DEVELOPMENT OF HIGH RESOLUTION GRIDDED DATASETS OF MONTHLY TEMPERATURE SINCE 1916 FOR SPAIN

Andrés CHAZARRA, Belinda LORENZO, Roser BOTEY

Spanish State Meteorological Agency (AEMET)
Leonardo Prieto Castro 8, 28071 Madrid, Spain
achazarrab@aemet.es

Abstract

This article describes the methodology used in the Spanish State Meteorological Agency (AEMET) for obtaining gridded datasets of monthly minimum, maximum and mean temperature with 1 × 1 km spatial resolution for Spain, covering the period 1916-2018. These datasets have been created for climate analysis and monitoring, and will be updated periodically to extend the time coverage. The data used to produce the grids have undergone a quality control process in order to remove or correct erroneous data. The spatial interpolation method consists on a multiple linear regression with ordinary kriging of the regression residuals, using elevation, easting, northing and distance to the coast as independent variables in the regression. The performance of the interpolation method and the accuracy of the grids are evaluated using a cross-validation approach to estimate the errors. Some examples of derived products are shown, as well as a temperature analysis over the 1916-2018 period in Spain based on the gridded datasets.

Key words: temperature, variability, climate change, geographic information system, grid.

1. INTRODUCTION

Recently, 1 x 1 km gridded datasets of monthly temperature for Spain covering the period 1961–2018 have been created and published by the Spanish State Meteorological Agency for climate analysis and monitoring (Chazarra et al., 2020).

These datasets are continuously updated to extend the time coverage to the present. Every month the new monthly provisional grids are generated using the temperature data available, and six months later the definitive grids are created using the complete network data.

New mean temperature series in Spain for the 1981-2010 reference period based on these gridded datasets have replaced the previous reference series used for climate monitoring, which were based on a set of 42 stations distributed along mainland Spain and the Balearic Islands.

In addition, some studies have been done to extend the gridded dataset back in time as long as possible using the same methodology. At present, temperature gridded datasets have been generated for mainland Spain and the Balearic Islands since 1916.

In this article, the methodology used for creating the grids since 1916 is described and the resulting 1916-2018 gridded datasets are analysed.

2. METHODOLOGY AND DATA

2.1. STUDY AREAS AND DATA

Two study areas were selected: the first one covers mainland Spain, the Balearic Islands and the autonomous cities of Ceuta and Melilla (mainland Spain area), whereas the second one covers the Canary Islands (Canary Islands area), as shown in Figure 1.

The coordinate reference systems used were ETRS89 / UTM zone 30N (EPSG: 25830) for the mainland Spain area and REGCAN95/UTM zone 28N (EPSG: 4083) for the Canary Islands area.

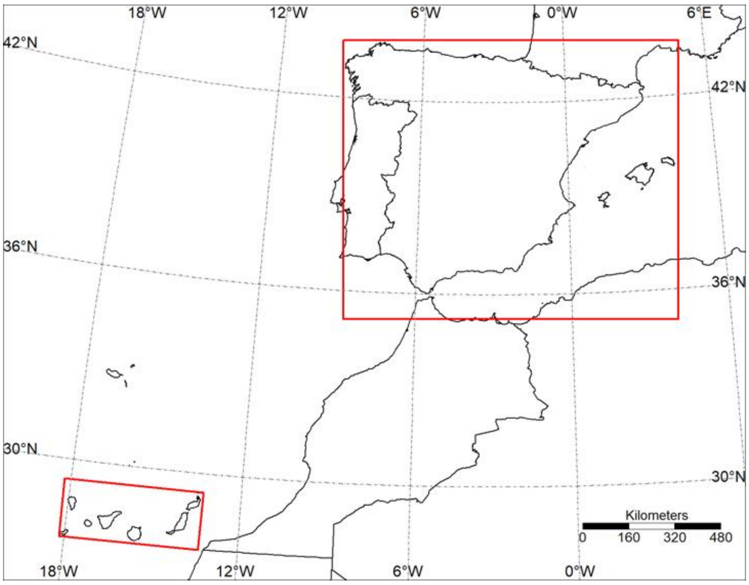

Fig. 1.: Study areas: mainland Spain area and Canary Islands area (in red color).

The variables considered for creating the grids were the monthly mean daily maximum temperature, the monthly mean daily minimum temperature and the monthly mean temperature. The monthly mean daily maximum and minimum temperature grids were spatially interpolated from the data, whereas the monthly mean temperature grids were derived by calculating the mean of the maximum and minimum temperature grids, in order to ensure the consistency of the three variable grids.

It should be noted that the station data were not homogenized before the spatial interpolation and, therefore, the resulting grids are not suitable for studying the climate variability in specific points. Nevertheless, they can be useful for climate monitoring in large areas such as provinces, autonomous regions or the whole country, where the effects of possible local inhomogeneities are expected to cancel each other out to a certain extent.

2.2. SELECTION OF THE STUDY PERIOD

Before selecting the time period for the study, all the monthly temperature data recorded in the Spanish National Climate Database since the beginning of the observations were analysed. The number of temperature stations with valid monthly temperature data for the two study areas are shown in figure 2.

In the mainland Spain area, the number of stations with monthly temperature data was fewer than 20 until the late 1870s. During the following decades, the number of stations increased slowly, and after 1911, the number of stations began to grow quickly and reached more than 100 in the second half of the 1910s. In the 1920s, the number of stations did not change significantly, but in the early 1930s the number rapidly rose, reaching 250 stations in 1936. During the Spanish Civil War (1936-1939) almost half of the stations stopped measuring data. After the war, the number of stations experienced a steady increase for several decades, reaching 1400 stations in 1975. Since then, the number of stations have been oscilating around 1600.

In the Canary Islands, the number of stations with valid monthly temperature data was fewer than 5 stations until 1945. During the late 1940s and the 1951-1960 decade the number of station oscilated between 15 and 30, and only after 1961 the number of valid data was constantly higher than 20.

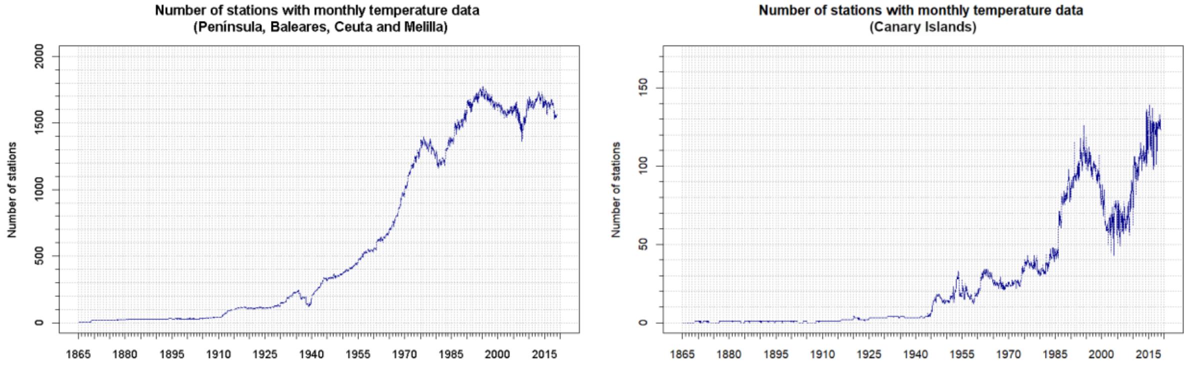

Fig. 2.: Number of stations with valid monthly temperature data in the two study areas since 1865.

After the analysis of the monthly temperature data avalaible, some spatial interpolations were carried out, by using the same interpolation method that will be used for creating the definitive grids, to identify the minimum number of data needed to generate valid grids, rejecting those grids that showed artifacts or noise due to a low data density. The visual analysis of the resulting grids since 1865 led to the conclusion that it is possible to generate high enough quality grids with 100 or more stations evenly distributed in the mainland Spain study area, and at least 20 stations in the Canary Islands area. These minimum numbers of stations were available after the first half of the 1910s in the mainland Spain area, and after 1961 in the Canary Islands area.

On the other hand, it is well known the existence of inhomogeneities in the historical temperature series of many countries due to the use, in the past, of meteorological shelters that differ from the current models (Parker, 1994). In Spain, during the XIXth century and the first years of the XXth century different screen models were used before the Stevenson screen became a standard. The most frequent pre-Stevenson shelter were the Montsouris and Glaisher open wooden stands, often called *facistoles* (lecterns) in Spanish because of their similar appearance to a choir lectern (Giménez, 1992).

According to some studies carried out in Spain (Brunet et al., 2004), in which temperatures simultaneously registered in both Stevenson and Montsouris shelters were compared, the daily maximum temperatures recorded under Montsouris stands are considerably overestimated (between 0.14 and 0.28 ºC in average, depending on the season and the location), whereas the daily minimum temperatures are slightly subestimated. Therefore, temperature data recorded under the old open stands should no be used for climate monitoring climate and especially for temperature trend analysis without a previous bias correction.

The Stevenson screen introduction in the Spanish meteorological network was gradual: although the replacement took place in the Astronomic Observatory of Madrid station in 1894, in other principal stations the Montsouris stands were not substituted by Stevenson shelters until the mid 1910s (Observatorio Central Meteorológico, 1918).

Taking all the above into account, January 1916 was finally chosen as the starting date for the study period in the mainland Spain area, since the number os stations was over the minimum number previously specified to obtain grids with enough quality, and the introduction of the Stevenson shelters had been already taken place in the main meteorological stations.

In the Canary Islands, the study was not finally extended back in time due to the low density of stations before the 1960s.

2.3. DATA QUALITY CONTROL

At present, several data quality control procedures are applied in the Spanish National Climate Database according to the World Meteorological Organization guidelines (WMO, 2018).

However, as these processes have evolved over time, it is possible to find in the historical temperature series anomalous data that have not undergone the present quality control procedures, so it is recommendable to incorporate additional data controls to improve the grid quality.
For this purpose, an automatic validation process has been applied to test the spatial consistency of the data. Every temperature data is compared to the estimated value obtained in that point by interpolating the neighbour data using the same method that will be used for creating the grids, and those data that significantly differ from the estimated value are rejected (Chazarra et al., 2020). Similar automatic control processes are currently applied in other European meteorological services for producing historical climate grids (Hollis et al., 2019).

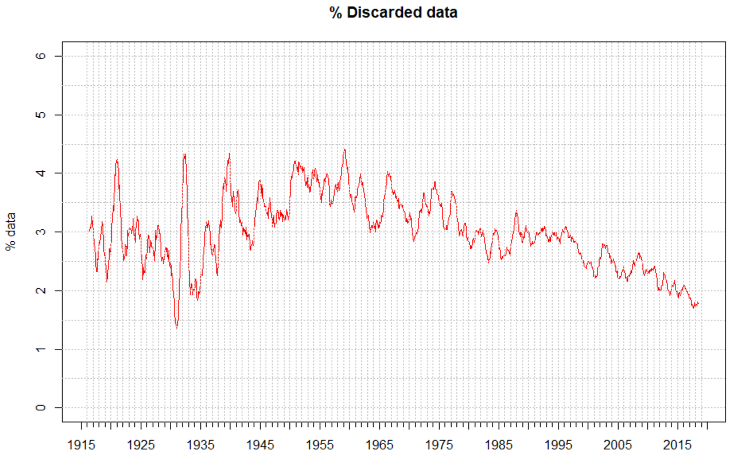

Fig. 3.: Percentage of monthly temperature data discarded in the automatic validation process in the mainland Spain area.

As shown in figure 3, the percentage of the discarded data was relatively small during the first two decades of the study period, highly oscillating around 3 %. In the late 1930s, there was a rise in that percentage into line with the increased number of stations, reaching 4 % in the early 1950s. Since then, the percentage of discarded data has constantly declined until the end of the study period, being fewer than 2 % in the recent years.
This analysis shows that the automatic control process works properly when there is a high density of stations so that it is possible to compare effectively each temperature data with its neighbour data. Therefore, the continuous decrease of the discarded data from 1960 up to now indicates a progressive increase in the quality of the temperature data of the Spanish National the National Climate Database. On the contrary, the relatively low percentage of the discarded data and the high variability of this figure in the first decades can be attributed to a small density of stations that makes difficult to find representative neighbourg data to compare. Consequently, it is to be expected a significantly lower quality in the first half of the study period grids, associated to the lower number of data and the reduced effectiveness of the automatic control process to detect anomalous data.

2.4. GRID INTERPOLATION AND ERROR ESTIMATION

The monthly gridded temperature datasets were created by spatial interpolation of the temperature data that have passed the quality control process, with 1 × 1 km spatial resolution.
The monthly minimum and maximum temperature data interpolation consisted on a multiple linear regression with ordinary kriging of the regression residuals, using elevation, easting, northing and distance to the coast as independent variables in the regression. An exponential model was used for the variogram adjustment.

The performance of the spatial interpolation method and the accuracy of each grid were estimated by leave-one-out cross validation, calculating for each gridded dataset the root mean square error (RMSE) and the mean absolute error (MAE), the most widely reported average-error measures (Willmott & Matsuura, 2006).
The monthly mean gridded datasets were derived by calculating the mean of the maximum and minimum temperature grids in each grid point. Aditionally, the error statistics were derived from the leave-one-out cross validation of the corresponding maximum and minimum temperature grids.

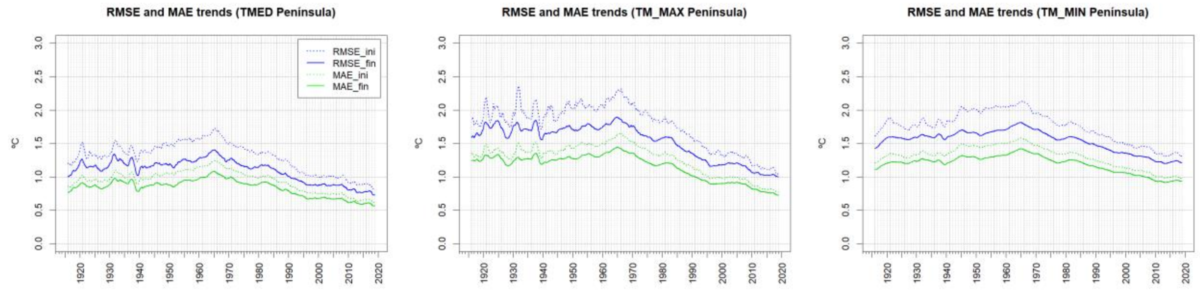

Fig. 4.: Trends of the root mean square error (RMSE) and the mean absolute error (MAE) for the mean, maximum and minimum temperatures before and after the automatic quality control process (mainland Spain area).

The trends of the root mean square error and the mean absolute error for the three considered variables, before and after the data quality control process, are shown in figure 4. As can be seen in that figure, the quality control slightly reduced the errors during the first two decades of the study period, when the density of the data is low. From 1940 to now, the reduction of the errors is much more evident. On the other hand, the magnitude of the errors was approximately constant or even showed a small increase until the mid 1960s, whereas from 1965 until the end of the study period the errors constantly decreased, indicating a continuous improvement of the data and grid quality since then.
The software used for creating the gridded dataset, including the automatic control process and the estimation of the errors, was the open source geographic information system SAGA GIS, version 6.3.0 (Conrad et al., 2015). All the procedures were programmed in R language (R Core Team, 2017) using several packages. These include, among others, RSAGA (Brenning et al., 2018), tmap (Tennekes, 2018) and raster (Hijmans, 2017).

3. RESULTS

Some examples of derived products are shown, as well as a temperature analysis over the 1916-2018 period in the mainland Spain area based on the gridded datasets.

3.1. GRIDDED DATASETS OF MONTHLY, ANNUAL AND SEASONAL TEMPERATURES 1916-2018

The gridded datasets of monthly mean, maximum and minimum temperatures for the study period 1916-2018 were created by following the procedure described above. As an example, the gridded datasets of April 1916 are shown in figure 5.

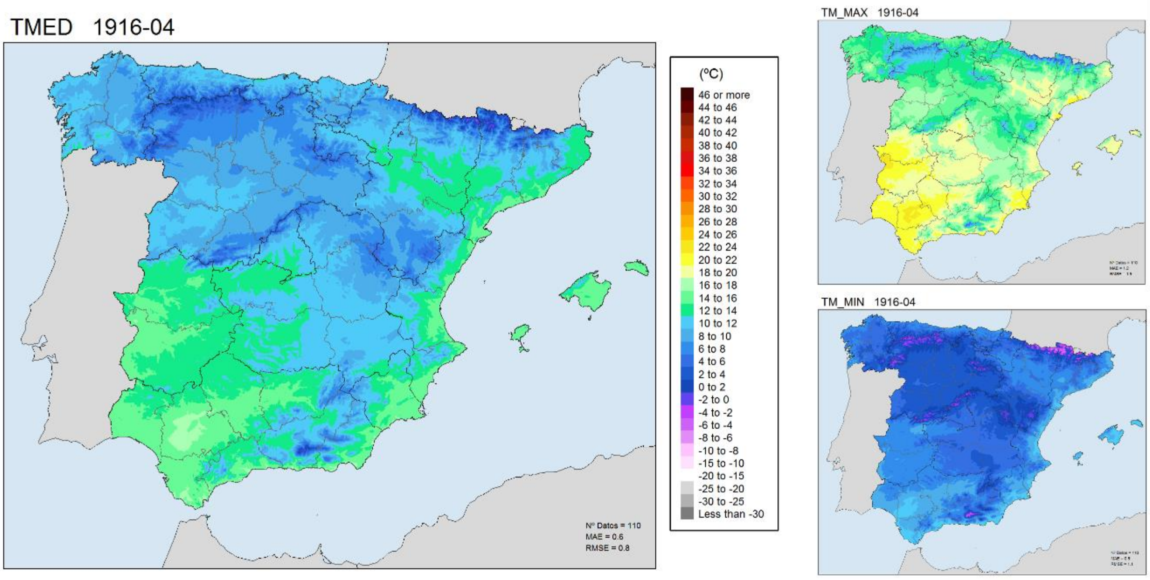

Fig. 5.: Gridded monthly mean (left), maximum and minimum temperatures (right) of April 1916.

Aditionally, annual and seasonal gridded datasets for the study period were calculated based on the monthly gridded datasets, as well as mean gridded datasets for each decade and each 30-year reference period in the study period for climate monitoring in Spain.

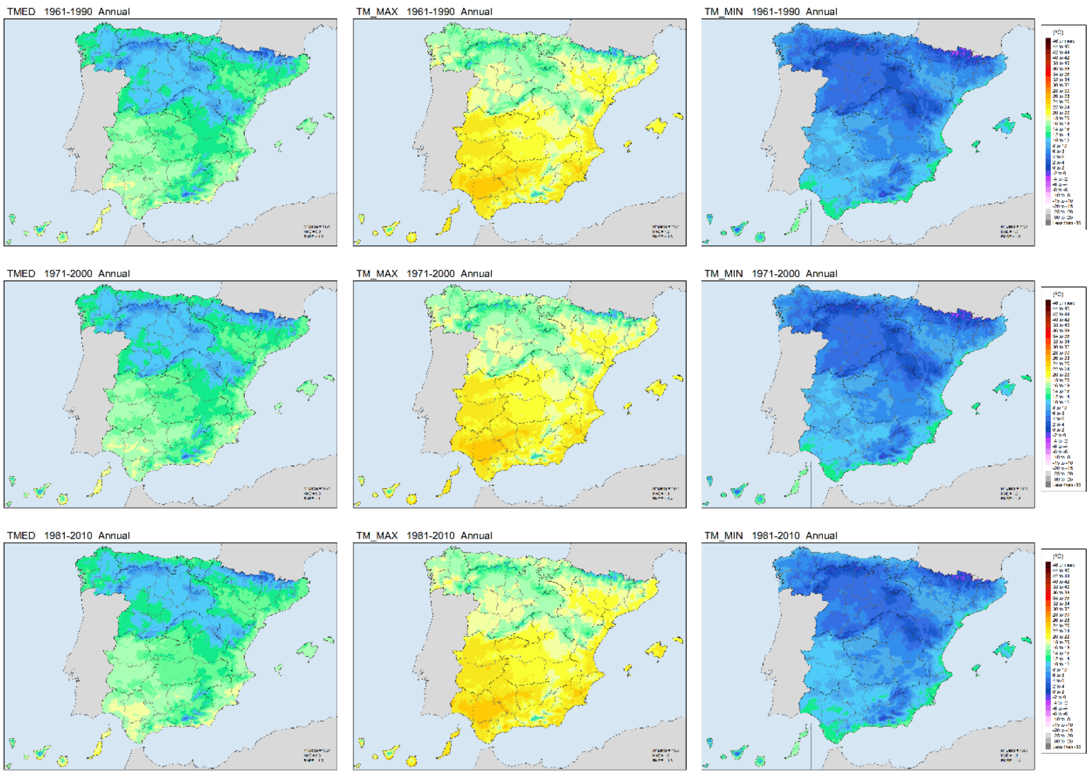

Fig. 6.: Example of several 30-year normal temperature grids.

These gridded temperature datasets have also been used to calculate the temperature anomaly series in mainland Spain for different reference periods, as shown in Figure 7.

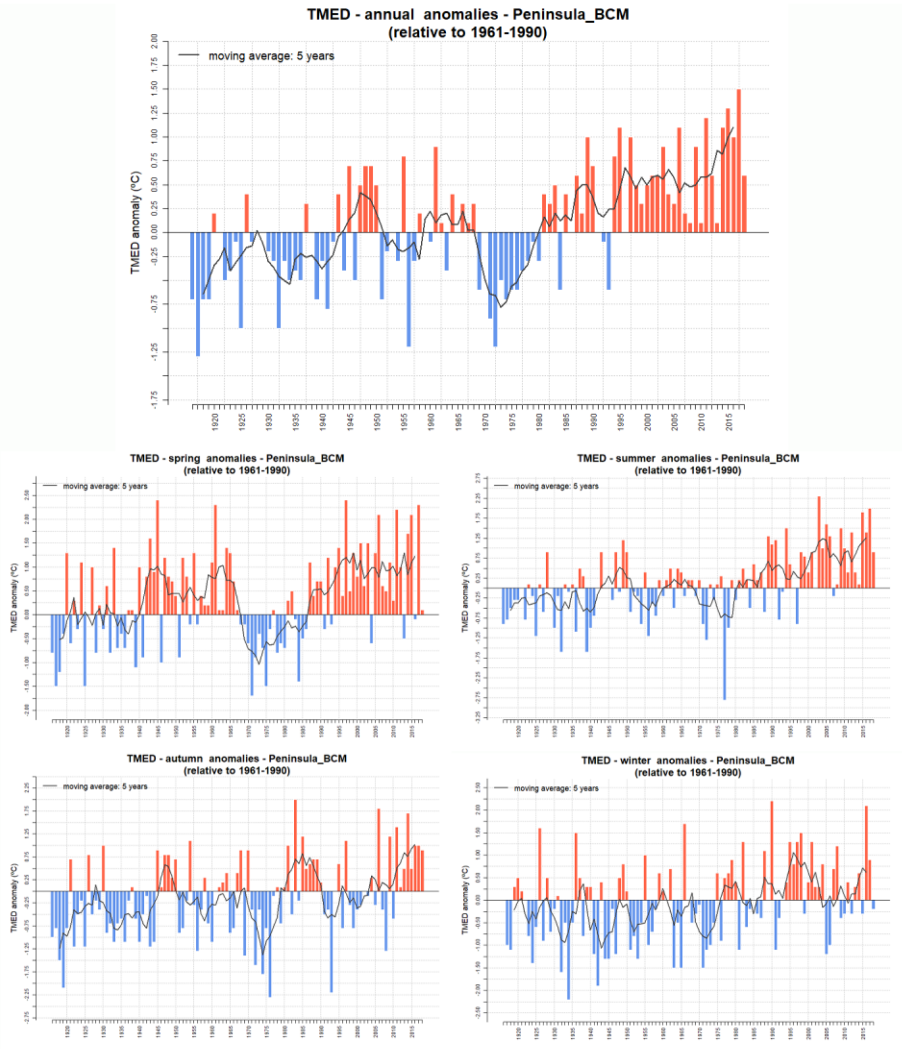

Fig. 7.: Annual and seasonal mean temperature anomaly series relative to 1961-1990 mean for the mainland Spain area.

3.2. TREND ANALYSIS OF THE ANNUAL AND SEASONAL TEMPERATURES IN THE MAINLAND SPAIN AREA 1916-2018

The annual and seasonal mean temperature series in the mainland Spain study area during the 1916-2018 period are shown in figure 8 (blue color), as well as the linear trends calculated by the method of least squares (red color). Additionally, the corresponding series and trends for the maximum and minimum temperatures are shown in figure 9. All the annual and seasonal trends for the 1916-2018 period were statistically significant with a significant level of 0.01 (table 1).

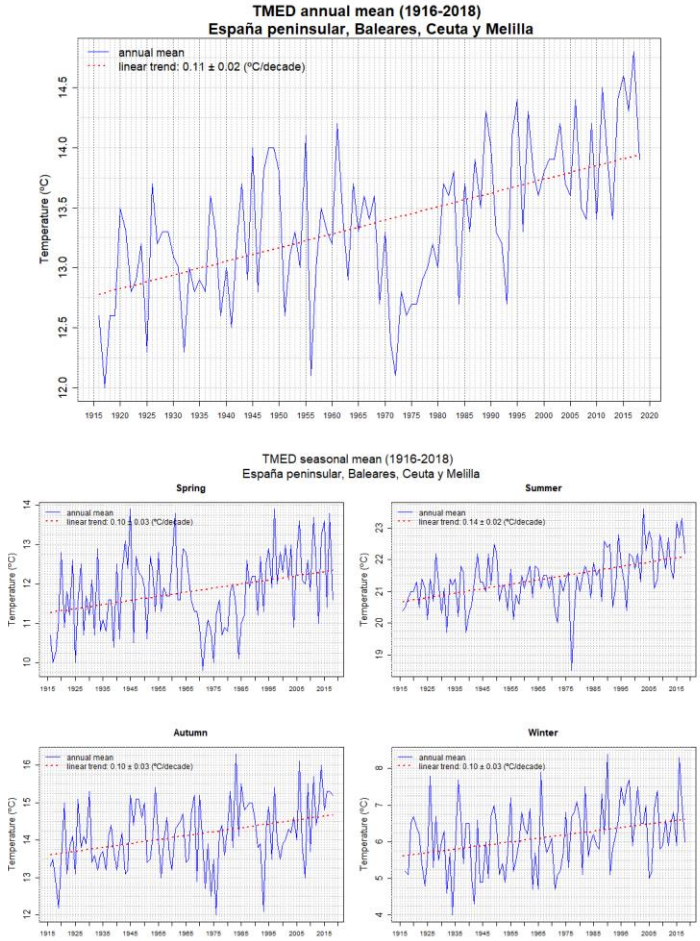

Fig. 8.: Annual and seasonal mean temperature series and trends in the mainland Spain area during the 1916-2018 period.

At first glance, there was a growing trend in the mean annual temperature series since the beginning of the study period until the early 1960s. A sudden cooling was observed during the late 1960s and the early 1970s, with a relative minimum in 1972. After that cold period, mean annual temperatures increased again, showing a noticeable warming in the mid and late 1970s and during the 1980 decade. In the early 1990s a short-term cold period of about three years occurred, which has been attributed to the Pinatubo eruption (González-Hidalgo et al., 2016). During the second half of the 1990s and the 2000 decade a slowdown in the warming was observed: although the mean annual temperatures were relatively high, they did not show any apparent growth trend. Finally, in the 2010s the warming trend has grown again, especially during the second half of the decade, in which the highest temperature values of the study period have been recorded.

According to the least square adjustment results, the mean annual warming in Spain has increased at 0.11 ± 0.02 ºC/decade during the 1916-2018 period. Maximum temperatures have increased at greater rates (0.16 ± 0.02 ºC/decade) than minimum temperatures (0.07 ± 0.02 ºC/decade), with a subsequent rise in the global diurnal temperature range of 0.09 ºC/decade. These results are similar to those of Brunet et al. (2007), which estimated from a series based on a set of stations a mean annual warming of 0.13 ± 0.03 ºC/decade for the mainland Spain in the 1901-2005 period, with larger rates of change for the maximum temperatures (0.17 ± 0.04 ºC/decade) than for the minimum temperatures (0.09 ± 0.03 ºC/decade).

Summer is the season that has contributed most to the annual warming (0.14 ± 0.02 ºC/decade), with a larger rate for the maximum temperatures (0.20 ± 0.03 ºC/decade) than for the minimum temperatures (0.08 ± 0.02 ºC/decade). Spring maximum temperatures has the second higher rate (0.17 ± 0.04 ºC/decade). The lowest warming rates were observed in the spring and winter minimum temperatures (0.04 ± 0.03 ºC/decade and 0.06 ± 0.04 ºC/decade, respectively).

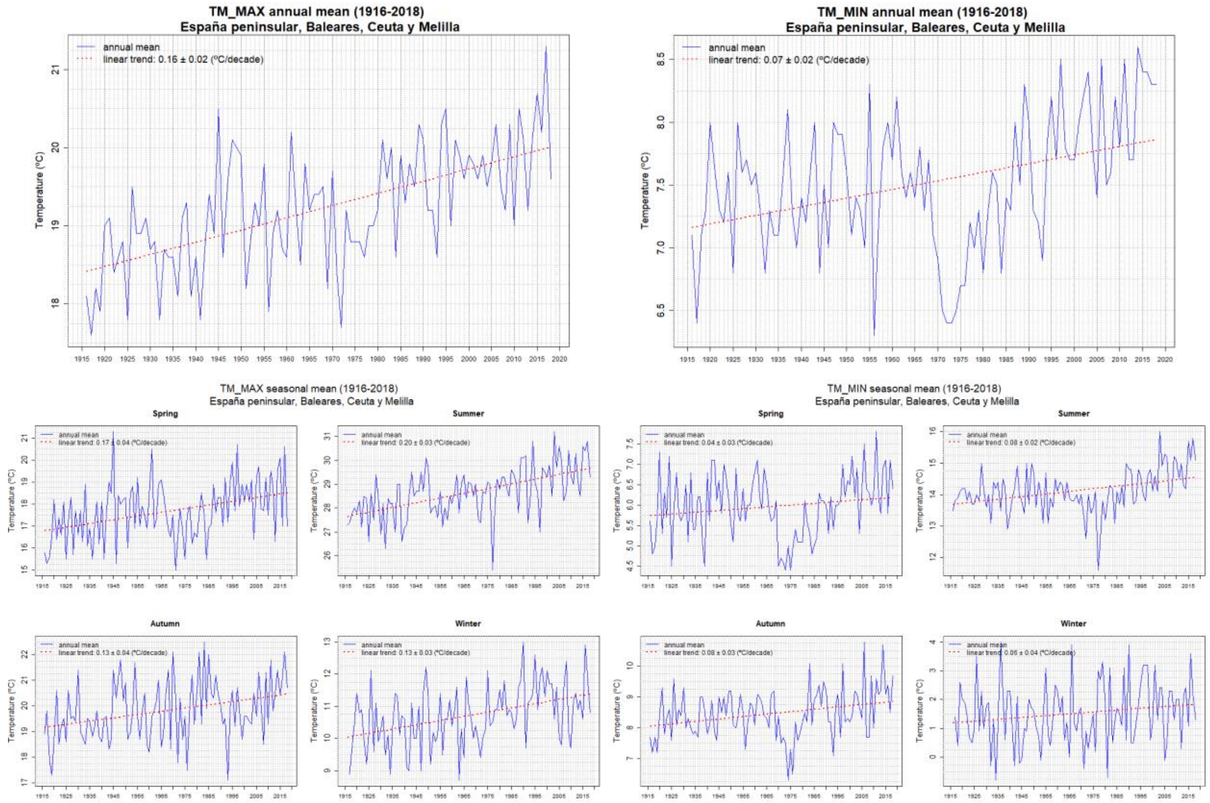

Fig. 9.: Annual and seasonal maximum and minimum temperature series and trends in the mainland Spain area during the 1916-2018 period.

Table 1.: Temperature trends in the 1916-2018 period and in 30-year periods (in bold, tendencies statistically significant at the 0.01 level; in bold and italics, tendencies statistically significant at the 0.05 level).

TMED (ºC)

	Spring	Summer	Autumn	Winter	Annual
1916-2018	*0.10 ± 0.03*	*0.14 ± 0.02*	*0.10 ± 0.03*	*0.10 ± 0.03*	*0.11 ± 0.02*
1921-1950	0.35 ± 0.19	0.28 ± 0.15	0.21 ± 0.15	0.00 ± 0.20	0.23 ± 0.09
1931-1960	0.22 ± 0.19	0.08 ± 0.15	0.18 ± 0.15	0.14 ± 0.19	0.16 ± 0.11
1941-1970	-0.13 ± 0.19	-0.02 ± 0.12	0.02 ± 0.16	0.27 ± 0.18	0.01 ± 0.11
1951-1980	***-0.54 ± 0.17***	-0.09 ± 0.15	-0.13 ± 0.17	0.23 ± 0.18	-0.15 ± 0.10
1961-1990	-0.17 ± 0.19	0.12 ± 0.16	0.32 ± 0.20	0.34 ± 0.19	0.14 ± 0.11
1971-2000	***0.81 ± 0.13***	***0.46 ± 0.17***	0.24 ± 0.21	**0.47 ± 0.18**	***0.49 ± 0.09***
1981-2010	**0.39 ± 0.16**	0.29 ± 0.16	-0.25 ± 0.19	0.07 ± 0.19	0.11 ± 0.09
1991-2018*	0.12 ± 0.21	0.36 ± 0.18	***0.60 ± 0.19***	0.05 ± 0.20	**0.29 ± 0.10**

TM_MAX (ºC)

	Spring	Summer	Autumn	Winter	Annual
1869-2018	*0.17 ± 0.04*	*0.20 ± 0.03*	*0.13 ± 0.04*	*0.13 ± 0.03*	*0.16 ± 0.02*
1921-1950	0.51 ± 0.27	***0.48 ± 0.19***	0.35 ± 0.20	0.12 ± 0.19	0.34 ± 0.13
1931-1960	0.33 ± 0.28	0.14 ± 0.19	0.25 ± 0.22	0.15 ± 0.18	0.20 ± 0.14
1941-1970	-0.04 ± 0.28	0.04 ± 0.16	0.03 ± 0.25	0.17 ± 0.19	0.04 ± 0.15
1951-1980	**-0.52 ± 0.24**	0.13 ± 0.18	0.10 ± 0.25	0.23 ± 0.17	-0.02 ± 0.12
1961-1990	-0.17 ± 0.25	0.21 ± 0.19	0.41 ± 0.26	**0.44 ± 0.17**	0.21 ± 0.13
1971-2000	***0.95 ± 0.21***	**0.42 ± 0.22**	-0.01 ± 0.27	***0.55 ± 0.15***	***0.47 ± 0.11***
1981-2010	0.38 ± 0.23	0.21 ± 0.20	-0.41 ± 0.23	-0.08 ± 0.19	0.02 ± 0.11
1991-2018*	0.01 ± 0.29	0.36 ± 0.24	***0.76 ± 0.21***	-0.02 ± 0.21	0.29 ± 0.13

TM_MIN (ºC)

	Spring	Summer	Autumn	Winter	Annual
1869-2018	*0.04 ± 0.03*	*0.08 ± 0.02*	*0.08 ± 0.03*	*0.06 ± 0.04*	*0.07 ± 0.02*
1921-1950	0.18 ± 0.16	0.12 ± 0.11	0.08 ± 0.12	-0.10 ± 0.23	0.07 ± 0.08
1931-1960	0.12 ± 0.15	0.03 ± 0.13	0.11 ± 0.11	0.13 ± 0.24	0.10 ± 0.10
1941-1970	-0.20 ± 0.13	-0.10 ± 0.11	0.02 ± 0.11	0.34 ± 0.20	-0.01 ± 0.10
1951-1980	***-0.54 ± 0.14***	**-0.32 ± 0.13**	**-0.37 ± 0.14**	0.22 ± 0.23	**-0.27 ± 0.10**
1961-1990	-0.19 ± 0.15	0.04 ± 0.15	0.20 ± 0.18	0.27 ± 0.25	0.06 ± 0.11
1971-2000	***0.66 ± 0.08***	***0.51 ± 0.13***	***0.48 ± 0.17***	0.41 ± 0.25	***0.51 ± 0.08***
1981-2010	***0.40 ± 0.11***	***0.37 ± 0.12***	-0.06 ± 0.18	0.18 ± 0.25	**0.21 ± 0.09**
1991-2018*	0.26 ± 0.14	**0.36 ± 0.13**	**0.47 ± 0.19**	0.09 ± 0.24	**0.29 ± 0.09**

4. CONCLUSIONS

At present, high resolution gridded datasets of monthly temperature for the mainland Spain area since 1916 are available, and for the Canary Islands since 1961. These datasets are continuously updated to extend the time coverage to the present.
The error analysis of the grids shows that the quality of the grids is lower during the first decades of the study period, and it progressively increases after the 1960s.
Due to possible local inhomogeneities, these datasets are not suitable for studying the climate variability in specific points, but they can be useful for climate monitoring in whole country or big areas during the last century. The results of the trend analysis of the annual and seasonal temperatures in the mainland Spain area in the period 1916-2018 are consistent with previous studies based on a set of stations.

References

Brenning, A., Bangs, D., & Marc, B. (2018). RSAGA: SAGA Geoprocessing and Terrain Analysis. R package version 1.1.0. https://CRAN.R-project.org/package=RSAGA

Brunet M., Bañón M., García F., Aguilar E., Saladié O., Sigró J., Asín J., López D. 2004. Una aproximación experimental tendente a la minimización del sesgo artificial asociado al tipo de garita meteorológica a través de la observación dual de la temperatura del aire. In Actas de las XXVIII Jornadas Científicas de la AME, Badajoz, 11–13 February 2004, Pub. Asociación Meteorológica Española y Ministerio de Medio Ambiente: La Meteorología y el Clima Atlánticos, Badajoz: 13.

Brunet, M., Jones, P., Sigró, J., Saladié, O., Aguilar, E., Moberg, A., . . . López, D. (2007). Temporal and spatial temperature variability and change over Spain during 1850–2005. J. Geophys. Res., 112, D12117. doi:10.1029/2006JD008249

Chazarra, A., Lorenzo, B., Rodríguez, C., Botey, R. (2020). Análisis de las temperaturas en España en el periodo 1961-2018. Madrid: Agencia Estatal de Meteorología (AEMET). doi: https://doi.org/10.31978/666-20-003-4

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., . . . Böhner, J. (2015). System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geosci. Model Dev., 8, 1991-2007. doi:10.5194/gmd-8-1991-2015

Giménez, J.M (1992). La Meteorología en el Observatorio Astronómico de Madrid. En “Doscientos años del Observatorio Astronómico de Madrid” (ISBN: 978-84-88405-00-6), editado por la Asociación de amigos del Observatorio Astronómico de Madrid.

González-Hidalgo, J., Peña-Angulo, D., Brunetti, M., & Cortesi, N. (2016). Recent trend in temperature evolution in Spanish mainland (1951–2010): from warming to hiatus. Int. J. Climatol., 36, 2405-2416. doi:10.1002/joc.4519.

Hijmans, R. J. (2017). raster: Geographic Data Analysis and Modeling. R package version 2.6-7. https://CRAN.R-project.org/package=raster

Hijmans, R. J. (2017). raster: Geographic Data Analysis and Modeling. R package version 2.6-7. https://CRAN.R-project.org/package=raster

Hollis, D., McCarthy, M., Kendon, M., Legg, T., & Simpson, I. (2019). HadUK-Grid—A new UK dataset of gridded climate observations. Geosci Data J.(6), 151-159. doi: https://doi.org/10.1002/gdj3.78

Observatorio Central Meteorológico (1918). Resumen de las observaciones meteorológicas efectuadas en las estaciones del Servicio Meteorológico Español durante el año 1916.

Parker, D. E. (1994), Effects of changing exposure of thermometers at land stations, Int. J. Climatol., 14, 1 – 31.

R Core Team. (2017). R: A language and enviroment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Tennekes, M. (2018). tmap: Thematic Maps. R package version 1.11-1. https://CRAN.Rproject.org/package=tmap

Willmott, C. J., & Matsuura, K. (2006). On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. International Journal of Geographical Information Science, 20(1), 89-102. doi:10.1080/13658810500286976

World Meteorological Organization (2018). Guide to Climatological Practices. OMM-nº100. 2018 edition.