

## A COMPLETE DAILY PRECIPITATION DATABASE FOR NORTHEAST SPAIN: RECONSTRUCTION, QUALITY CONTROL AND HOMOGENEITY

Sergio M. Vicente-Serrano<sup>2</sup>, Santiago Beguería<sup>1</sup>, Juan I. López-Moreno<sup>2</sup>, Miguel A. García-Vera<sup>3</sup>,  
and Petr Stepanek<sup>4</sup>

<sup>1</sup>Estación Experimental de Aula Dei, CSIC, Zaragoza, Spain

<sup>2</sup>Instituto Pirenaico de Ecología, CSIC, Zaragoza, Spain

<sup>3</sup>Oficina de Planificación Hidrológica. Confederación Hidrográfica del Ebro, Zaragoza, Spain

<sup>4</sup>Czech Hydrometeorological Institute, Brno, Czech Republic

**Abstract:** This paper reports the procedure used in creating a homogeneous database of daily precipitation in northeast Spain. The source database comprised 3106 daily precipitation observatories, with data ranging from 1901 to 2002. Firstly, a reconstruction of the series was performed. Data from adjacent observatories were combined to provide long temporal coverage. Data gaps were filled using values from nearest neighbor observatories. A distance threshold was set to avoid the introduction of spurious information in the series. Secondly, the reconstructed series were subjected to a quality control process. Empirical percentiles corresponding to each precipitation observation were compared to the percentiles corresponding to the closest neighbor observatory, and a threshold difference was set to identify questionable extremes. After careful inspection of each case, 0.1% of the data was rejected and replaced with information from the nearest neighbor. Thirdly, the homogeneity of the series was checked using the standard normal homogeneity test. This allowed detection of inconsistencies present in the original database or introduced by the reconstruction process. Four parameters were assessed at a monthly level: amount of precipitation, number of rainy days, daily maxima, and number of days above the 99th percentile. A total of 43% of the series had some periods of inhomogeneity and were discarded. The final database comprised 828 series with varying time coverages. The greatest number of stations existed during the 1990s, but more than 300 series contained information from the 1960s, and 34 series contained a complete record since 1920. Comparisons of the spatial variability of several parameters describing the daily precipitation characteristics were made. The results showed that the

final database had improved spatial coherence. The process described here is proposed as a model for developing a standard procedure for the construction of databases of daily climate data.

**Key-words:** Daily precipitation, database, quality control, homogenization, SNHT, extreme events, climatic variability, climatic change, Spain.

## **1. Introduction**

Climate hazards cause human and economic impacts on a global scale. Floods and droughts in particular have major human and economic effects (Bruce, 1994; Obasi, 1994), and both are closely related to precipitation intensity, frequency and duration. An increase in precipitation-related natural hazards and consequent economic loss have been reported by several authors (Karl and Easterling, 1999; Meehl et al., 2000; Peterson et al., 2002; Aguilar et al., 2005; Hundedcha and Bardossy, 2005; Moberg and Jones, 2005), and can be attributed to an increase in the vulnerability of society to these events (Kunkel et al., 1999). Climate change models indicate that the frequency and magnitude of extreme precipitation events could increase markedly in a number of regions during this century (e.g. Boronenat et al., 2006; Frei et al., 2006; Pall et al., 2007). This is likely to be accompanied by changes in the frequency distribution of precipitation (Katz and Brown, 1992; Easterling et al., 2000), affecting both the intensity and duration of precipitation-related extreme events, such as drought and floods.

Accurate estimates of the risk of extreme precipitation events and droughts are needed for agricultural and water management, land planning, public works programs and in other sectors. Extreme value analysis techniques are normally used to obtain frequency estimates of dangerous phenomena, and to enable the probability of occurrence and the return period of extreme events of a given magnitude and/or duration to be assessed (e.g. Smith, 1989 and 2003).

The analysis of precipitation-related processes requires spatially dense databases with high frequency precipitation records (daily or sub-daily). The vast majority of available information is at

a daily time scale, and sub-daily information is not widely available. Long time series datasets are the most valuable, as the reliability of frequency estimations is closely related to the sample size used during the analysis process (Porth et al., 2001). Long time series are also necessary in analysis of the temporal variability and trends of extreme events, and to estimate the risk and probability of these events.

Long-term, dense and reliable daily precipitation databases are uncommon for several reasons. Changes in the location of observatories within the same locality are frequent, resulting in fragmented or inconsistent data series. Human error can occur during the process of observation, and in the transcription and digitization of data (Reek et al., 1992). In addition, measurements at a meteorological station can vary as a consequence of instrument deterioration or replacement, variations in the time of observations, and changes in the surrounding environment. These factors increase noise in the data, and can lead to inhomogeneities that make the data unusable (Peterson et al., 1998; Beaulieu et al., 2007).

To overcome these problems, and to construct a reliable database for extreme value analysis, a process of reconstruction, quality control and homogenization of precipitation data is needed. This approach is common for monthly precipitation series, and several complete and homogeneous precipitation databases have been created (e.g. González-Rouco et al., 2000; González-Hidalgo et al., 2004; Brunetti et al., 2006). Several homogeneous databases of temperature series have also been generated at a daily time resolution (e.g. Brunet et al., 2006), and some approaches have been developed for homogenizing this variable (Allen and DeGaetano, 2000; Vincent et al., 2002; Brandsma and Könen, 2006).

However, despite its importance no complete protocols are available for processing daily precipitation datasets. Various procedures have been developed for filling data gaps in daily precipitation series (Karl et al., 1995; Eischeid et al., 2000), and for automatic (Feng et al., 2004) and manual (Griffiths et al., 2003; Aguilar et al., 2005) quality control of the datasets. These

procedures are widely applied before analysis of daily precipitation data. Nevertheless, testing for inhomogeneities in daily precipitation datasets is not common, and although some reports exist (e.g. Schmidli and Frei, 2005; Tolika et al., 2007) they have usually been in relation to precipitation volume and not precipitation frequency, which is also crucial for guaranteeing the homogeneity of daily precipitation datasets (Wijngaard et al., 2003; Viney and Bates, 2004). Although there have been some recent advances in the creation of quality-controlled and homogenized daily precipitation datasets (e.g. Peterson et al., 2008), complete procedures for optimizing all the available information are lacking, as typically only the longest and most complete series are used. This suggests inefficient use of the available data, and the possibility of obtaining spatially dense information, particularly in countries where highly fragmented series not fulfilling the standard criteria of length and completeness exist among neighboring areas.

Some studies in the Iberian Peninsula have focused on the creation of daily precipitation databases for spatial and temporal studies of climate change and climatic hazards. Romero et al. (1998) created 410 complete daily precipitation series for the Spanish Mediterranean provinces, using information derived from 3366 individual series. This analysis was mainly focused on filling gaps, and the homogeneity of the resultant series was not checked. Moreover, the final dataset is temporally limited, covering only the period from 1964 to 1993. Lana et al. (2004) developed a daily precipitation database for 1950–2000 from 75 rain gauges in Catalonia (northeast Spain). Homogeneity was checked using monthly totals, but there were gaps in the dataset. Although some global daily precipitation databases exist (e.g. Peterson et al., 1997; Gleason, 2002), the spatial density of data is very poor in most regions. For example, the Global Daily Climatology Network (GDCN; <http://www.ncdc.noaa.gov/oa/climate/research/gdcn/gdcn.html#precip>) includes only 19 precipitation observatories for the Iberian Peninsula, which is inadequate for capturing the large spatial variability that characterizes precipitation in this region. Klein-Tank et al. (2002) also

developed a database for daily precipitation in Europe, but the spatial density is too low (10 observatories in Spain) to enable adequate spatial analysis.

This paper presents a process for reconstruction of a spatially dense database of daily precipitation records for the northeast part of Spain, using data since 1900 from the archives of the Spanish National Institute of Meteorology. The process included selection of suitable observatories for the reconstruction, gap filling, identification of anomalous and questionable records, and homogeneity testing. The objective was to construct a spatially dense, continuous, long and reliable database for climate studies, reducing as much as possible the noise-to-signal ratio and eliminating all likely inconsistencies.

## **2. Database**

The original database comprised 3106 daily precipitation observatories in the study area, whose activities spanned the period of existence of the Spanish National Institute of Meteorology (1900–2002). The boundaries of the database correspond to administrative limits, and include 18 provinces in the northeast of Spain with a total area of 159,423.7 km<sup>2</sup> (Figure 1). The spatial density is very high (one observatory per 51.3 km<sup>2</sup>), although variation in the density among regions is also high.

Normalization of instruments is very important to ensure the temporal homogeneity of the data. Pluviometers used in the Spanish observation network are approved and normalized by the National Institute of Meteorology, which provides the instruments and guarantees their uniformity. The Hellmann pluviometer was officially adopted in 1911 for the Spanish precipitation network. This device is characterized by a hollow cylinder with a funnel formed at one end, and is placed on a pole at a height of 1.5 m. The opening of the pluviometer has a surface area of 200 cm<sup>2</sup>, and it has a brass ring with a beveled edge to ensure the surface area remains constant and to avoid splash. Measurements are taken twice daily, at 9:00 h and 14:00 h. Since 1911 the measurement protocol

has remained unchanged. Only two of the data series used in this study contained precipitation data prior to 1911. Therefore, the database was not affected by changes in instrumentation or measurement protocol.

The original database series were highly variable in terms of record lengths and the quantity and duration of data gaps. No information was available about data acquisition and the history of each observatory (metadata), including changes in location, measurement conditions, observers, and observation times. Metadata are very useful for assessing the quality of a series and to identify possible errors, but are frequently missing in raw climate databases.

The data collection periods for the majority of observatories were very short (374 had less than 60 months of complete records, and 1723 had 240 months or less). Only 286 stations had more than 600 months of records (50 years of data). The database was also very fragmented, with several observatories located in the same locality, but covering different periods.

### **3. Methodology**

The method comprised three main steps. The first involved reconstruction of the precipitation series with the objective of deriving continuous and long-term series by combining short duration series from nearby observatories, and the filling of gaps by using auxiliary information obtained from nearby observatories. The second step was a quality control assessment of the reconstructed series to identify and substitute anomalous and questionable records in the database (negative precipitation, extreme precipitation events, some zero values, and records that differed markedly from values recorded in neighboring observatories). The third step tested the homogeneity of the reconstructed series using four parameters of the series. This enabled identification of complete series and removal of periods for which data were not homogeneous; the latter was carried out to avoid the presence of spurious information in the final dataset. The three steps, the decisions taken and the products obtained are described in detail below.

### ***3.1. Reconstruction and gap filling***

The reconstruction of a single long time series from a number of shorter series from neighboring observatories enabled optimization of the highly fragmented daily precipitation data typical of many datasets. Reconstruction relied on the assumption that the cessation of data recording at one observatory, and the establishment of one or more new observatories close to the previous one, results in two or more data series which are usually not useful for climate analysis as a consequence of their short duration. Nevertheless, if the observatories are sufficiently close the differences in precipitation records are usually very small, so data from the shorter series can be combined into a single series, which is ascribed to the last observatory that collected data. It is important to note that this approach is only valid where observatories are separated by short distances, and it is assumed that the combined series can exhibit inhomogeneities due to the reconstruction process. These inhomogeneities need to be identified and removed from further analyses.

A review of the literature showed that no general criterion exists for the selection of observatories suitable for reconstruction. Lana et al. (2004) considered a minimum of 31 years of data in a 50-year period was necessary for a series to be included in a daily precipitation database for Catalonia. In a study involving the Spanish Mediterranean coast, Romero et al. (1998) used only those series with less than 10% of data missing in a 30-year period. Similarly, Eischeid et al. (2000) set a maximum of 48 months of missing data in a 40-year period as the criterion for rejecting a data series in western USA. From a set of 181 series, Haylock and Nicholls (2000) used only those that had less than 10 days missing per year for at least 80 years in the 88-year period they considered. Moberg and Jones (2005) were more restrictive in a trend analysis for the whole of Europe, selecting only those series with less than 3 years with data gaps in a total of 89 years.

We followed different criteria to select suitable data series for reconstructions as a function of temporal duration, the period covered and the data gaps. Data series that covered a period of less than 15 years were considered too short to be suitable for reconstruction. This group of

observatories (1106; labeled *Z*) were reserved for use in reconstructions of data from other observatories. The exception was where *Z* observatories included data from 2000–2002; these series were considered of high value as they could complement long-term data from nearby observatories that had ceased data collection in previous years. A total of 37 series matched this criterion. The remaining series (1963) were divided into two groups. Series *A* included those series (1094) covering a period of more than 25 years with less than 10 years of data gaps. Series *B* included those series (869) covering a period of 15 to 25 years, or a period of more than 25 years but with more than 10 years of data gaps. Series *A* and *B* were considered suitable for reconstruction.

The next step was to fill data gaps in the time period covered by each series. Approaches to gap filling in daily climate series (Eischeid et al., 2000) involve consideration of only the data in the series, or involve use of the data from nearby observatories. Karl et al. (1995) and Brunetti et al. (2001) filled missing values by generating random rainfall amounts, based on the probability distributions of the variables studied. The goal of this procedure was not to give a realistic estimate of the unknown daily values, but to obtain a data series of equal length without changing the probability distributions of rainfall amounts. Nevertheless, for other applications it is more reliable to use methods based on the values recorded at nearby observatories (Paulhus and Kohler, 1952; Eischeid et al., 2000). We focused on methods based on the information from neighboring observatories, and tested three different procedures: the nearest neighbor, inverse distance weighted interpolation, and linear regression methods.

To compare the methods, artificial data gaps were created by randomly removing 1% of the available observations from the *A* and *B* observatories. A total of 1963 series were tested, involving creation of 181,861 artificial data gaps using the nearest neighbor, inverse distance weighted interpolation, or linear regression method. After applying the three methods to these data, the root mean square error (RMSE) of the reconstructed gaps was used to choose the best method.



i) In the nearest neighbor method data gaps were filled directly with data from the closest observatory that had information. To apply this method, two criteria were established: the nearest neighbor had to be within a radius of 15 km of the target observatory, and the correlation (Pearson's R) between the daily precipitation series from both observatories had to be higher than 0.5, with a minimum of 3 years of common data. These criteria were based on the average distance among observatories for the complete dataset, and the average correlations among observatories at different distances. Descriptive statistics showed that the average number of neighboring observatories within a radius of 5 km was very small (3), but increased to nine with a 10 km radius (Fig. 2). However, the availability of observatories was highly variable across regions. Thus, 25% of the observatories has less than six neighbors within a radius of 10 km. To overcome this problem we selected a threshold radius of 15 km, for which only 5% of the observatories had less than 10 neighbors. This threshold was not large enough to lead to important differences in the precipitation conditions among observatories.

As a distance of 15 km may not have ensured similarity in precipitation conditions between two sites (e.g., in the case of strong elevation differences), we established an additional criterion in fixing the distance threshold, based on the correlation between the two series. The average correlation in daily precipitation between pairs of observatories with a minimum of 3 years of common data decreased rapidly as a function of distance from 1 to 50 km (average R from 0.78 to 0.45, Fig. 3). At greater distances the decrease was slower but sustained. At a distance of 15 km the average correlation was  $R = 0.62$ , but for greater distances (e.g. 25 km) the correlations were lower (average  $R = 0.57$ ) in order to achieve a higher number of neighbors. In contrast, for shorter distances (e.g. 10 km,  $R = 0.67$ ) the number of neighbors decreased markedly, as indicated above. Thus, a threshold distance of 15 km appeared to be a good compromise.

- ii) For interpolation from neighboring data series we selected a local method based on the inverse distance weighting (IDW):

$$z(x_j) = \frac{\sum_{i=1}^n z(x_i) \cdot d_{ij}^{-r}}{\sum_{i=1}^n d_{ij}^{-r}}$$

where  $z(x_j)$  is the predicted value according to the weighted average of the data at points  $z(x_1), z(x_2), \dots, z(x_n)$ . The distance ( $d$ ) between  $z(x_i)$  and  $z(x_j)$  is the weighting factor, and we used an  $r$  value of 2. We fixed a maximum distance of 15 km for the interpolation.

- iii) In the linear regression method, missing data were obtained by determining the most correlated single independent series. To avoid negative values and to retain the zeros, the regression line was forced to pass through the origin, providing a model only with a slope coefficient. This approach has been used to reconstruct daily temperature series (e.g. Allen and DeGaetano, 2001), as this variable is not affected by abrupt spatial changes, and varies gradually in space. Linear regression is very suited to obtaining reliable dependence models among a candidate observatory and auxiliary observatories used in the reconstruction. Additional problems arise with daily precipitation series, as these usually show lower correlations even among close observatories (Auer et al., 2005).

Of the three methods the nearest neighbor method provided the best results, with an average RMSE of 1.05 mm (with a range of 0.23–5.7 mm between stations). The IDW approach had an average RMSE of 1.23 mm (range 0.31–7.2), and the linear regression method had an average RMSE of 1.31 mm (range 0.30–6.9).

The performance of the three methods could not be evaluated solely on the basis of the RMSE, because a high RMSE value can mask important changes in the frequency of rainy days and extreme values. The series from the Fabra observatory in Barcelona illustrates this problem (Fig. 4).

This series was of good quality, had complete records (see details about the observatory in Rodríguez et al., 1999), and had a large number of neighboring observatories. The entire series was reconstructed between 1913 and 2002. The conclusions derived from the Fabra observatory can be generalized to the other observatories. Among the three methods, the IDW and regression methods reduced the frequency of extreme values, and increased the frequency of events less than 50 mm. The nearest neighbor method more accurately reconstructed the frequency of the most extreme events in the series. In comparison with the original series, the IDW reconstruction noticeably decreased the total number of dry days and increased the number of rainy days, affecting the number and duration of dry and wet spells and the average precipitation per day. This was a direct consequence of neighbor averaging, as a single neighboring series with a daily record above zero resulted in the reconstructed series changing from a dry to a wet day. The IDW method included contributions from poorly correlated series, as all the observatories within 15 km were included. An alternative strategy may have been to weight the neighboring data according to the correlation coefficient rather than the distance. However, this would not have avoided a decrease in the number of rainy days, as this is intrinsic to the averaging nature of the method, and independent of the weighting criterion chosen. In contrast, the regression method overestimated the number of dry days and underestimated the number of wet days. The nearest neighbor method provided statistics closest to the original record, and also maintained the distribution characteristics of the original series better than the other methods. This evidence favored the nearest neighbor method over more sophisticated procedures involving several neighbors, and we therefore chose this method for gap filling in the dataset. For this purpose we used the  $Z$  series within a 15 km radius of the target  $A$  series, and with a correlation coefficient higher than 0.5. If some gaps remained in the  $A$  series after this process, we also used the  $B$  series (lacking information until 2000–2002) within the 15 km radius. The  $B$  series used to fill gaps in some  $A$  series were discarded in subsequent reconstructions. If no data were available within a 15 km radius, we rejected all the data prior to the gap so as to

avoid potential inaccuracies introduced by using data far from the target observatory. The unique exception to this was the 1936–1939 period, during which the instability caused by the civil war markedly reduced the number of observatories. In the few observatories for which gaps remained during these years, the earlier information was not deleted. A total of 862 observatories from the 1094 original *A* series were filled following this criterion. The remaining observatories (232) had the data removed before the gaps.

The remaining *B* series were also completed using the *Z* series and neighboring *B* series, always located within a radius of 15 km. The *Z* series were used first, and the remaining gaps were completed using the *B* series. To avoid redundant information we gave preference to the *B* series with data until 2000–2002. If both series did not reach this date, the shorter series was used to complete the longer series. The series used for gap filling were subsequently discarded. The data gaps in the remaining 37 *Z* series with data until 2000–2002 were filled following the same procedure.

After the gap filling procedure there were 1663 complete series comprising 1094 (*A*), 532 (*B*) and 37 (*Z*, until 2000–2002). Many of the completed series covered different periods. As the objective was to obtain a complete and reliable series up to the present, we performed a reconstruction procedure to create new series updated from near-complete series that covered different periods. As this procedure was a key issue for the creation of the database, the reconstruction process was done manually with the aid of a geographical information system in which the location, the data period and the topography were available. This was a unique step in the creation of the database. Although an automatic procedure would be desirable for performing this step, we were unable to find an optimal approach that allowed merging of the series. The topographical diversity (including elevation, topographic barriers, different atmospheric influences among neighboring valleys and other factors), and the need to avoid redundant information, necessitated use of a manual process for this reconstruction.

Long series from observatories without data up to 2000–2002 were assigned to observatories with data updated to 2000–2002, which were located in the same or nearby municipalities (always within a radius of 15 km) and had similar topographic conditions (less than 100 m difference in elevation), using the nearest neighbor method. Those series that finished before 2000 and lacked neighbors meeting these criteria were eliminated from the final dataset. When daily information was coincident among two or more observatories for the same day, data of the observatories containing data up to 2000–2002 were preferred. The spatial location of the reconstructed series was assigned to the location of the observatory with data updated to 2000–2002.

The result of this manual reconstruction process was 934 observatories with complete records until 2000–2002. Therefore, of the 3106 original observatories, 2172 were used in the reconstruction process and thereafter discarded according to the criteria described above. Of the 934 observatories with complete records, 383 (41%) were reconstructed or combined with other observatories to provide data covering more than 20 years; 229 (24.5%) were reconstructed with 5–20 years of data; and 322 (34.5%) had data gaps less than 5 years. The spatial distribution of the reconstructed series (Fig. 5) showed a homogeneous distribution in the study area, although a higher density of data series was present in some areas, such as in the south of the Lérida, Barcelona and Castellón provinces, and the center of Huesca and Álava provinces. The spatial density was lower than in the original series (170.7 km<sup>2</sup> per observatory, compared to 51.3 km<sup>2</sup> for the original dataset). All the provinces are well covered, and the presence of bias in the spatial distribution of the observatories used in reconstructions was discounted.

The number of data series available varied with respect to the starting date. Only two series were available with complete data from 1901. For 1920 onwards there were 68 series, and after 1940 there was a progressive increase from 206 series to a maximum of 934 in 1988. After 2000 there was a decrease in the number of available series until 2002 (December 2002 = 834 series).

### ***3.2. Quality control***

The objective of quality control was to identify erroneous or questionable records in the climate datasets. Errors in climate series are a common problem arising from sources such as the condition of instruments, and in data processing including collection, transcription and digitizing (Reek et al., 1992). As a consequence of the simplicity of rain gauges, very few instrument errors were expected in the databases used for this work. Nevertheless, errors arising from the data processing are equivalent to other climatic variables such as temperature.

Several criteria have been proposed for identifying erroneous data in climate series. As most errors are due to outliers (exceptionally high or low spurious values), some authors identify questionable values from a fixed threshold derived from the average and standard deviation of the series. The data of the previous and following day can also be used to identify anomalous spikes in variables in a series, such as temperature (Gleason, 2002). However, this approach cannot be applied to precipitation series due to its high temporal and spatial variability. A better approach to identifying outliers is comparison of daily values among neighboring observatories. Several procedures can be followed for this spatial comparison. Feng et al. (2004) used linear regression for daily precipitation data among each observatory and the 5 most correlated observatories. The regression residuals were used to identify questionable values. Griffiths et al. (2003) identified major outliers in the rainfall records, and manually assessed whether the values were related to real meteorological events such as flooding or heavy rainfall events. They also checked for internal consistency against records from nearby localities. A similar manual approach was followed by Manton et al. (2001) in the South Pacific, Brunetti et al. (2001) in Italy, and Aguilar et al. (2005) in Central America.

Due to the high density of data available, we were able to follow an approach based on comparison of the rank of each data record with the average rank of the data recorded in adjacent observatories. The original daily precipitation series were converted to percentiles, after eliminating the zero values; this accounted for more than 60% of the data. Each precipitation value was replaced by its

corresponding percentile, according to the complete series. After transformation the zero values were assigned a zero percentile. For each data series we selected the observatories located within a radius of 20 km, and set a criterion of a minimum of four observatories as a condition for performing the test. Where this criterion was not met, the daily value of the target observatory was not compared.

Only records above the 99th percentile were checked in the first stage. The maximum allowed difference between a candidate observation and the average values of the percentiles in the neighboring observatories was set at 60 percentile units. If the difference was higher than this the candidate observation was considered questionable. These values were flagged and substituted with data from the closest series. In the second stage, the records below the 99th percentile were compared to the average of the neighboring series. In this case a difference of 70 percentile units was set as the threshold for identifying questionable data, and values exceeding this were flagged and substituted with data from the closest observatory.

Another common source of error is the inclusion of false zero values (Viney and Bates, 2004). Hence, zero values coinciding with substantial precipitation in nearby observatories were flagged following a similar approach, and if the average percentile in the neighboring observatories was higher than 50, the zero value of the target observatory was substituted with data from the closest observatory.

The thresholds described above were set after analyzing a number of series and showing that these values optimized the identification of questionable data. Different thresholds may be required in places where the climate characteristics are different.

Following Reek et al. (1992), we also checked series for the occurrence of identical values (excluding zeros) on at least seven consecutive days. These data were also flagged and substituted.

On average the proportion of data substituted using the above criteria was 0.1% in each observatory (range 0–1.04%) (Fig. 6). Only 47 series from a total of 934 had more than 0.4% of the data

replaced. The highest proportion of data rejected in any one series using the described process was 1.04% (Fig. 6). Most of the replacements (63.8%) corresponded to zero values (Fig. 6). These values are similar to those reported in other studies. For example, Feng et al. (2004) found an average of 0.03% of the data questionable, while Reek et al. (1992) reported a figure of 0.04%.

As the methodology described can affect the probability distribution of the most extreme records in a series, a test was performed using standard methods for extreme value analysis. For this purpose we calculated the L-coefficients of skewness and kurtosis of the data series before and after the quality control process. Partial duration series (PDS) or series of peaks over a threshold were extracted in order to isolate only the extreme values (Beguiría, 2005). Given a precipitation series  $X = \{x_1, x_2, \dots, x_n\}$ , where  $x_n$  is the observation on a given day, the PDS  $Y = \{y_1, y_2, \dots, y_j\}$  consists of the exceedences of the original series over a predetermined threshold,  $x_0$ :

$$y_j = x_i - x_0 \quad \forall \quad x_i > x_0 .$$

Therefore, the size of the series obtained depends on the value of the threshold,  $x_0$ . For each series, the values corresponding to the 90th and 95th percentiles before and after the quality control process were used as thresholds for constructing the PDS.

The L-coefficients of skewness ( $\tau_3$ ) and kurtosis ( $\tau_4$ ) were calculated as follows:

$$\tau_3 = \frac{\lambda_3}{\lambda_2}$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2},$$

where  $\lambda_2$ ,  $\lambda_3$  and  $\lambda_4$  are the L-moments of the PDS series. These were obtained from the probability-weighted moments (PWMs) of the series, using the formulae:

$$\lambda_1 = \alpha_0$$

$$\lambda_2 = \alpha_0 - 2\alpha_1$$



$$\lambda_3 = \alpha_0 - 6\alpha_1 + 6\alpha_2$$

$$\lambda_4 = \alpha_0 - 12\alpha_1 + 30\alpha_2 - 20\alpha_3 .$$

The PWMs of order  $s$  were calculated as:

$$\alpha_s = \frac{1}{N} \sum_{i=1}^N (1 - F_i)^s x_i ,$$

where  $F_i$  is an empirical frequency estimator corresponding to the data  $x_i$ .  $F_i$  was calculated following Hosking (1990):

$$F_i = \frac{i - 0.35}{N} ,$$

where  $i$  is the range of  $x_i$  in the PDS arranged in ascending order, and  $N$  is the number of data records.

We found that the relationship between the values of  $\tau_3$  and  $\tau_4$  before and after the quality control process was approximately linear, and noticeable changes were observed in only a few series (Fig. 7). This provides evidence that the quality control process did not significantly affect the statistical characteristics of the extremes, with the exception of a few observatories that had greater differences from the surrounding series.

### **3.3. Homogeneity testing**

A common problem in climate data series is the presence of inhomogeneities. The majority of these appear as abrupt changes in the average values, but also appear as changes in the trend of the series (Alexandersson and Moberg, 1997). Inhomogeneities in climate series can result in substantial misinterpretation of the behavior and evolution of climate. Inhomogeneities can arise from human causes such as changes in the location of the observation station, alteration of the surrounding environment, observer changes and instrumental replacement (Karl and Williams, 1987). Accumulation of daily precipitation over several days is another important problem that can introduce inhomogeneity into daily precipitation series (Viney and Bates, 2004), and the

reconstruction of time series through the union of two or more series (the approach followed in this study) is a common source of inhomogeneities (Lanzante, 1996; Peterson et al., 1998). If a series is identified as non-homogeneous, use of the data for trend and variability analysis becomes questionable, and it is usually discarded.

A variety of methods have been developed to identify inhomogeneities in climate data series (see reviews in Peterson et al., 1998 and Beaulieu et al., 2007). There are two general types of homogenization procedure: i) absolute, which considers only the information in the time series being tested; and ii) relative, in which data from other observatories are also used. The latter procedure is more reliable as it involves comparison of the temporal evolution of a candidate series with that of a reference series created from correlated series nearby.

The majority of methods are focused on monthly, seasonal and annual data (Peterson et al., 1998). There is no standard approach for daily precipitation series because of the high spatial and temporal variability of this variable, and the difficulties in correcting the series if inhomogeneities are found. For this reason, the homogeneity tests applied to daily precipitation series can only identify the temporal inhomogeneities in the series enabling elimination of the periods and/or series which are not homogeneous.

Given the lack of methods for directly testing the homogeneity of daily precipitation series, the most common approach is to apply the techniques used for monthly precipitation series, after transformation of the daily series to a monthly equivalent (e.g. Brunetti et al., 2000; Feng et al., 2004; Lana et al., 2004; Schmidli and Frei, 2005; Tolika et al., 2007).

While this approach is valid only if the volume of precipitation is analyzed, inhomogeneities in daily precipitation can be much more complex, as inhomogeneities can affect other parameters. For example, attributing multi-day rainfall accumulations to a single day is a common problem in daily data series (Viney and Bates, 2004). This practice reduces the number of rainy days and increases the average precipitation per rainy day, and may cause significant changes in the recorded

frequency distribution of daily precipitation series, and the length of dry and wet spells. For series with many multi-day accumulations, Suppiah and Hennessy (1996) found an effect on temporal trends in percentiles when accumulations were either distributed or ignored. Changes in the observation protocol (e.g. through a change of observer) can produce inhomogeneities in the frequency of rainy days without causing an inhomogeneity in the monthly precipitation record. Therefore, there is a need to test the precipitation volume, and also the precipitation frequency and intensity. Wijngaard et al. (2003) tested the homogeneity of daily precipitation records by means of wet day count series rather than precipitation amounts. They argued that wet day counts have lower variability than series comprising annual amounts, and hence the former facilitate easier detection of inhomogeneities.

In this study we tested the homogeneity of the reconstructed and quality-controlled daily precipitation series using four monthly parameters: i) monthly precipitation amount, ii) monthly average number of rainy days above 1 mm, iii) monthly maximum precipitation, and iv) number of days above the 99.5th percentile. Following Wijngaard et al (2001), we adopted a 1 mm threshold for the second criterion because using a lower threshold (e.g. any precipitation) usually leads to a high rate of false inhomogeneities, caused solely by errors in measuring very low amounts. Calculation of the 4th criterion at a monthly time scale would have yielded a sequence of mostly zeros, but as we describe below, the homogeneity testing was performed seasonally and annually. As homogeneity testing was performed using averages of long time periods, this approach helped to identify changes in the frequency of the most extreme precipitation events, which could have been due to inhomogeneities in the dataset.

Of the several methods for detecting inhomogeneities in climate series, we used the standard normal homogeneity test (SNHT) developed by Alexandersson (1986) for single breaks. This is the most widely used test for detecting inhomogeneities in climate series (e.g. Keiser and Griffiths, 1997; Moberg and Bergstrom, 1997; González-Hidalgo et al., 2002). Various comparative studies of

interpolation methods have shown that this method is better than other approaches, and facilitates detection of small breaks and multiple breaks in a series (Easterling and Peterson, 1992; Ducre-Robitaille et al., 2003).

As the reliability of inhomogeneity detection increased through the use of relative homogeneity methods based on information from neighboring stations, we calculated reference series for each observatory. Although a single neighboring series of good quality can be used as a reference (Keiser and Griffiths, 1997), it is very difficult to ensure that a series to be used as a reference to other series is completely homogeneous. In this study we used the approach of Peterson and Easterling (1994), as modified by González-Hidalgo et al. (2004), which uses several neighboring stations to create a reference series for each of the four parameters analyzed. The probability of inhomogeneities is therefore minimized, since all the series are considered as a whole.

To create the reference series we considered all the observatories within a radius of 50 km from the candidate observatory, according to:

$$P_{R,i} = \frac{\sum_{x=1}^n P_{x,i} \cdot w_x}{\sum_{x=1}^n w_x},$$

where  $P_{R,i}$  is the observation for the reference series in month  $i$ ,  $P_{x,i}$  is the observation at observatory  $x$  in month  $i$ , and  $w_x$  is a weighting factor. Peterson and Easterling (1994) used the coefficient of correlation between the candidate series and each surrounding series as the weighting factor. However, they considered that the presence of discontinuities in the series could alter the coefficients of correlation, so they calculated the correlation from the series of differences according to:

$$D_i = P_{i+1} - P_i,$$

where  $D$  is the difference between the two series for month  $i$ , and  $P$  is the observation corresponding to month  $i$ .

Correlations were calculated using monthly precipitation series; hence 12 coefficients of correlation were obtained for each observatory. We discarded those observatories with any month having a correlation coefficient lower than 0.6. Finally, the weighting factor used for each observatory was the average of the correlation coefficients obtained for the 12 monthly series. The ProClimDB software (Štěpánek, 2007a) was used to automate calculation of the 3736 reference series (4 for each observatory).

The AnClim software (Štěpánek, 2007b) was used in the application of the SNHT to each observatory and parameter. The test was applied to seasonal and annual series of the four parameters, since this approach yields better results than using only monthly series. For each seasonal and annual series a  $T$  series was obtained using the SNHT. If the value of  $T$  in each month exceeded a certain threshold, the series was flagged as inhomogeneous. The threshold  $T$  value can be set to any given confidence level ( $\alpha$ ), and in this study a value of  $\alpha = 0.05$  was used (see values in Alexandersson and Moberg, 1997). As a consequence of the substantial length of some climate series, some short inhomogeneous periods could be hidden after testing. To avoid this problem a sequential splitting procedure was applied after each 30 years of data, to detect short inhomogeneous periods (Štěpánek, 2004).

As a consequence of the large quantity of information, we established an automatic criterion to accept or reject inhomogeneities. Flagged inhomogeneities were accepted only when they appeared in the annual series and a minimum of two seasonal series. Since the temporal location of the inhomogeneities can vary within a range of some years, a maximum difference of eight years was allowed between inhomogeneities found in the annual and seasonal series. Those data series with two or more inhomogeneities were removed from the dataset. In series in which one unique inhomogeneity was found, the period prior to the inhomogeneity was also removed, as correcting inhomogeneous periods in daily precipitation records is exceedingly difficult. We preferred to lose some information but retain the remaining high quality data for subsequent climatic studies.

As explained above, we tested for series homogeneity using four variables. Firstly we tested the homogeneity in the series of precipitation amounts. After removing the inhomogeneous series and periods, the remaining series were tested for inhomogeneities in the number of rainy days, and subsequently for inhomogeneities in the maximum values and the number of events above the 99.5 percentile. A total of 260 inhomogeneous series were found using monthly precipitation amounts, and 74 of these were discarded because they contained two or more inhomogeneities (Table 1). A total of 157 inhomogeneous series were found at the second step, and 32 of these were discarded. Finally, 25 inhomogeneities were found in the extreme series, but no series were discarded. At the completion of the entire process, 407 series were found to be inhomogeneous, corresponding to 43.6% of the total series. For 301 series, the period prior to the inhomogeneity was deleted, and 106 series were completely discarded because they had two or more temporal inhomogeneities. We also analyzed the impact of data reconstruction on the homogeneity of the series, to determine if inhomogeneities were introduced in the process of reconstruction and data filling using data from nearest neighbors. Table 1 shows the number of inhomogeneous series detected in series with more than 20 years of reconstructed data, stations with a reconstructed period of 5–20 years, and series having less than 5 years with data gaps. Of the 383 stations with reconstructions exceeding 20 years, 192 were inhomogeneous. Of the 229 stations with reconstructions of 5–20 years, 113 were inhomogeneous, and of the 322 non-reconstructed series filled only for periods less than 5 years, 137 series were inhomogeneous. As expected, a larger percentage of inhomogeneous series was found for long reconstructions (50.1% of total) than for short reconstructions. However, there were no large differences between the series with reconstructions shorter than 20 years and the non-reconstructed series (49.3 and 42.5% of the total, respectively). This result indicates that the data gap filling process can introduce inhomogeneities to the series. Nevertheless, with careful selection of neighbors using restrictive distance and correlation criteria, the number of artificial

inhomogeneities added during the process could be minimized. In any case, these inhomogeneities were identified and eliminated during the four-step homogeneity testing.

The time evolution of the number of inhomogeneities showed an irregular distribution (Fig. 8). A higher number of inhomogeneities was found in the decades of the 1920s and 1930s, but also in the decade of the 1960s. Nevertheless, with the exception of the year 1968, the number of inhomogeneities per year from 1960 to 1990 was very regular, affecting 1–2% of the available series.

In general we found that for some significant inhomogeneities detected in monthly totals in some years it was very difficult to decide whether the temporal inhomogeneity in the series was real, and to which year it should be attributed. However, if seasonal and annual data were incorporated, inhomogeneities were more apparent. Therefore, we used the daily precipitation series aggregated in seasons and years to identify temporal inhomogeneities, which could be recognized in the variation of the  $T$  values (see the example in Fig. 9). Fig. 10 shows the results of applying the SNHT to the l'Ametlla de Mar series using the seasonal and annual precipitation amount series, and also the series of number of rainy days. This example shows the situation found in some observatories, whereby no inhomogeneities were found in the series of precipitation amount, but significant inhomogeneities were found in the series of the number of rainy days (the middle of the 1970s in this example). Examination of the time evolution of the series and the associated  $T$  statistic did not provide any extra information. Comparison with the reference series clearly showed an accumulation of the same precipitation amounts in fewer days during the period 1950–1974. After this period the number of rainy days was again very similar to the reference series. This error is explained by the attribution of several days of rainfall to a single day, which could affect the frequency distribution of precipitation events, the average precipitation per event, and the duration of dry and wet spells, as discussed earlier. Therefore, the inhomogeneous period prior to the

detected break was removed from the final dataset. This inhomogeneity would not have been identified if only monthly amounts had been used.

A final example shows a case in which an inhomogeneity was found in the series of monthly maxima for 1961 (Fig. 15). However, it was not detected in the series of precipitation amount or number of rainy days, and was attributed to errors in the rain gauge recording of the most extreme events.

#### **4. Results**

In this section we present some results of comparisons of the characteristics of the database at various stages of the process. The usefulness of the database for several types of climate analysis is discussed. Some basic analyses were also performed on the spatial distribution of precipitation in the study area, to illustrate the improved performance of the final database relative to earlier stages.

As a tradeoff in establishing strict quality control criteria, a reduction in the spatial density of data occurred (Fig. 12). Of the original 3106 series, the final database consisted of 828 series covering the study area comprehensively. Some areas had lower data density, including the Pyrenean Range in the Lérida province, the north of the Castellón province, and the central areas of the Burgos province.

The availability of data in the database varied greatly as a function of the length of the data series (Fig. 13). For example, after the reconstruction process a total of 207 series starting in 1940 were available. After homogeneity testing this was reduced to 117 series (i.e. 56%). More importantly, there was a decrease from 471 to 291 series in the 1960s. Although the amount of data discarded in the homogenization process was large, the remaining series were adequate in terms of the quality and temporal homogeneity of the database. The spatial coverage of the database was also affected by variation in the length of the series (Fig. 14). Although the number of series with records since 1920 was very satisfactory (34), the majority of these were located in the Cataluña (24) and Aragón (8) regions. For series with data since 1935 (92) there was an increase in the number of provinces



represented, but the data density remained higher in the Cataluña region. This situation reduces the usefulness of the database for reliable spatial analysis, although long-term studies focused in the regions with good data density can still be undertaken. The number of series available from 1950 increased noticeably to 190, providing good coverage of the study area at adequate density. This temporal coverage is sufficient for trend analysis, considering the spatial variation in this factor. From 1965 a total of 331 series were available, providing very good data density with few spatial gaps. This high spatial density enables detailed spatial analysis of precipitation fields, and the recording period is long enough for extreme events analysis.

To assess how the quality control and homogeneity testing procedures affected the final database, we undertook several analyses using the series available between 1960 and 2000. The variables analyzed were: i) average precipitation per rainy day, ii) average duration of wet spells, iii) average duration of dry spells, iv) number of days with precipitation  $> 0$  mm, and v) number of days with precipitation greater than 75 mm (extreme events). Preliminary statistical analysis showed that, in general, the range and variability of the precipitation parameters analyzed were reduced after the quality control process and homogeneity test, but the average values remained similar. This does not necessarily mean that the spatial average over large areas is unbiased. However, among close observatories the precipitation differences (over a few days or long periods) caused by erroneous measurements or inhomogeneities were removed during the process, making the data more consistent. The procedure avoids spurious results in temporal analysis of precipitation records, but can also reveal spurious differences among neighboring observatories in future climate studies. We performed a spatial analysis of the dataset at various steps in the process to assess their influence on spatial coherence of the precipitation parameters. For this purpose we used semivariance plots from daily precipitation series, which represent the spatial self-correlation of a variable (Fig. 15). The semivariance is half the variance of the differences between all possible points within a given distance. It is normally assumed that the semivariance increases with increasing distance lag, in

accordance with the basic geostatistical principle that closer objects are more similar. We additionally considered that high-quality series will have a lower semivariance than poor-quality series, especially over shorter distance lags. Both hypotheses were confirmed in our test. For all precipitation parameters the semivariance increased as a function of the distance lag, and the homogenous database consistently had a lower semivariance, especially with distance lags less than 25 km. However, the difference among databases was less evident for some precipitation parameters, including the dry spell duration. The results of this analysis support the view that the spatial coherence of the homogeneous dataset is better than in the original and intermediate stages. This was also evident when correlation coefficients were calculated between pairs of daily precipitation observatories separated by various distances (spatial self-correlogram; Fig. 16). The relationship among neighboring series was noticeably improved after the homogeneity testing process, as shown by higher R-coefficients. This was a consequence of removal of those data series that differed markedly from adjacent series. Therefore, we believe that the quality control and homogeneity testing processes described in this paper improved the quality and spatial coherence of the dataset, especially in relation to future spatial and temporal regional analyses.

## **5. Conclusions**

We have described a process for creating a spatially dense database of daily precipitation in northeast Spain. The main contribution of the research relates to effective use of available data through a combined process of reconstruction, quality control and homogeneity testing. This is significant as very few examples exist of the construction of quality-controlled databases with daily time resolution and regional coverage. The usual approach has been to select long-term and reliable series, and to discard fragmented or short-term data. This results in a significant loss of information, and detrimentally affects the spatial density of the data, reducing the usefulness of the dataset for spatially explicit analyses. Our approach involved reconstruction of spatially close data series to generate new and unique series. This substantially reduced the loss of data involved in the

alternative procedure, but introduced the risk of creating spurious and inconsistent data series. To ensure the quality of the final database, the reconstructed series were subjected to quality control and homogeneity testing processes, consisting of several stages. Analyses of the final database included single site and spatial analyses, and confirmed its coherence.

The methodology described in this paper can be readily adapted for use worldwide with other databases having similar characteristics (high spatial density, daily temporal resolution), particularly in areas where long-term precipitation series are rare and fragmented series covering different periods in the same locality are common, as in Spain. Although the methodology involves several steps, the data filling and homogeneity testing processes are completely automated, and can be adapted to other regions without modification. For quality control purposes, some decisions must be made in advance, such as the optimum threshold values. It is likely that the values used in this work would yield good results with other datasets, but this should be assessed by users through a heuristic process. The only manual procedure in creating the database was during selection of the neighboring series to be merged in the reconstruction process. Although an automated process could have been used, with the aid of a geographical information system, we chose to oversee this crucial step directly. The experience of the researcher and a good knowledge of the regional climate could be very useful in developing an automated process, based on the distance between observatories and using other auxiliary information layers, such as elevation. In this paper we have detailed the decisions taken to construct a quality dataset in relation to problems which are typically encountered. It should not be assumed that all the decisions made have universal validity, as it is possible that some of the threshold values or the procedures adopted will need to be modified for application to other datasets, or according to the specific purposes of the research. We encourage other researchers to make similar assessments prior to undertaking any climate study, and we urge discussion of the specific problems associated with the use of high frequency climate databases.

The database described on this paper is available for use by scientists for research purposes. Anyone interested in using the data is encouraged to contact the authors at the following email addresses: sbegueria@eead.csic.es and svicen@ipe.csic.es.

## Acknowledgements

We would like to thank the National Institute of Meteorology (Spain) for providing the precipitation data base used in this study. This work has been supported by the following projects: CGL2005-04508/BOS (financed by the Spanish Commission of Science and Technology and FEDER), PIP176/2005 (financed by the Aragón Government), and “Programa de grupos de investigaciónlower consolidados”(BOA 48 of 20-04-2005), also financed by the Aragón Government. Authors would like to thank to Dr. José C. González-Hidalgo and José M. García-Ruiz for their helpful comments.

## References

- Aguilar, E., Peterson, T.C., Ramírez-Obando, P., Frutos, R., Retana, J.A., Solera, M., Soley, J., González García, I., Araujo, R.M., Rosa Santos, A., Valle, V.E., Brunet, M., Aguilar, L., Álvarez, L., Bautista, M., Castañón, M., Herrera, L., Ruano, E., Sinay, J.J., Sánchez, E., Hernández Oviedo, G.I., Obed, F., Salgado, J.E., Vázquez, J.L., Baca, M., Gutiérrez, M., Centella, M., Espinosa, J., Martínez, D., Olmedo, B., Ojeda Espinoza, C.E., Núñez, R., M. Haylock, M., Benavides, H. and Mayorga, R., (2005): Changes in precipitation and temperature extremes in Central America and northern South America, 1961–2003. *Journal of Geophysical Research-Atmosphere*, 110, D23107, doi:10.1029/2005JD006119.
- Alexandersson, H., (1986): A homogeneity test applied to precipitation data. *Journal of Climatology*, 6: 661-675.
- Alexandersson, H. and Moberg, A., (1997): Homogenization of Swedish temperature data. Part I: Homogeneity test for lineal trends. *International Journal of Climatology*. 17: 25-34.
- Allen, R.J. and DeGaetano, A.T., (2001): Estimating missing daily temperature extremes using an optimized regression approach. *International Journal of Climatology*, 21: 1305-1319.
- Auer, I., Reinhard Böhm, Anita Jurkovi, Alexander Orlik, Roland Potzmann, Wolfgang Schöner, Markus Ungersböck, Michele Brunetti, Teresa Nanni, Maurizio Maugeri, Keith Briffa, Phil Jones, Dimitrios Efthymiadis, Olivier Mestre, Jean-Marc Moisselin, Michael Begert, Rudolf Brazdil, Oliver Bochnicek, Tanja Cegnar, Marjana Gaji-apka, Ksenija Zaninovi, eljko Majstorovi, Sándor Szalai, Tamás Szentimrey, Luca Mercalli (2005): A new instrumental precipitation dataset for the greater alpine region for the period 1800-2002. *International Journal of Climatology*, 25: 139-166.
- Beaulieu, C., Ouarda, T, and Seidou, O. (2007): A review of homogenization techniques for climate data and their applicability to precipitation series. *Hydrological Sciences Journal-Journal des Sciences Hydrologiques* 52: 18-37.

- Beguiría, S. (2005): Uncertainties in partial duration series modelling of extremes related to the choice of the threshold value. *Journal of Hydrology*, 303(1-4): 215-230.
- Boroneant, C., Plaut, G., Giorgi, F., and Bi, X., (2006): Extreme precipitation over the Maritime Alps and associated weather regimes simulated by a regional climate model: Present-day and future climate scenarios, *Theoretical and Applied Climatology*, 86: 81-99.
- Brandsma, T., and Können, G.P., (2006): Application of nearest-neighbor resampling for homogenizing temperature records on a daily to sub-daily level. *International Journal of Climatology* 26: 75-89.
- Bruce, J.P., (1994): Natural disaster reduction and global change. *Bulletin of the American Meteorological Society*. 75: 1831-1835.
- Brunet, M., Saladié, O., Jones, P., Sigró, J., Aguilar, E., Moberg, A., Lister, D., Walther, A., Lopez, D. and Almarza, C., (2003): The development of a new dataset of Spanish Daily Adjusted Temperature Series (SDATS). *International Journal of Climatology*, 26: 1777-1802.
- Brunetti, M., Buffoni, L., Maugeri, M. and Nanni, T., (2000): Precipitation intensity trends in northern Italy. *International Journal of Climatology*, 20: 1017-1031.
- Brunetti, M., Colacino, M., Maugeri, M. and Nanni, T., (2001): Trends in the daily intensity of precipitation in Italy from 1951 to 1996. *International Journal of Climatology*, 21: 299-316.
- Brunetti, M., Maugeri, M., Monti, F. and Nanni, T., (2006): Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series. *International Journal of Climatology*, 26: 345-381.
- Ducré-Robitaille, J.F., Vincent, L.A. and Boulet, G., (2003): Comparison of techniques for detection of discontinuities in temperature series. *International Journal of Climatology*, 23: 1087-1101.
- Easterling, D.R. and Peterson, T.C., (1992): Techniques for detecting and adjusting for artificial discontinuities in climatological time series: a review. *Proc. Fifth Int. Meeting on Statistical Climatology* (Toronto, Canada). J28-J32. Steering Comitee for International Meetings of Statistical Climatology.
- Easterling, D.R., Meehl, G.A., Parmesan, C., Changnon, S.A., Karl, T.R., and Hearn, L.O., (2000): Climate extremes: observations, modelling and impacts. *Science*, 289: 2068-2074.
- Eischeid, J.K., Pasteris, P.A., Diaz, H.F., Plantico, M.R. and Lott, N.J., (2000): Creating a serially complete, national daily timeseries of temperature and precipitation for the western United States. *Journal of Applied Meteorology*, 39: 1580-1591.
- Feng, S., Hu, Q., Qian, W., (2004): Quality control of daily meteorological data in China, 1951-2000: a new dataset. *International Journal of Climatology*, 24: 853-870.
- Frei, C., Schöll, R., Fukutome, S., Schmidli, J. and Vidale, P., (2006): Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *Journal of Geophysical Research-Atmosphere*, Volume 111, Issue D6, CiteID D06105
- Gleason, E., (2002): Global daily climatology network, V1.0. National Climatic Data Center, 151 Patton Ave. Asheville, NC.
- González-Hidalgo, J.C., De Luis, M., Stepanek, P., Raventós, J. y Cuadrat, J.M., (2002): Reconstrucción, estabilidad y procesos de homogeneizado de series de precipitación en ambientes de elevada variabilidad pluvial. In *La información climática como herramienta de gestión ambiental* (Cuadrat, J.M, Vicente, S.M. and Saz, M.A Eds.): 47-57. Zaragoza.
- González Hidalgo J.C., De Luis M., Vicente Serrano S.M, Saz M.A., Štěpánek P., Raventós J., Cuadrat J.M., Creus J.M. and Ferraz J.A., (2004): Monthly Rainfall Data Base for Mediterranean coast of Spain: reconstruction and quality control. World Meteorological organization. *TD No 1236*. Geneve: 105-116.

- González-Rouco, J. F., Jiménez, J.L., Quesada, V. and Valero, F., (2001), Quality control and homogeneity of precipitation data in the southwest of Europe, *Journal of Climate*, 14: 964–978.
- Griffiths, G.M., Salinger, M.J., and Leleu, I., (2003): Trends in extreme daily rainfall across the South Pacific and relationship to the South Pacific Convergence Zone. *International Journal of Climatology*, 23: 847-869.
- Haylock, M. and Nicholls, N., (2000): Trends in extreme rainfall indices for an updated high quality data set for Australia, 1910-1998. *International Journal of Climatology*, 20: 1533-1541.
- Hosking, J.R.M., (1990): L-Moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of Royal Statistical Society B*, 52: 105-124.
- Hundecha, Y. and Bardossy, A. (2005): Trends in daily precipitation and temperature extremes across western Germany in the second half of the 20th century. *International Journal of Climatology*, 25: 1189-1202.
- Karl, T.R., and Williams, C.N., (1987): An approach to adjusting climatological time series for discontinuous inhomogeneities. *Journal of Climate and Applied Meteorology*. 26: 1744-1763.
- Karl, T.R., Knight, R.W., and Plummer, N., (1995): Trends in high-frequency climate variability in the twentieth century. *Nature*, 377: 217-220.
- Karl, T.R. and Easterling, D.R., (1999): Climate extremes: selected review and future research directions. *Climatic Change*, 42: 309-325.
- Katz, R.W. and Brown, B.G., (1992): Extreme events in a changing climate. Variability is more important than averages. *Climatic Change*, 21: 289-302.
- Keiser, D.T. and Griffiths, J.F., (1997): Problems Associated with Homogeneity Testing in Climate Variations Studies: A Case Study of Temperature in the Northern Great Plains, USA. *International Journal of Climatology*, 17: 497-510
- Klein-Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M. Pashiardis, S. Hejkrlik, L. Kern-Hansen, C. Heino, R. Bessemoulin, P. Müller-Westermeier, G. Tzanakou, M. Szalai, S. Pálsdóttir, T. Fitzgerald, D. Rubin, S. Capaldo, M. Maugeri, M. Leitass, A. Bukantis, A. Aberfeld, R. van Engelen, A. F. V. Forland, E. Mietus, M. Coelho, F. Mares C., Razuvaev, V. Nieplova, E. Cegnar, T. Antonio López, J. Dahlström, B. Moberg, A. Kirchhofer, W. Ceylan, A. Pachaliuk, O. Alexander, L. V. and Petrovic, P., (2002): Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment, *International Journal of Climatology*, 22: 1441-1453.
- Kunkel, K.E., Pielke, R.A. and Changnon, S.A., (1999): Temporal fluctuations in weather and climate extremes that cause economic and human health impacts: a review. *Bulletin of the American Meteorological Society*. 80: 1077-1098.
- Lana, X., Martínez, M.D., Serra, C. and Bargaño, A., (2004): Spatial and temporal variability of the daily rainfall regime in Catalonia (northeastern Spain), 1950-2000. *International Journal of Climatology*, 24: 613-641.
- Lanzante, J.R., (1996): Resistant, robust and non-parametric techniques for the analysis of climate data: theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology*. 16: 1197-1226.
- Manton, M.J., Della-Marta, P.M Haylock, M.R. Hennessy, K.J. Nicholls, N. Chambers, L.E. Collins, D.A. Daw, G. Finet, A. Gunawan, D. Inape, K. Isobe, H. Kestin, T.S. Lefale, P. Leyu, C.H. Lwin, T. Maitrepierre, L. Ouprasitwong, N. Page, C.M. Pahalad, J. Plummer, N. Salinger, M.J. Suppiah, R. Tran, V.L. Trewin, B. Tibig, I. and Yee D., (1998): Trends in extreme daily rainfall and temperature in Southeast Asia and the South Pacific: 1961-1998. *International Journal of Climatology*, 21: 269-284.

- Meehl, G.A., Karl, T., Easterling, D.R., Changnon, S., *et al.*, (2000): An introduction to trends in extreme weather and climate events: observations, socioeconomic impacts, terrestrial ecological impacts, and model projections. *Bulletin of the American Meteorological Society*. 81: 413-416.
- Moberg, A. and Bergstrom, H., (1997): Homogenization of Swedish temperature data. Part III: The long temperature records from Uppsala and Stockholm. *International Journal of Climatology*. 17: 667-699.
- Moberg, A. and Jones, P.D., (2005): Trenes in indices for extremes in daily temperature and precipitation in central and western Europe, 1901-99. *International Journal of Climatology*, 25: 1149-1171.
- Obasi, G.O.P., (1994): WMO's role in the international decade for natural disaster reduction. *Bulletin of the American Meteorological Society*. 75: 1655-1661.
- Pall, P., Allen, M.R., and Stone, D.A., (2007): Testing the Clausius-Clapeyron constraint on changes in extreme precipitation under CO2 warming. *Climate Dynamics*, 28: 351-363.
- Paulhus, J.L., and Kohler, M.A., (1952): Interpolation of missing precipitation records, *Monthly Weather Review*, 80: 129-133.
- Peterson, T.C. and Easterling, D.R., (1994): Creation of homogeneous composite climatological reference series. *International Journal of Climatology*. 14: 671-679.
- Peterson, T.C., Dann, H. And Jones, P., (1997): Initial selection of GCOS surface network. *Bulletin of the American Meteorological Society*, 78: 2145-2152.
- Peterson, T.C., Easterling, D.R., Karl, T.R., *et al.*, (1998): Homogeneity adjustments of in situ atmospheric climate data: a review. *International Journal of Climatology*. 18: 1493-1517.
- Peterson, T.C., Taylor, M.A., Demeritte, R., Duncombe, D.L., Burton, S., Thompson, F., Porter, A., Mercedes, M., Villegas, E., Semexant Fils, R., Klein Tank, A., Martis, A., Warner, R., Joyette, A., Mills, W., Alexander, L. and Gleason, B., (2002): Recent changes in climate extremes in the Caribbean region. *Journal of Geophysical Research – Atmospheres* 107(D21): 4601. DOI: 10.1029/2002JD002251.
- Peterson, T.C., Zhang, X., Brunet-India, M. and Vázquez-Aguirre, J.L., (2008): Changes in North American extremes derived from daily weather data. *Journal of Geophysical Research-Atmospheres*, 113, D07113, doi:10.1029/2007JD009453
- Porth, L. S., Boes, D. C., Davis, R. A. Troendle, C. A. and King, R. M., (2001): Development of a technique to determine adequate sample size using subsampling and return interval estimation. *Journal of Hydrology*, 251: 110–116.
- Reek, T., Doty, S.R. and Owen, T.W., (1992): A deterministic approach to the validation of historical daily temperature and precipitation from the cooperative network. *Bulletin of the American Meteorological Society*, 73: 753-762.
- Rodríguez, R., Llasat, M.C. and Wheeler, D., (1999): Analysis of the Barcelona precipitation series, 1850-1991. *International Journal of Climatology*. 19: 787-801.
- Romero, R., Guijarro, J. A., Ramis, C. and Alonso, S., (1998): A 30-year (1964-1993) daily rainfall data base for the Spanish Mediterranean regions: first exploratory study. *International Journal of Climatology*, 18: 541-560.
- Schmidli, J. and Frei, C., (2005): Trends of heavy precipitation and wet and dry spells in Switzerland during the 20th century. *International Journal of Climatology*, 25: 753-771.
- Smith, R. L., (1989): Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statistical Science*, 4, 367–393.
- Smith, R.L., (2003): Statistics of extremes, with applications in environment, insurance and finance. Dept. of Statistics, University of North Carolina, Chapel Hill, NC, 62 pp. [Available online at <http://www.stat.unc.edu/postscript/rs/semstatrls.pdf>.]
- Štěpánek, P. (2004): Homogenization of air temperature series in the Czech Republic during a period of instrumental measurements. In: *Fourth seminar for homogenization and quality*

- control in climatological databases* (Budapest, Hungary, 6-10 October 2003), WCDMP-No. 56. WMO, Genova. 117-133. <http://www.wmo.ch/web/wcp/wcdmp/reports/WCDMP-56.pdf>
- Štěpánek, P. (2007): *ProClimDB - software for processing climatological datasets*. CHMI, regional office Brno.
- Štěpánek, P. (2007b): *AnClim - software for time series analysis (for Windows)*. Dept. of Geography, Fac. of Natural Sciences, Masaryk University, Brno. 1.47 MB.
- Suppiah, R. and Hennessy, K.J. (1996): Trends in the intensity and frequency of heavy rainfall in tropical Australia and links with the Southern Oscillation. *Australian Meteorological Magazine*, 45: 1–17.
- Tolika, K., Maheras, P., Vafiadis, M., Flocas, H.A., and Arseni-Papadimitriou, A., (2007): Simulation of seasonal precipitation and raindays over Greece: a statistical downscaling technique based on artificial neural networks (ANNs). *International Journal of Climatology*. DOI: 10.1002/joc.1442
- Vincent, L.A., Zhang, X., Bonsal, B.R. and Hogg, W.D., (2002): Homogenization of daily temperatures over Canada. *Journal of Climate*, 15: 1322-1334.
- Viney, N.R. and Bates, B.C., (2004): It never rains on Sunday: the prevalence and implications of untagged multi-day rainfall accumulations in the Australian high quality data set. *International Journal of Climatology*, 24: 1171-1192.
- Wijngaard, J. B., Klein Tank, A. M. G. and Können, G. P., (2003): Homogeneity of 20th century European daily temperature and precipitation series. *International Journal of Climatology*, 23: 679-692.



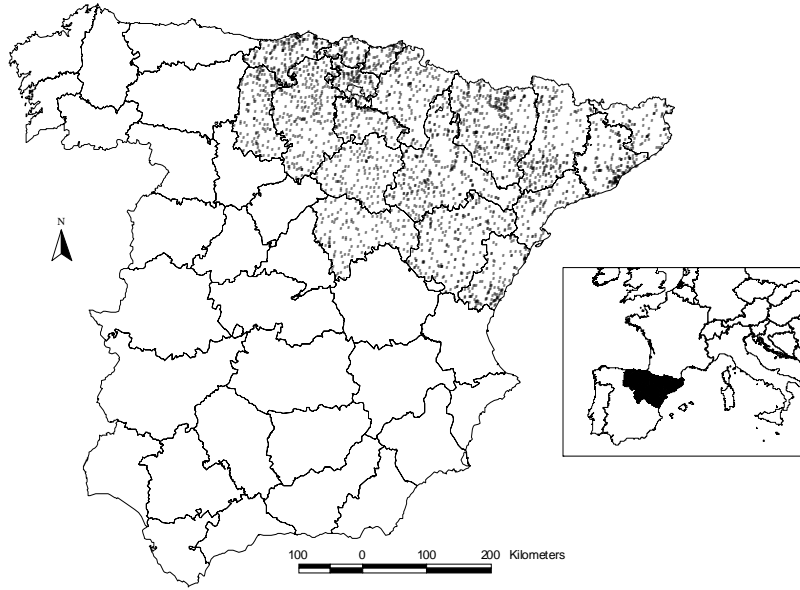


Figure 1. Location of the study area and spatial distribution of the original daily precipitation observatories.

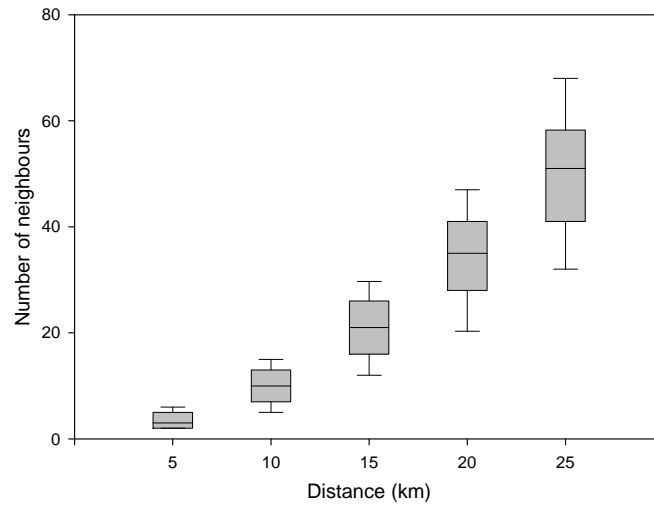


Figure 2. Box plot of the number of neighbors at different distance lags. The 10th and the 90th percentiles are shown by the lower and upper whiskers, respectively. The 25th and 75th percentiles are shown by the lower and upper limits of the boxes, respectively. The line within the box represents the median.

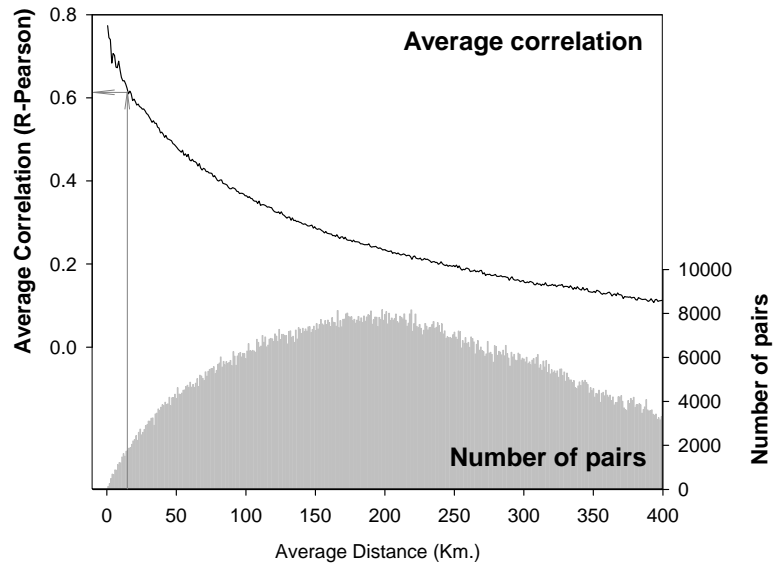


Figure 3. Average correlation between the series of daily precipitation at different distance lags for all possible pairs (left axis), and histogram of station separation (right axis). The arrows indicate the average correlation between observatories at a distance of 15 km. The bin size for the pair distance is 1 km.

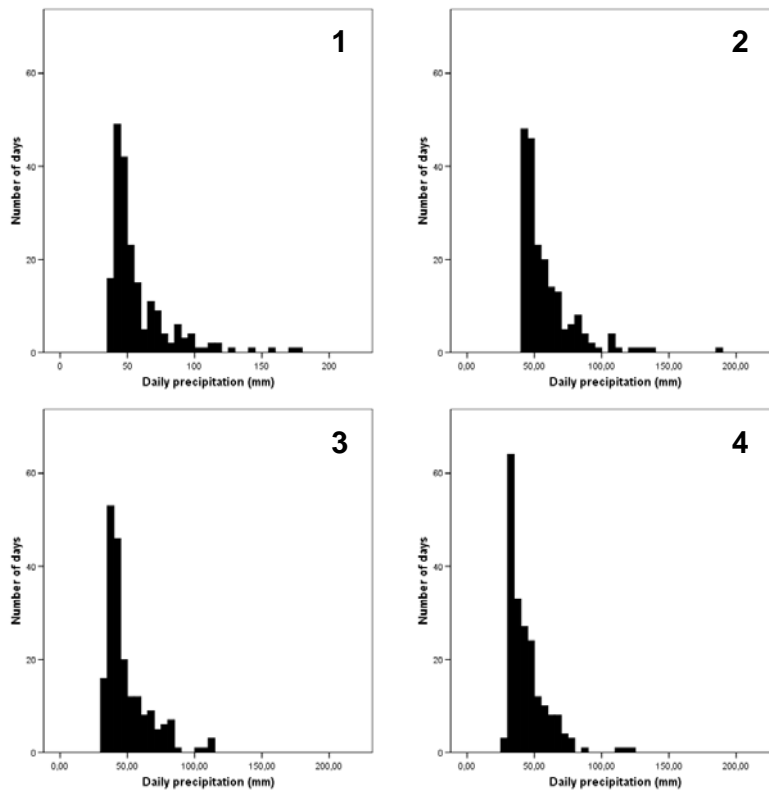


Figure 4. Histograms of the frequencies of the 200 highest records of the Fabra observatory (Barcelona). 1) original series, 2) nearest neighbor method, 3) IDW, and 4) regression method.

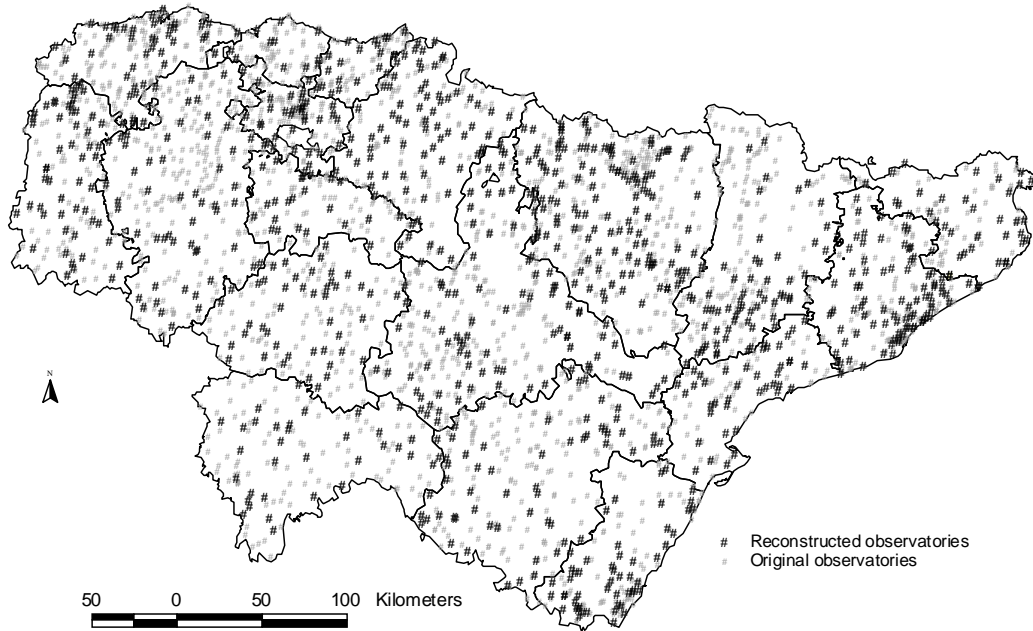


Figure 5. Location of the data series available after the reconstruction process (black). The original observatories are shown in grey.

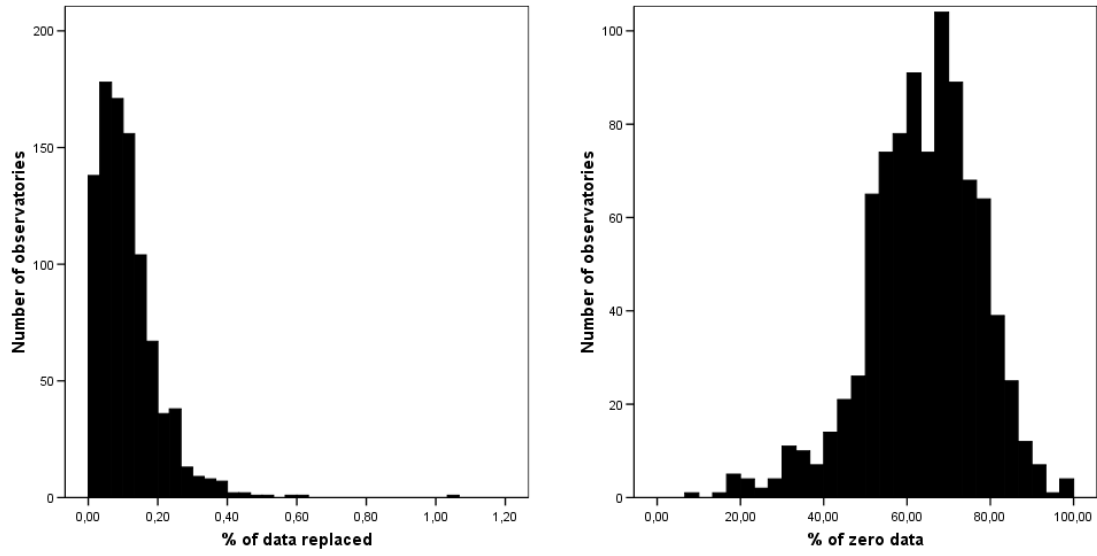


Figure 6: Frequency histograms of the data replaced in each series, and the percentage of total substituted data that corresponded to zero values.

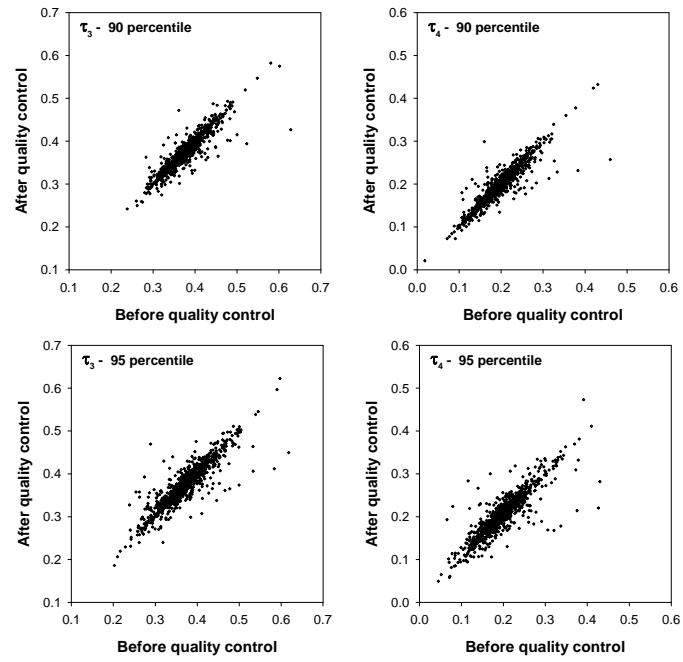


Figure 7. Relationship between the L-coefficients of skewness and kurtosis in the partial duration series at the 90th and 95th percentiles, before and after the quality control process.

HOMOGENEITY TESTING PROCEDURE	NUMBER OF INHOMOGENEITIES		SERIES ELIMINATED	
Precipitation amount	260		74	
Number of days with precipitation > 1 mm.	157		32	
Monthly maxima and number of days with precipitation above 99.5 percentile	25		0	
RECONSTRUCTED PERIOD	TOTAL	INHOMOGENEOUS	% TOTAL	% INHOMOGENEOUS
> 20 years	383	192	41.0%	43.4%
> 5 years < 20 years	229	113	24.5%	25.5%
< 5 years	322	137	34.5%	30.1%
Total	934	442	100.0%	100.0%

Table 1. Results of homogeneity testing. The number and percentage of inhomogeneous series for reconstructed and non-reconstructed series are also shown.



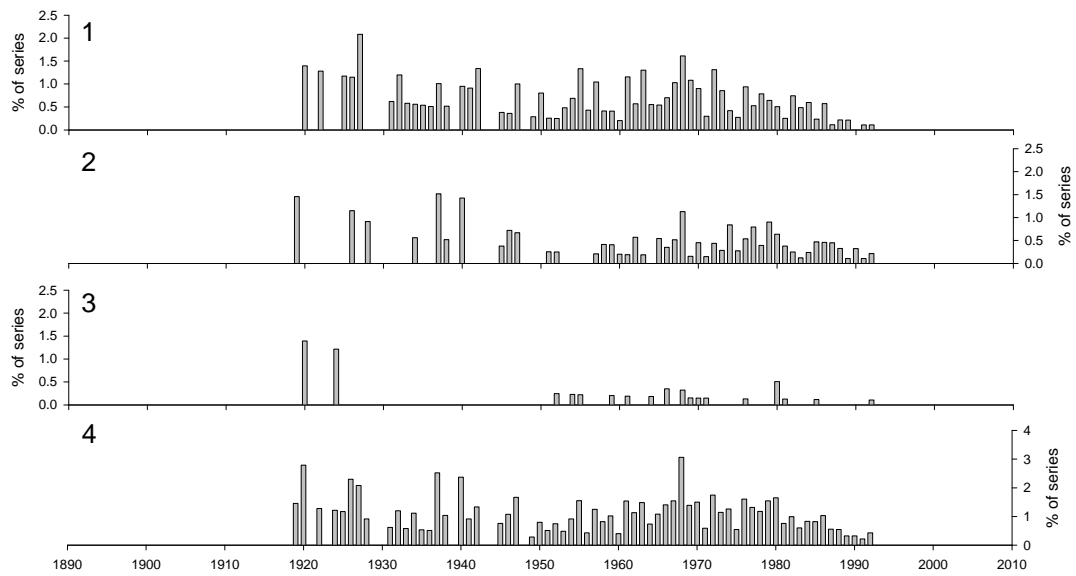


Figure 8. Percentage of inhomogeneous series with respect to the number of series available for each year: 1) precipitation amount, 2) number of rainy days, 3) monthly maximum and number of days above 99.5th percentile, 4) total.

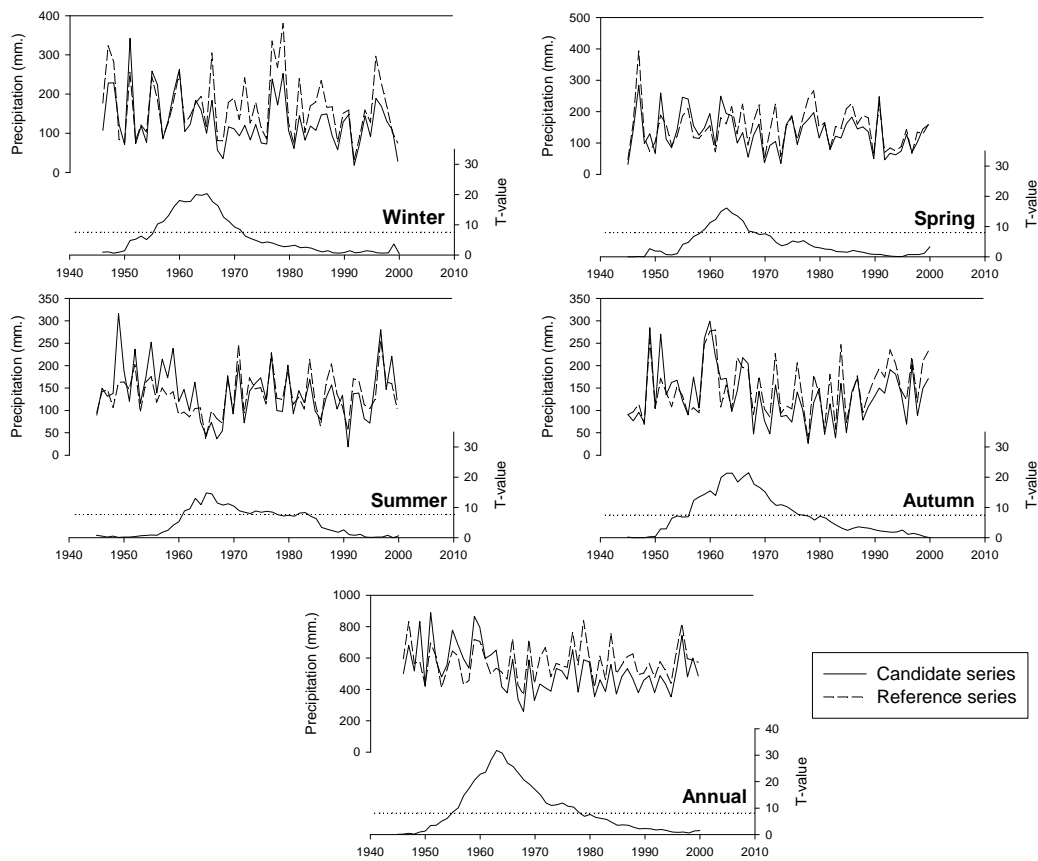


Figure 9. Seasonal and annual series of monthly precipitation amounts at the El Burgo de Osma (La Rasa) observatory. The series of T-values and the limit of confidence (dotted line) are also shown.

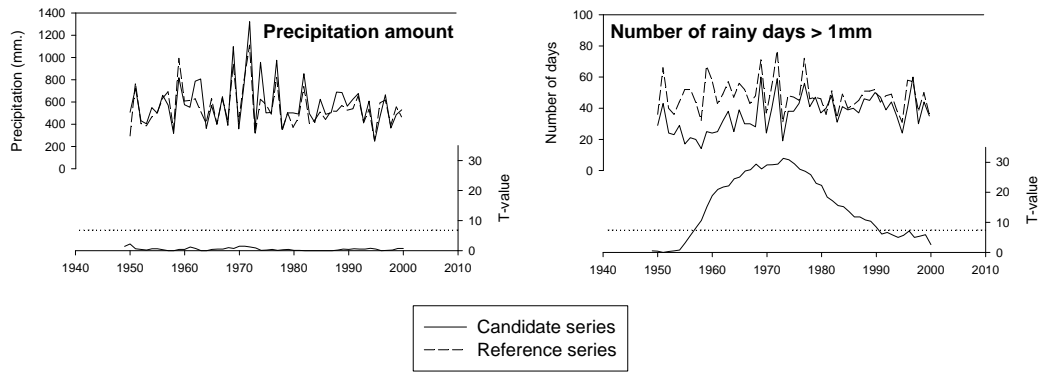


Figure 10. Annual series of precipitation amounts and number of rainy days at the l'Ametlla de Mar observatory. The series of T-values and the limit of confidence (dotted line) are also shown.

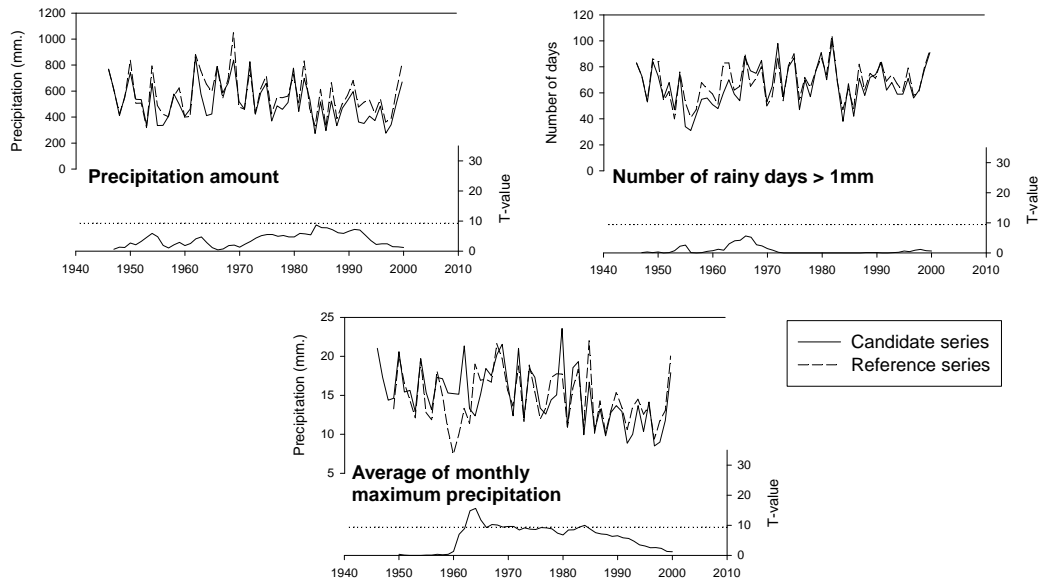


Figure 11. Annual series of precipitation amounts, number of rainy days and the average maximum precipitation at the Cedejas de la Torre observatory. The series of T-values and the limit of confidence (dotted line) are also shown.

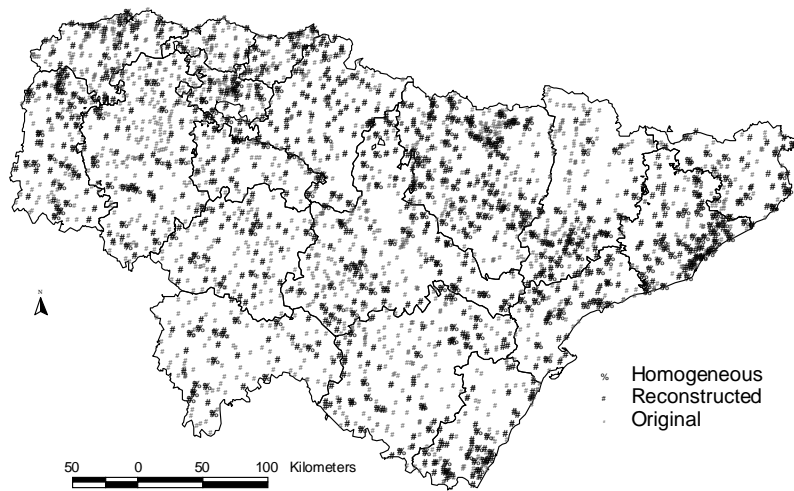


Figure 12. Homogeneous daily precipitation observatories in the study area after reconstruction, the quality control process, and homogenization testing.

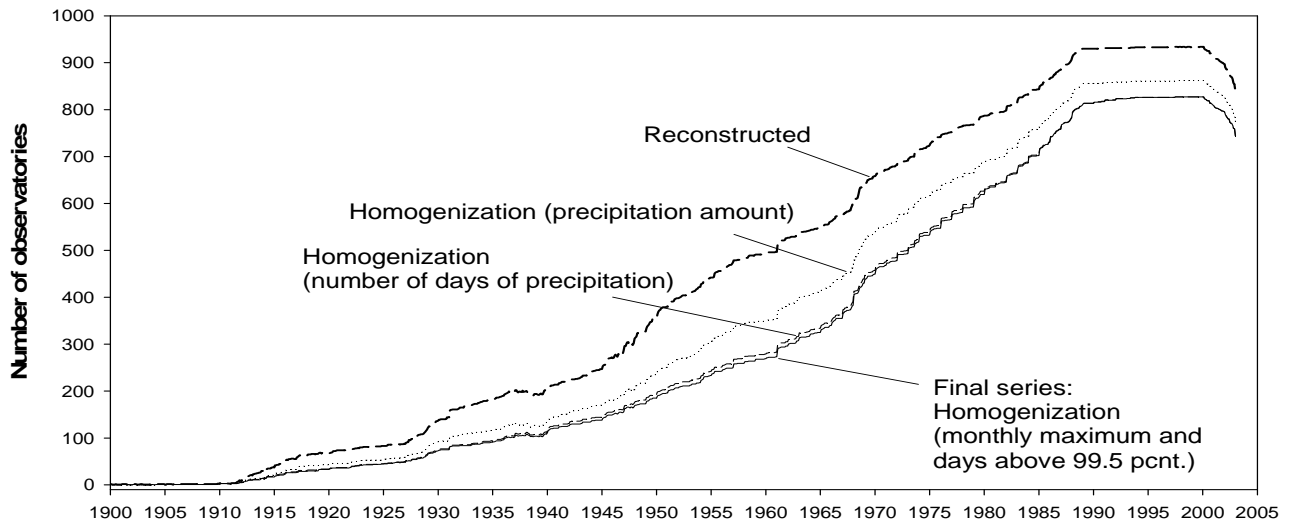


Figure 13. Time variation of the number of data series at various stages in construction of the database and quality control process.

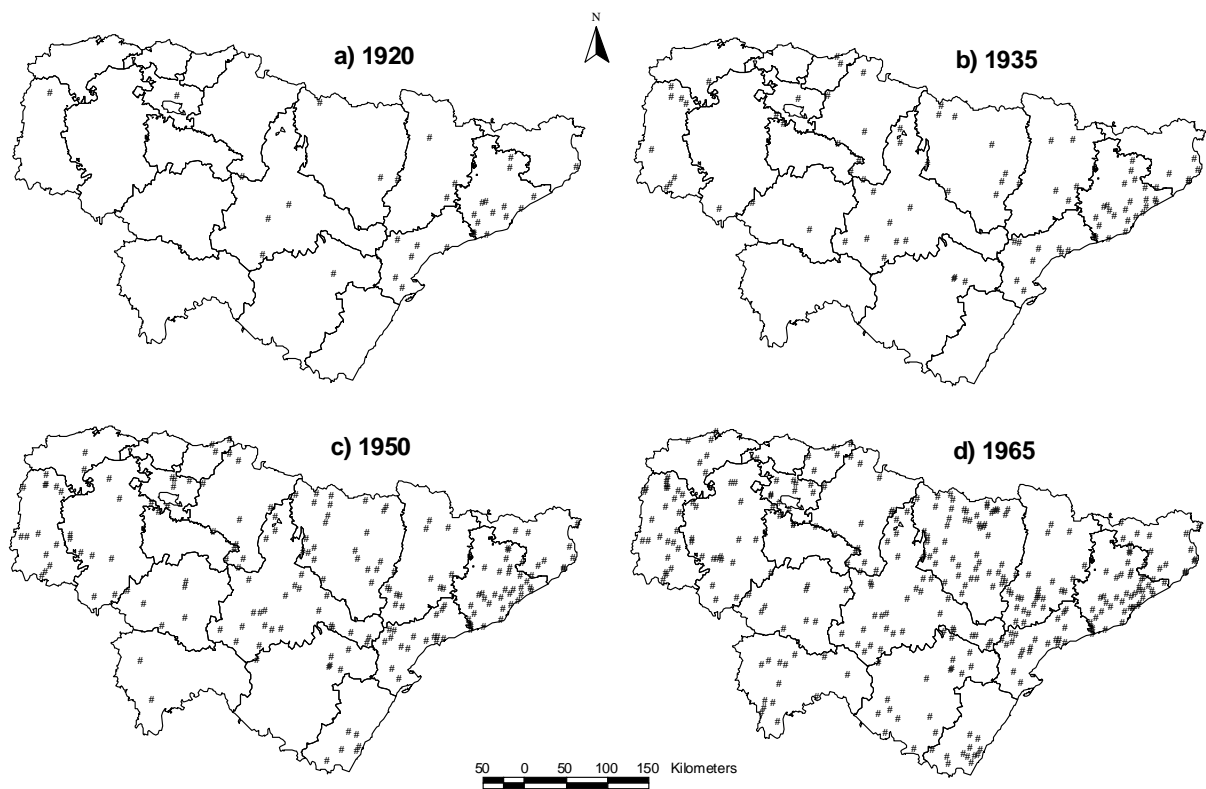


Figure 14. Spatial distribution of homogeneous data series starting in or before: a) 1920, b) 1935, c) 1950 and d) 1965.

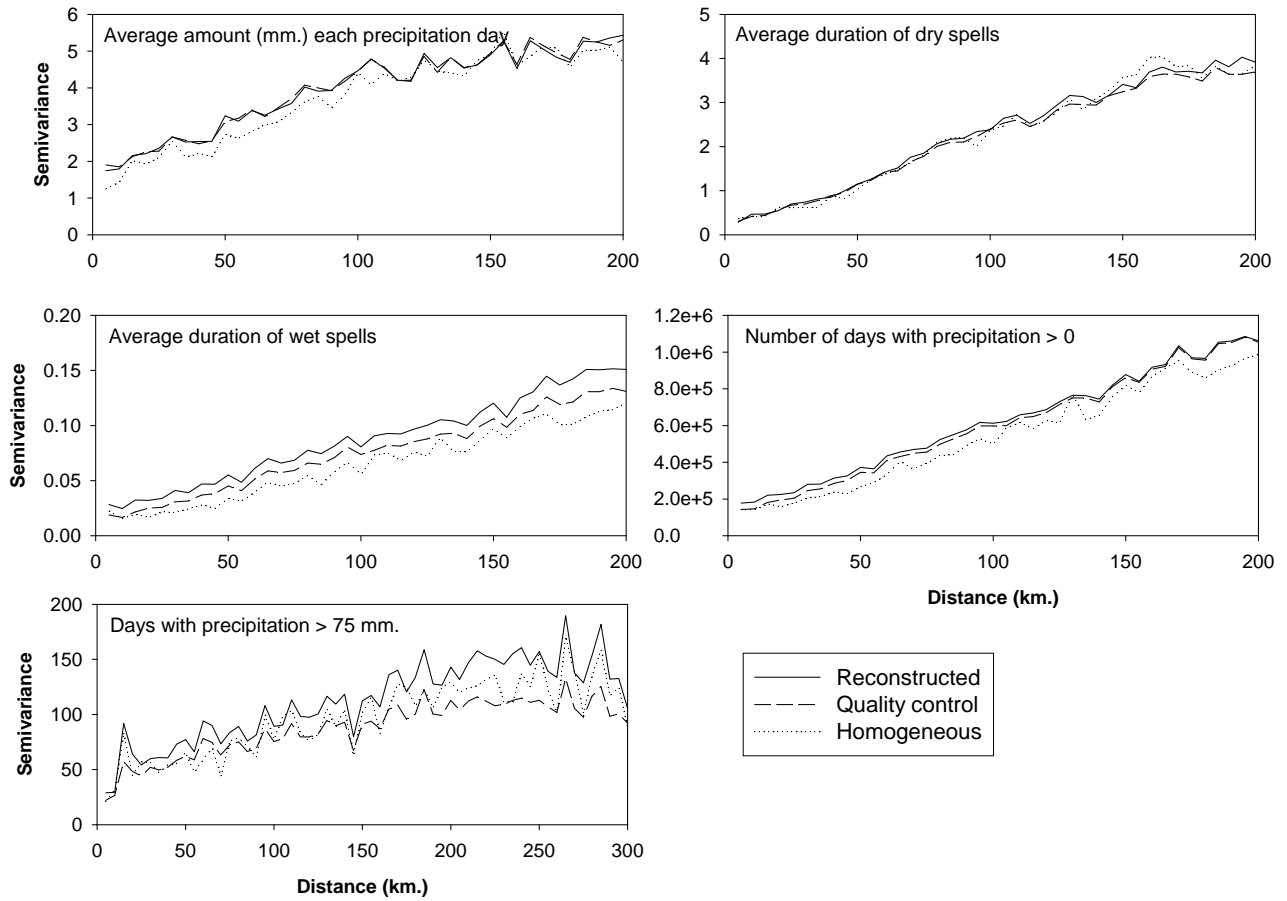


Figure 15. Semivariance of several precipitation parameters in the reconstructed, quality-controlled and homogeneous datasets.



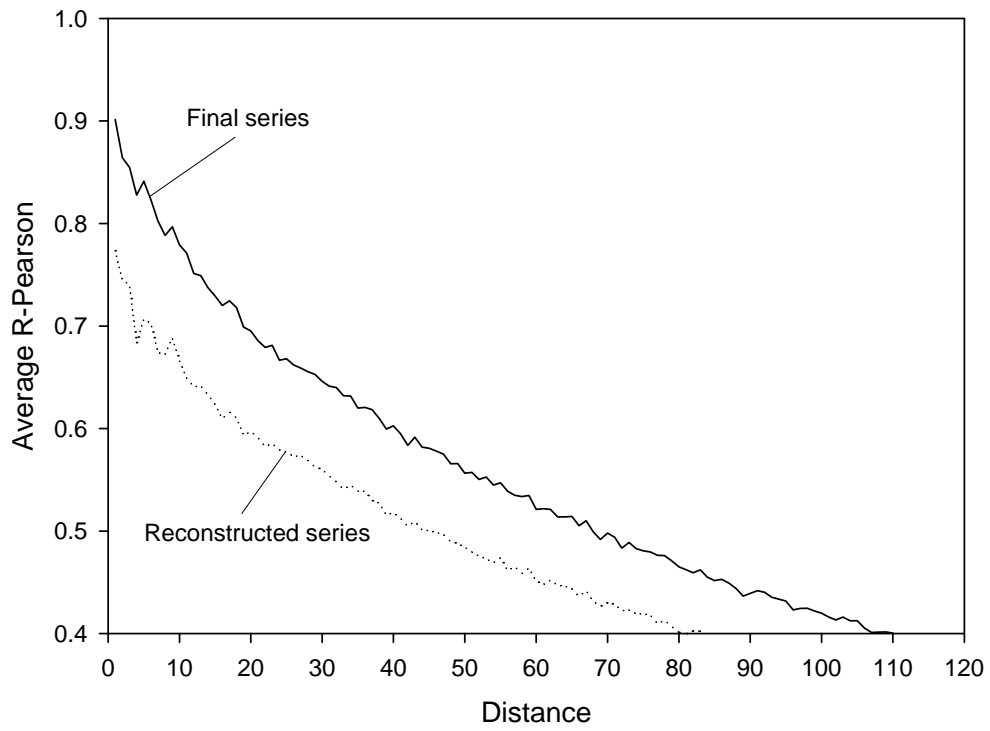


Figure 16. Average Pearson's R correlations among the time series as a function of the distance lag after reconstruction and homogeneity testing (final series).