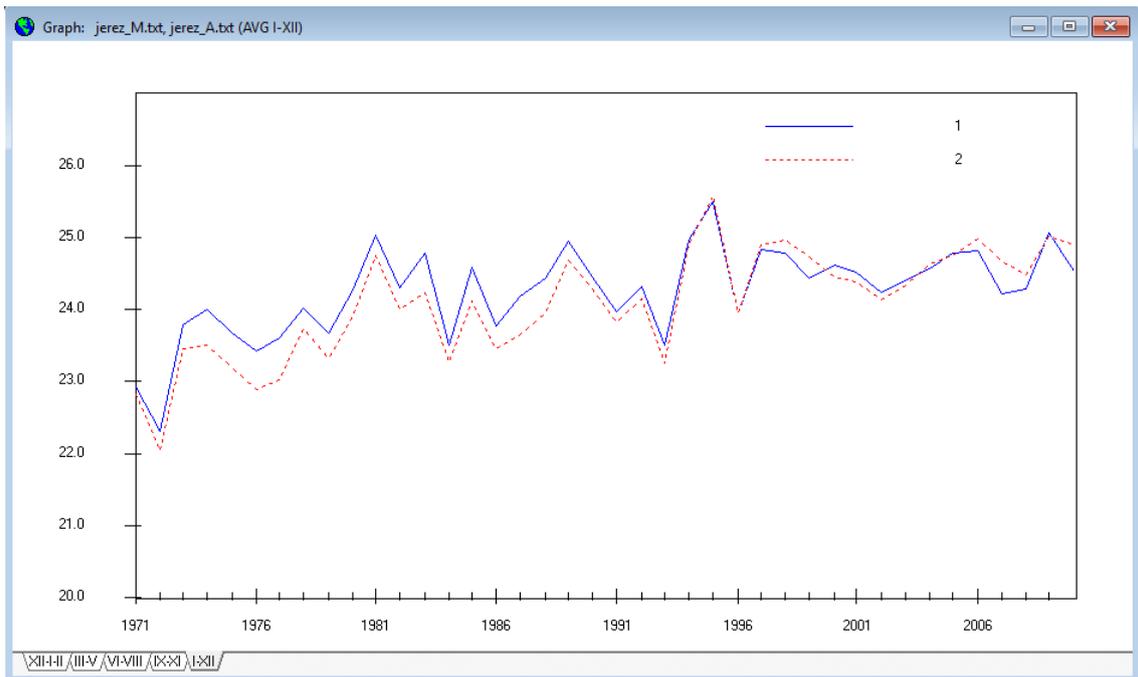


Figura 13. Gráfico de comparación de valores medios anuales de tmax en Jerez (1-MOTEDAS; 2-AEMET).

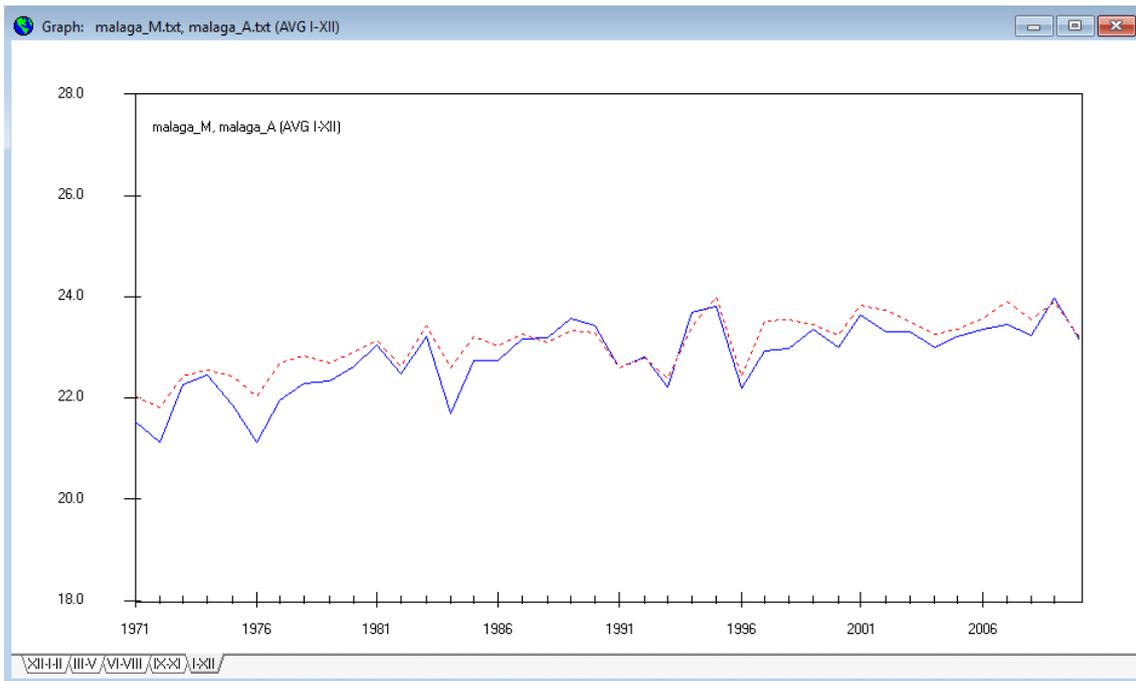


Figura 14. Gráfico de comparación de valores medios anuales de tmax en Málaga (1-MOTEDAS; 2-AEMET).

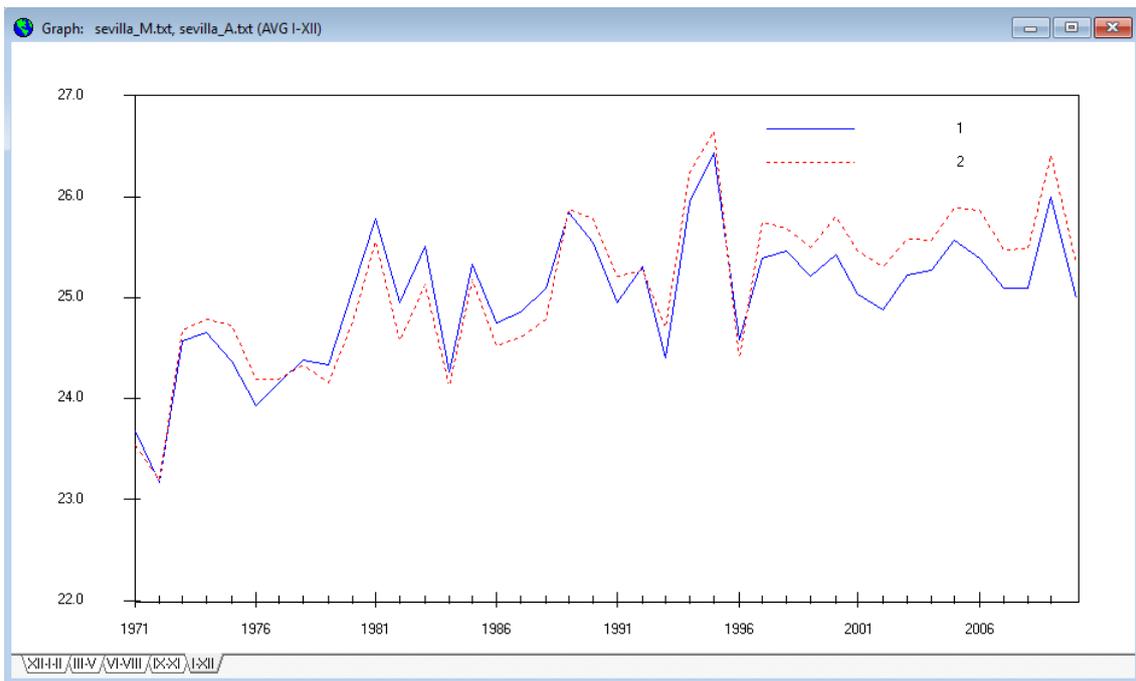


Figura 15. Gráfico de comparación de valores medios anuales de tmax en Sevilla (1-MOTEDAS; 2-AEMET).

Para comprobar numéricamente la existencia de inhomogeneidades, se ha realizado el test de Alexanderson (SHNT) de homogeneidad relativa (Shift versión II), tomando la serie MOTEDAS como serie de referencia. El software utilizado ha sido Anclim, desarrollado por Petr Stepanek (Stepanek, 2003). Los resultados anuales se muestran en las figuras 16 a 19, que aparecen a continuación.

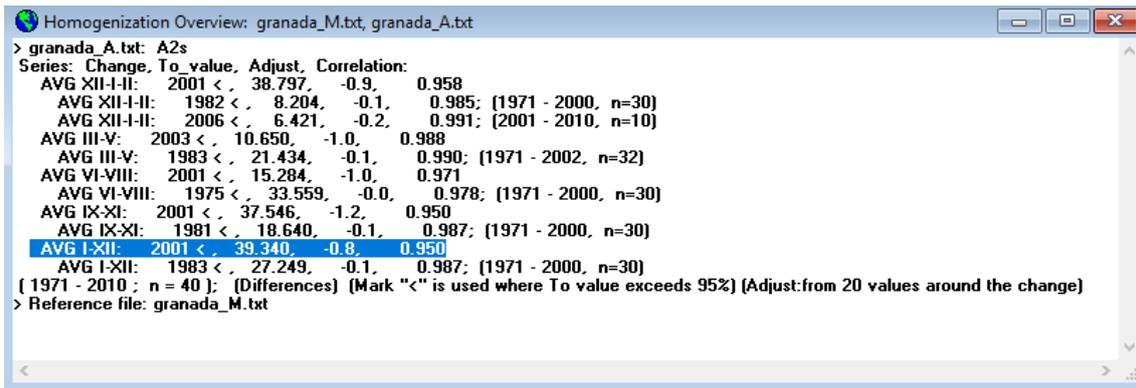


Figura 16. Resultados del test de homogeneidad para la serie anual de Granada Base Aérea (5514).

En la serie de Granada se detecta una posible homogeneidad (significación del 95%) a partir del año 2001, con un ajuste de  $-0.8$  °C. Este resultado confirma lo observado en el gráfico de la figura 12.

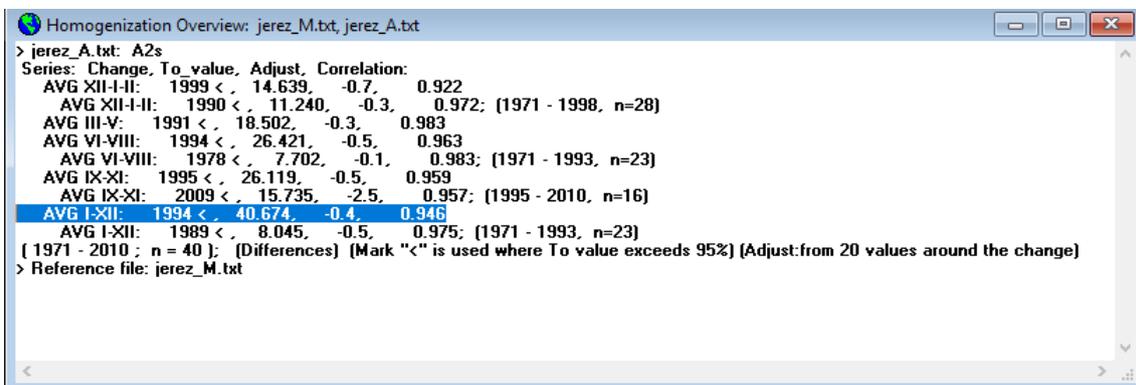


Figura 17. Resultados del test de homogeneidad para la serie anual de Jerez Aeropuerto (5960).

En la serie de Jerez se detecta una posible homogeneidad (significación del 95%) a partir del año 1994, con un ajuste de  $-0.4$  °C. Este resultado confirma lo observado en el gráfico de la figura 13.

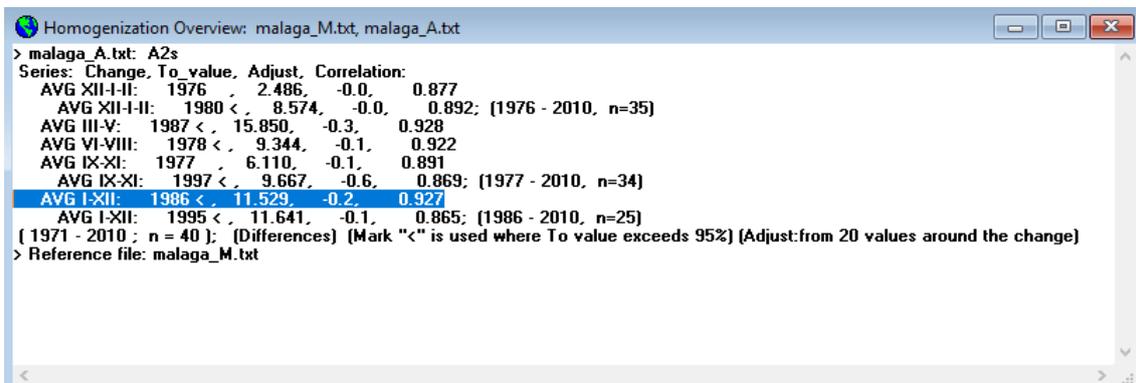


Figura 18. Resultados del test de homogeneidad para la serie anual de Málaga Aeropuerto (6155A).

En la serie de Málaga se detecta una posible homogeneidad (significación del 95%) a partir del año 1986, con un ajuste de  $-0.2$  °C. Este resultado viene a confirmar lo observado en el gráfico de la figura 14. También se obtiene un cambio en la serie a partir de 1995, hecho que se observa igualmente en el gráfico, que precisamente muestra un comportamiento anómalo en relación con el conjunto de la serie entre los años 1986 y 1995.

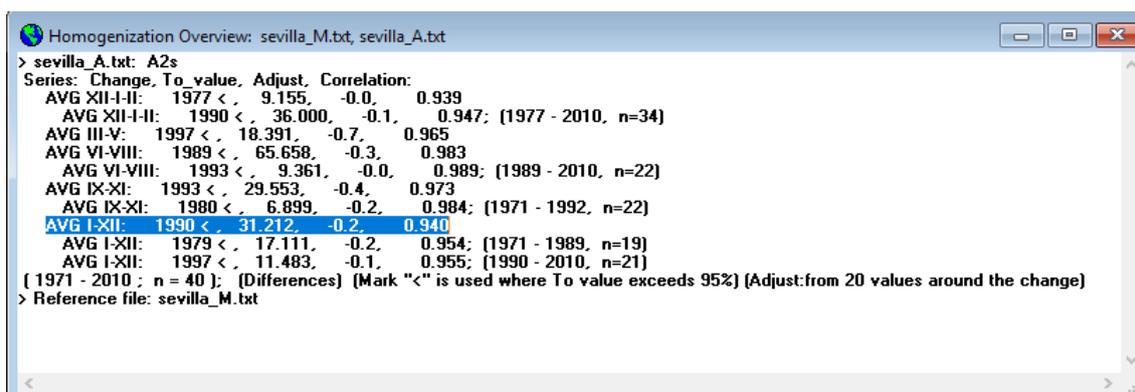


Figura 19. Resultados del test de homogeneidad para la serie anual de Sevilla Aeropuerto (5783).

En la serie de Sevilla se detecta una posible homogeneidad (significación del 95%) a partir del año 1990, con un ajuste de  $-0.2$  °C. Este resultado confirma lo observado en el gráfico de la figura 15.

La observación de los gráficos, así como las posibles inhomogeneidades detectadas por el test, indicarían que **la tendencia positiva de la serie es generalmente mayor que la tendencia real cuando se calcula sobre las series no homogeneizadas**. Esto es así porque, en la mayoría de los casos, las series no homogeneizadas de la AEMET presentan valores más altos que la serie MOTEDAS a partir del salto de inhomogeneidad. Por ello, en el siguiente epígrafe, se abordará el cálculo y comparación de las tendencias en las series de la AEMET originales y en las mismas series tras su homogeneización, así como en las series de referencia (MOTEDAS).

## Análisis de tendencias

Para el cálculo de las tendencias se ha utilizado el método de ajuste por mínimos cuadrados. Se ha calculado la tendencia para dos periodos diferentes: El periodo global (1971-2010 para MOTEDAS y 1971-2017 para AEMET) y dos sub-periodos (1971-2000 para MOTEDAS y AEMET y 2001-2010 para MOTEDAS y 2001-2017 para AEMET). Se mostrarán a continuación los resultados obtenidos para el mes de junio, ya que es uno de los meses en los que se obtienen los mayores aumentos de la temperatura, tanto en los datos objeto de este control como en general, como se refleja en otros estudios (Abaurrea et al, 2006; Barbancho et al., 2012; Del Río et al., 2012; Mestre et al., 2015; González Hidalgo et al., 2016).

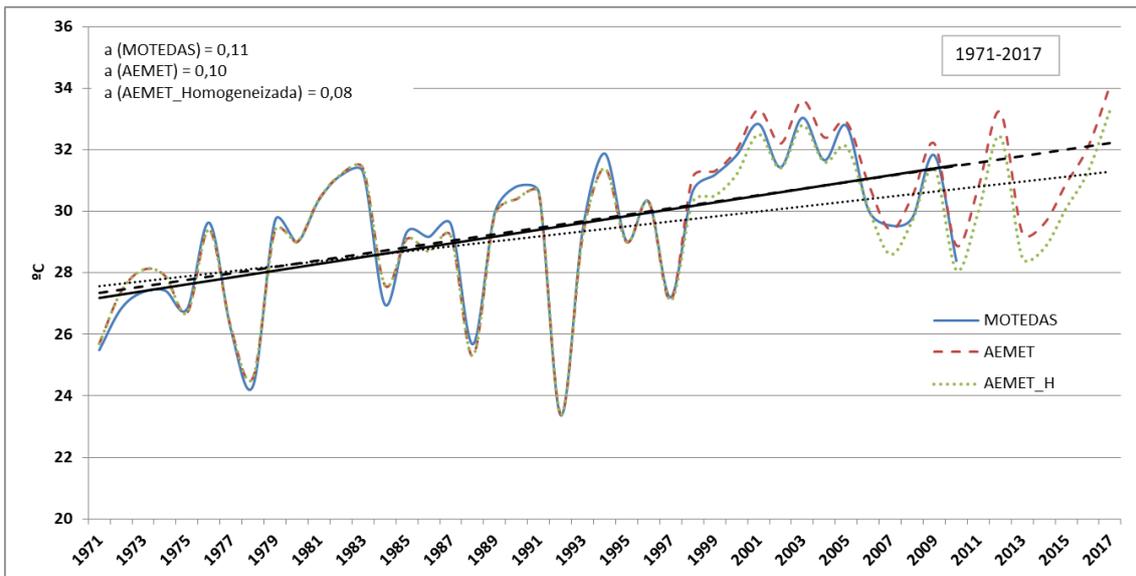


Figura 20. Tendencias de la temperatura media mensual de las máximas en Granada Base Aérea en el mes de junio en las tres series analizadas en el periodo 1971-2017. (a= cambio anual).

Si comparamos el gráfico anterior con el gráfico proporcionado por la Agencia (Fig.21), observamos como la comparación de las medias de dos periodos en una serie con tendencia no puede ser considerada una herramienta adecuada ya que, al ser la media un valor aritmético abstracto dependiente de la distribución de los valores y del rango de los mismos, va a producir dos medidas sesgadas por la posición del periodo en la serie, más baja en aquella que contiene los valores iniciales de la señal y más alta en el periodo final de la misma. Adicionalmente la ruptura en el año 2000 implica que el periodo de comparación tiene solamente 17 años, lo cual no cumpliría con la norma de periodos mínimos de 30 años que puedan ser representativos estadísticamente.

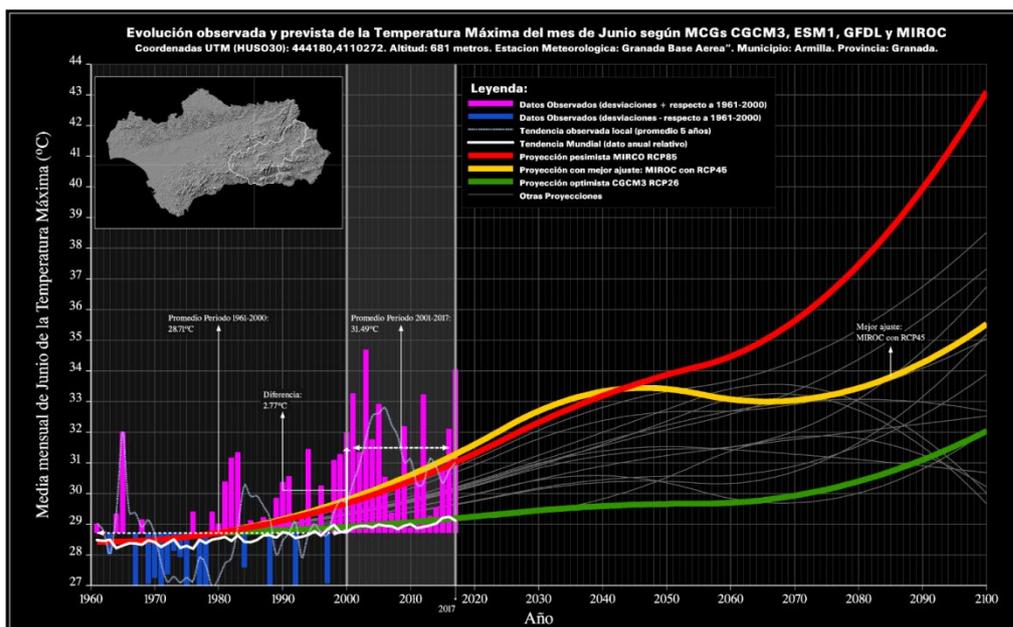


Figura 21. Gráfico en Granada Base Aérea en el mes de junio proporcionado por la Agencia.

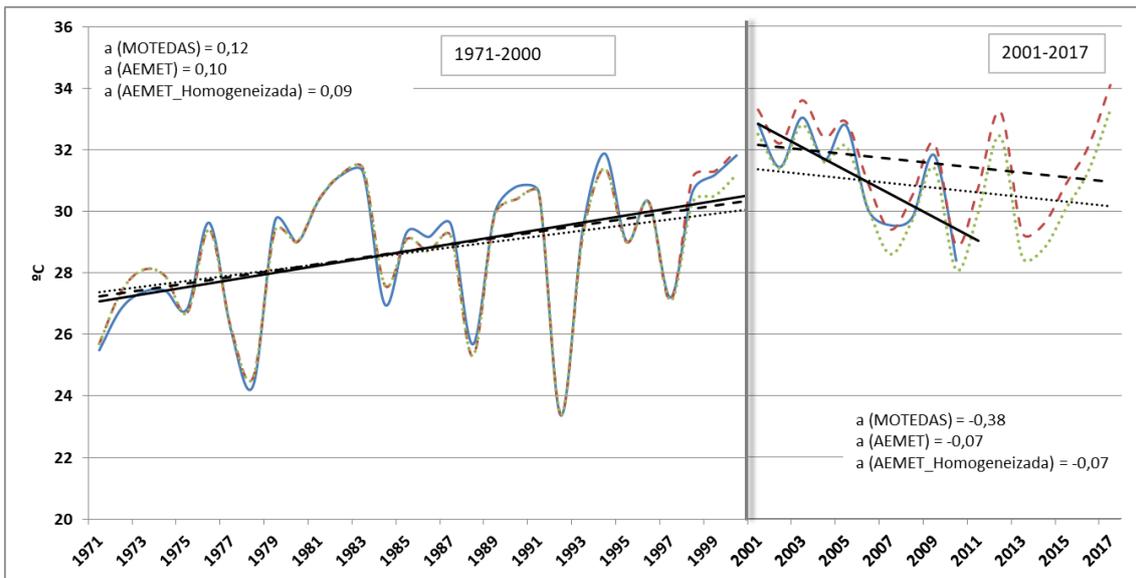


Figura 22. Tendencias de la temperatura media mensual de las máximas en Granada Base Aérea en el mes de junio en las tres series analizadas en el periodo 1971-2017 y 2001-2017. (a= cambio anual).

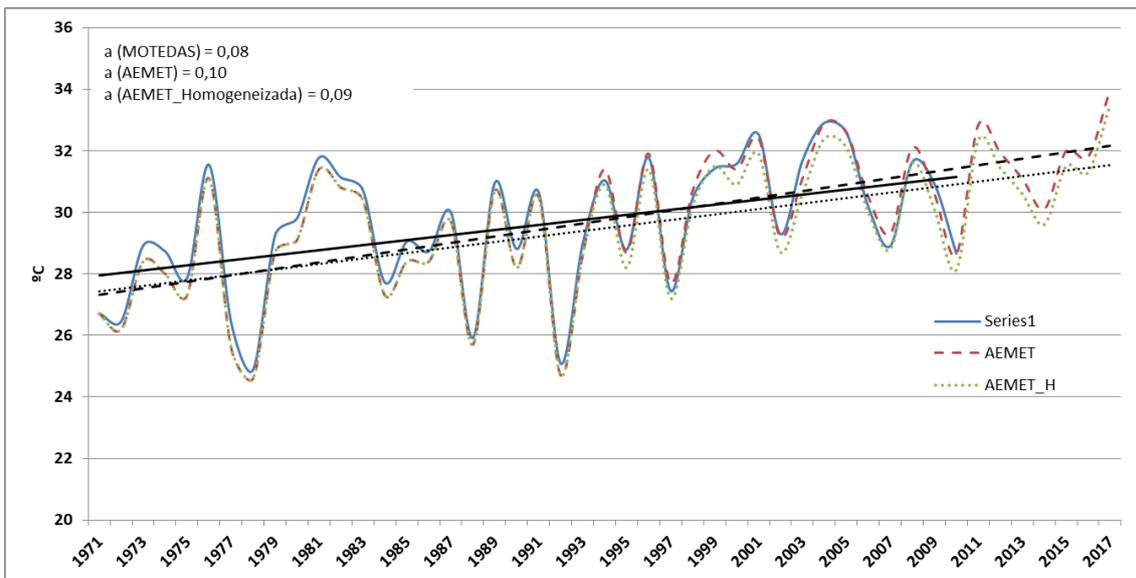


Figura 23. Tendencias de la temperatura media mensual de las máximas en Jerez Aeropuerto en el mes de junio en las tres series analizadas en el periodo 1971-2017. (a= cambio anual).

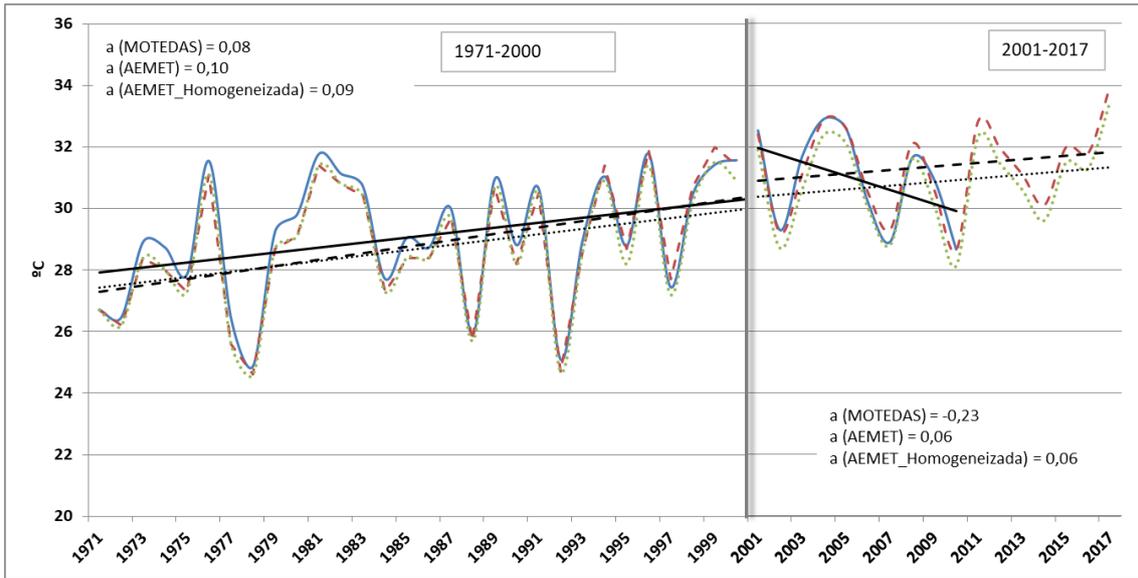


Figura 24. Tendencias de la temperatura media mensual de las máximas en Jerez Aeropuerto en el mes de junio en las tres series analizadas en el periodo 1971-2017 y 2001-2017. (a= cambio anual).

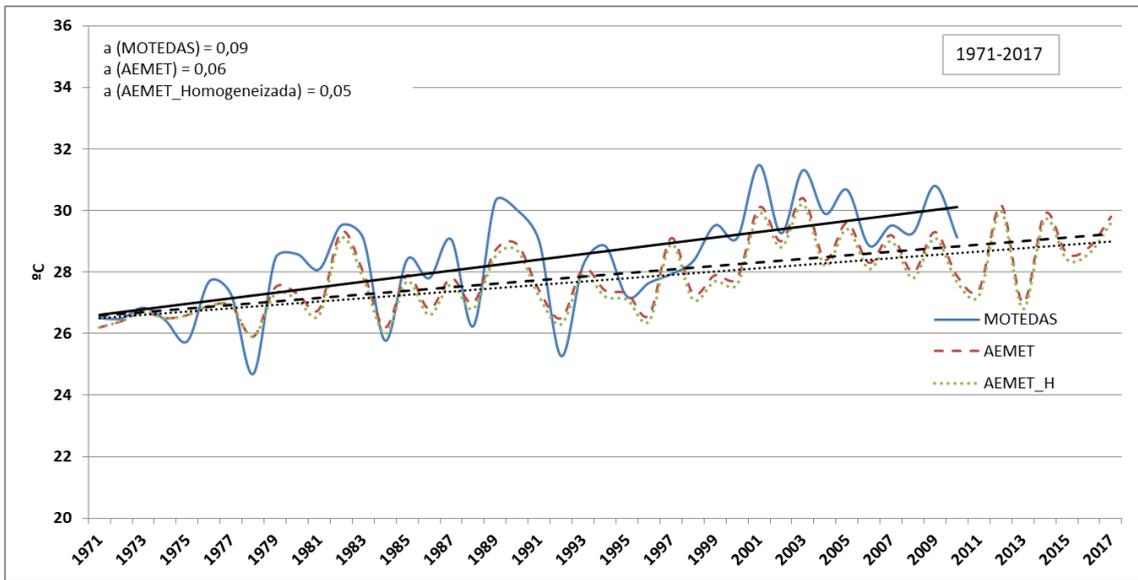


Figura 25. Tendencias de la temperatura media mensual de las máximas en Málaga Aeropuerto en el mes de junio en las tres series analizadas en el periodo 1971-2017. (a= cambio anual).

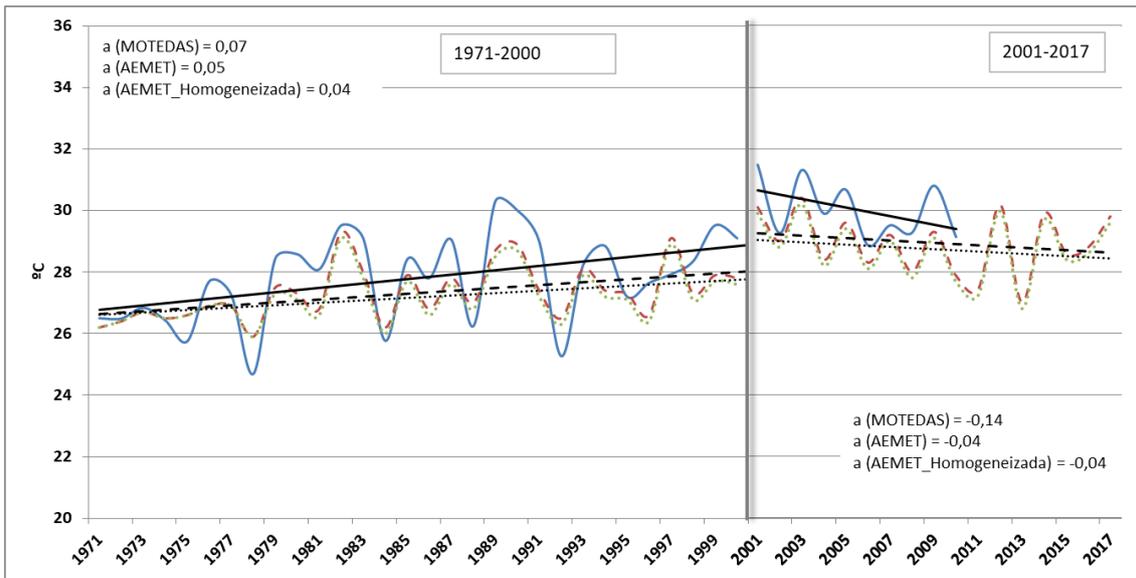


Figura 26. Tendencias de la temperatura media mensual de las máximas en Málaga Aeropuerto en el mes de junio en las tres series analizadas en el periodo 1971-2017 y 2001-2017. (a= cambio anual).

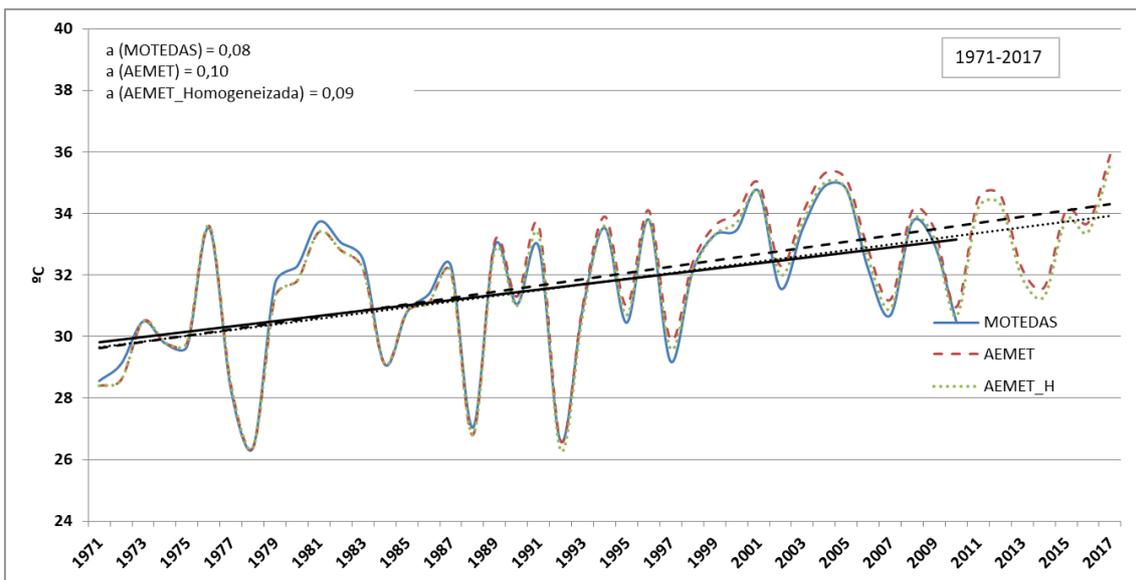


Figura 27. Tendencias de la temperatura media mensual de las máximas en Sevilla Aeropuerto en el mes de junio en las tres series analizadas en el periodo 1971-2017. (a= cambio anual).

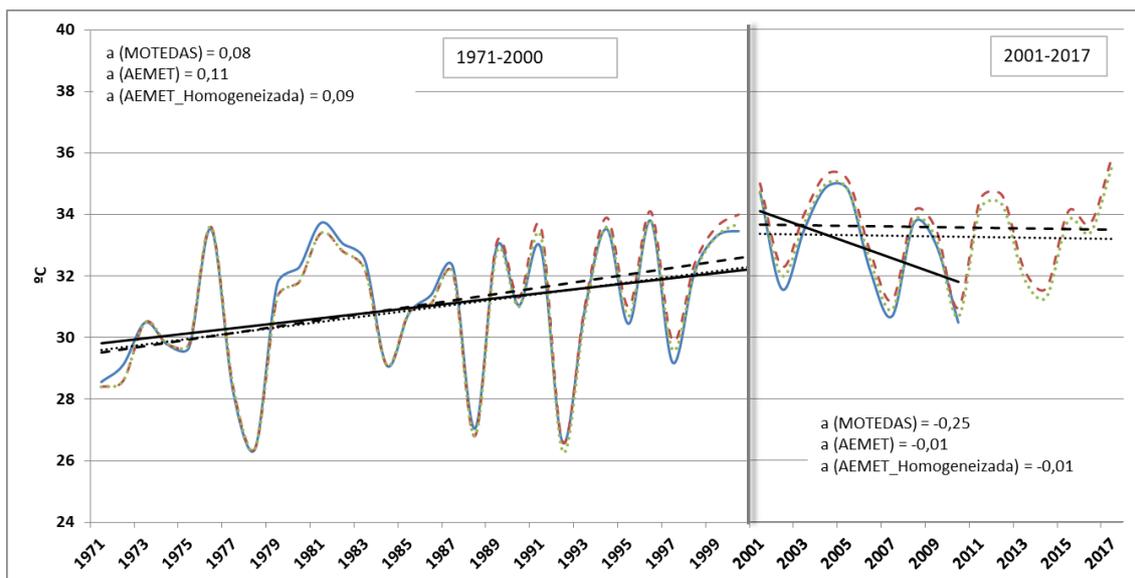


Figura 28. Tendencias de la temperatura media mensual de las máximas en Sevilla Aeropuerto en el mes de junio en las tres series analizadas en el periodo 1971-2017 y 2001-2017. (a= cambio anual).

En el caso de la serie de MOTEDAS, para el periodo global 1971-2017, en el conjunto de las cuatro estaciones, se observan tendencias que van de 0.8º a 1.1º/Década (en adelante, /D.), mientras que las tendencias de las series homogeneizadas de la AEMET oscilan entre 0.5º y 0.9º/D.

Estos resultados concuerdan en general con los encontrados en diferentes estudios sobre la evolución de las temperaturas en España durante las últimas décadas, especialmente en las épocas primaveral y estival: Sigró et al. (2008) indican una tendencia media en el territorio peninsular de 0.84º/D (media junio-julio) en el periodo 1973-2005, Ceballos et al. (2007), de 1.07º/D (junio) en ciudades de Castilla y León en el periodo 1961-2006, o Ríos et al. (2012), que calculan un incremento medio de 0.4º/D en el mes de junio en la península ibérica.

Las tendencias del periodo 1971-2000 arrojan valores muy similares a los del periodo global, mientras que en los últimos periodos (2001-2010 y 2001-2017) no parecen existir tendencias significativas o son en general ligeramente negativas. Este último hecho ha sido también constatado por Ríos et al (2012), que sólo encuentran tendencias significativas (negativas) en los meses de enero y marzo durante el periodo 1994-2010.

## Conclusiones y propuestas

- El análisis de los datos indica una tendencia significativa al incremento de las temperaturas máximas durante el periodo 1971-2010, que está en consonancia con los resultados de otros trabajos consultados en la literatura científica sobre la evolución de las temperaturas en la península ibérica.
- Los periodos 2001-2010 parecen mostrar un comportamiento térmico más neutro, sin embargo, no resulta adecuado el cálculo de las tendencias en estos periodos tan cortos por no ser estadísticamente significativas.

- Consideramos que el hecho de comparar la media de estos últimos periodos con la media del periodo 1961-1990 o del periodo 1971-2000 produce un resultado distorsionado y alarmista de la evolución térmica de las primeras décadas del siglo XXI. Esto es debido a la propia existencia de la tendencia, ya que la media del periodo de referencia siempre será mucho menor que la del último periodo, incluso existiendo, como parece, una estabilización en la evolución térmica de los últimos años.

## Referencias

Abaurrea J., Álvarez E., Asín J., Cebrián A.C. y Centelles A. (2006): Evolución de las temperaturas máxima y mínima diaria en Zaragoza, Barcelona y Madrid en la segunda mitad del siglo XX. En: Cuadrat J., Saz M., Vicente Serrano S.M., Lanjeri S., De Luis M., González-Hidalgo J. (eds.) *Clima, Sociedad y Medio Ambiente*. V Congreso de la AEC. Asociación Española de Climatología, Zaragoza, pp 15–26. URL.: <http://hdl.handle.net/20.500.11765/8730>.

Ceballos A., Morán E, Quirós M. (2007). Evolución de las temperaturas y precipitaciones en las capitales de Castilla y León en el período 1961-2006. *Polígonos. Revista de Geografía*. 17: 59-81. ISSN: 1132-1202.

Gonzalez-Hidalgo, J. C., Peña-Angulo, D., Brunetti, M. y Cortesi, N. (2015): MOTEDAS: a new monthly temperature database for mainland Spain and the trend in temperature (1951–2010). *International Journal of Climatology*, 35(15), 4444-4463. DOI.: <https://doi.org/10.1002/joc.4298>.

Gonzalez-Hidalgo, J. C., Peña-Angulo, D., Brunetti, M. y Cortesi, N. (2016): Recent trend in temperature evolution in Spanish mainland (1951–2010): from warming to hiatus. *International Journal of Climatology*, 36(6), 2405-2416. DOI.: <https://doi.org/10.1002/joc.4519>.

Del Río S., Herrero L., Pinto-Gomes C. y Penas A. (2011). Spatial analysis of mean temperature trends in Spain over the period 1961–2006. *Glob. Planet. Change*. 78: 65-75. DOI: 10.1016/j.gloplancha.2011.05.012.

Del Río, S., Cano-Ortiz, A., Herrero, L. y Penas, A. (2012). Recent trends in mean maximum and minimum air temperatures over Spain (1961–2006). *Theoretical and Applied Climatology*, 109(3-4), 605-626. DOI.: <https://doi.org/10.1007/s00704-012-0593-2>.

Mestre, I.; Casado, M.J. y Rodríguez, E. (2015). Tendencias observadas y proyecciones de cambio climático sobre España. Capítulo 2. En Herrero A. y Zavala M.A. (eds), *Los Bosques y la Biodiversidad frente al Cambio Climático: Impactos, Vulnerabilidad y Adaptación en España*. MAGRAMA, pp. 87 -98.

Ríos D., Peñas D. y del Río S. (2012). Comparative analysis of mean temperature trends in continental Spain over the period 1961–2010. *Int. J. Geobot. Res.* 2: 41–55.

Sigró J, Brunet M, Aguilar E (2008) Evolución de la temperatura máxima estival en España. Parte I: Tendencias y patrones espaciales. En: Sigró J, Brunet M, Aguilar E (eds) VI Congreso

Internacional AEC. Cambio Climático regional y sus impactos. vol Serie A. Artyplan, Tarragona, pp 307–318.

Stepanek, P. (2003): AnClim - software for time series analysis. Dept. of Geography, Fac. of Natural Sciences, MU, Brno. 1.47 MB.

ANEXO I. Script para la descarga de datos mensuales de cualquier variable de la API AEMET OpenData.

```
1 #Autor-Jose I. Alvarez Francoso
2 import httplib
3 import json
4 import csv
5 import time
6 import io
7 import sys
8 def carga_aemet(inicio,fin,variable):
9     conn = httplib.HTTPSConnection("opendata.aemet.es")
10    headers = {
11        'cache-control': "no-cache"
12    }
13    with io.open('C:\Users\Jose\Documents\ELCC5\completas\estaciones_andal.json', encoding=
14        'ISO 8859-1') as data_file:
15        data = json.load(data_file)
16        archivo = open('C:\Users\Jose\Documents\ELCC5\completas\datos_mensuales_' + variable +
17            '_' + str(inicio) + '-' + str(fin) + '.csv','wb')
18        fin2= int(fin) + 1
19        inicio2 = int(inicio)
20        for agno in range(inicio2, fin2):
21            print ('Procesando agno : ' + str(agno) + "...")
22            i=-1
23            for estacion in data:
24                i=i+1
25                codigo = data[i]["indicativo"]
26                conn.request("GET",
27                    "/opendata/api/valores/climatologicos/mensualesanuales/datos/aniocini/"+str(agno)+
28                    "/aniofin/"+str(agno)+"/estacion/"+codigo+"?AQUI_VA_LA_API_KEY", headers=headers)
29                res = conn.getresponse()
30                if res:
31                    # print('Estacion codigo : ' + codigo )
32                    data1 = res.read()
33                    data2 = data1.decode("ISO 8859-1",'ignore')
34                    data3= json.loads(data2)
35                    res_state = data3['estado']
36                    if res_state == 200:
37                        conn.request("GET", data3['datos'], headers=headers)
38                        res= conn.getresponse()
39                        datos2 = res.read().decode('utf-8','ignore')
40                        # Validar json. Aunque se espere 1 minuto, al hacer un cierto numero de peticiones, la
41                        # API devuelve un 429 y el json es invalido, por lo que se vuelve a hacer la peticion
42                        try:
43                            json.loads(datos2)
44                            datos3= json.loads(datos2)
45                            wr = csv.writer(archivo, delimiter=';')
46                            for item in datos3:
47                                if variable in item:
48                                    print ('Processing station ' + codigo + ' and year ' + str(agno))
49                                    wr.writerow([[item['indicativo']]+[item['fecha']]+[item[variable]]])
50                        except ValueError as error:
51                            print ('Processing year ' + str(agno))
52                            print("invalid json: %s" % error)
53                            print("I keep trying, hang on please...")
54                            time.sleep(60)
55                            i=i-1
56                    if res_state == 429:
57                        print ('Processing year ' + str(agno))
58                        time.sleep(60)
59                        i=i-1
60                    if res_state == 404:
61                        print ('No data for this station, I skip it...')
62                    # if i%10 == 0:
63                    #     time.sleep(60)
64                archivo.close()
65 if __name__ == '__main__':
66     # El usuario tiene que definir agno inicio, fin y variable
67     if len(sys.argv) < 4 or len(sys.argv) > 4:
68         print "uso: <agno inicio><agno fin><variable:tm_mes,tm_max,tm_min>"
69         raise SystemExit
70     inicio = sys.argv[1]
71     fin = sys.argv[2]
72     variable = sys.argv[3]
73     carga_aemet(inicio,fin,variable)
74     raise SystemExit
```

## ANEXO II. Script para la importación de los datos netCDF SPAIN02.

```
1 import sys
2 import datetime
3 import psycopg2
4 from osgeo import gdal, ogr, osr
5 from osgeo.gdalconst import GA_ReadOnly, GA_Update
6 def restart_line():
7     sys.stdout.write('\r')
8     sys.stdout.flush()
9 def import_spain02():
10     # Registra drivers gdal
11     gdal.AllRegister()
12     # Driver de postgis, para poder crear la tabla
13     driver = ogr.GetDriverByName('PostgreSQL')
14     # Hora de comienzo
15     hora_ini = datetime.datetime.strptime(datetime.datetime.now(), '%Y-%m-%d %H:%M:%S')
16     print 'Iniciado a las: ' + hora_ini
17     # Leemos la primera banda del dataset (el mes que se indique)
18     n_meses = 792
19     archivo = '<RUTA_AL_ARCHIVO_NETCDF>'
20     dataset = gdal.Open(archivo, GA_ReadOnly)
21     bandln = 1
22     cols = dataset.RasterXSize
23     rows = dataset.RasterYSize
24     bands = dataset.RasterCount
25     # Creamos la tabla Postgis para nuestros datos en el esquema public
26     table_name = 'test.spain02_tmax_1950_2015'
27     con = psycopg2.connect("<parametros de la conexion postgresql/postgis>")
28     cur = con.cursor()
29     orden_exec = "CREATE TABLE " + table_name + " (id VARCHAR(20), lat NUMERIC, lon NUMERIC,
30     mes INTEGER, agno INTEGER, tmax NUMERIC)"
31     cur.execute(orden_exec)
32     # Definimos tamaño de pixel y origen de coordenadas
33     pixelWidth = 0.1
34     pixelHeight = 0.1
35     xOrigin = -9.35
36     yOrigin = 43.85
37     #definimos variables para campos mes y agno
38     mes_val = 1
39     agno_val = 1950
40     n_iter = n_meses + 1
41     #iteramos bandas
42     #contador
43     id_n=0
44     for b in range(bandln,n_iter):
45         id_n+=1
46         percent = id_n * 100 / n_meses
47         sys.stdout.write('porcentaje completado: ' + str(percent))
48         sys.stdout.flush()
49         restart_line()
50         band = dataset.GetRasterBand(b)
51         data = []
52         band_data = band.ReadAsArray(0, 0, cols, rows)
53         data.append(band_data)
54     # Iteramos filas y columnas y definimos x e y para crear la geometria
55     for r in range(rows):
56         y = yOrigin - (r * pixelHeight)
57         for c in range(cols):
58             x = xOrigin + (c * pixelWidth)
59             # Gestion de los null
60             value = data[0][r,c]
61             id_value = str(int(x*100)) + 'x' + str(int(y*100)) + 'y'
62             if value > -99:
63                 # Para cada celdilla, anadimos un punto en la capa Postgis
64                 orden_exe = "INSERT INTO " + table_name + " VALUES('" + id_value + "','" + str(y) +
65                 "','" + str(x) + "','" + str(mes_val) + "','" + str(agno_val) + "','" + str(value) +
66                 "','"
67                 cur.execute(orden_exe)
68                 # Definimos la geometria de la capa y finalizamos su creacion
69                 if mes_val == 12:
70                     mes_val = 1
71                     agno_val+=1
72                 else:
73                     mes_val+=1
74                 con.commit()
75                 print 'Finalizado a las: ' + datetime.datetime.strptime(datetime.datetime.now(),
76                 '%Y-%m-%d %H:%M:%S')
77 if __name__ == '__main__':
78     # El usuario tiene que definir al menos dos parametros: la cadena de conexion Postgis
79     GDAL, agno de comienzo
80     if len(sys.argv) <1 or len(sys.argv) > 1:
81         print "uso: <Sin parametros>"
82         raise SystemExit
83     import_spain02()
84     raise SystemExit
```