

## Quality Control and Homogeneity of Precipitation Data in the Southwest of Europe

J. FIDEL GONZÁLEZ-ROUCO

*Departamento de Astrofísica y Ciencias de la Atmósfera, UCM, Madrid, Spain, and  
GKSS Forschungszentrum, Geesthacht, Germany*

J. LUIS JIMÉNEZ

*Departamento de Astrofísica y Ciencias de la Atmósfera, UCM, Madrid, Spain*

VICENTE QUESADA

*Departamento de Estadística e Investigación Operativa, UCM, Madrid, Spain*

FRANCISCO VALERO

*Departamento de Astrofísica y Ciencias de la Atmósfera, UCM, Madrid, Spain*

(Manuscript received 19 July 1999, in final form 19 April 2000)

### ABSTRACT

A quality control process involving outliers processing, homogenization, and interpolation has been applied to 95 monthly precipitation series in the Iberian Peninsula, southern France, and northern Africa during the period 1899–1989. A detailed description of the procedure results is provided and the impact of adjustments on trend estimation is discussed.

Outliers have been censored by trimming extreme values. Homogeneity adjustments have been developed by applying the Standard Normal Homogeneity Test in combination with an objective methodology to select reference series.

The spatial distribution of outliers indicates that they are due to climate variability rather than measurement errors. After carrying out the homogeneity procedure, 40% of the series were found to be homogeneous, 49.5% became homogeneous after one adjustment, and 9.5% after two adjustments. About 30% of the inhomogeneities could be traced to information in the scarce history files.

It is shown that these data present severe homogeneity problems and that applying outliers and homogeneity adjustments greatly changes the patterns of trends for this area.

### 1. Introduction

In the last few decades there has been an increasing concern regarding the changes in climate caused by anthropogenic emissions of greenhouse gases and aerosols (Houghton et al. 1992, 1995). This interest has been mainly based on the results of climate models. In a similar fashion, there has been a rise in the activity directed toward climate change detection in observed data (Santer et al. 1995; Hegerl et al. 1996). It is essential, for this purpose, to use high quality observations. Thus outliers and homogenization arise as important issues.

Historically, the identification of outliers has been the primary emphasis of quality control work (Filippov 1968; Grant and Leavenworth 1972). Outliers are observations very distant from the mean value that can be due to measurement errors or to extreme meteorological events. Several approaches that focus on temporal and/or spatial variability can be applied in order to identify outliers and diagnose whether they are erroneous or not (Barnett and Lewis 1994; Eischeid et al. 1995; Peterson et al. 1998a). When outliers are undoubtedly erroneous measurements those extreme data can be rejected and the problem is converted into one of missing data treatment. When outliers have a physical background the question arises whether they should be corrected or not (Barnett and Lewis 1994). On the one hand, extreme data carry very valuable climatological information that should not be dismissed. On the other hand, many statistical techniques of common use are not resistant, that is, they are sensitive to the presence of outliers that can

---

*Corresponding author address:* J. Fidel González-Rouco, GKSS Forschungszentrum, GUA, GEB. 38, Max-Planck-Str. 1, 21502 Geesthacht, Germany.  
E-mail: fidel@gkss.de

affect the estimation of sample statistics (Lanzante 1996). Such is the case of all moments-based statistics (Peterson et al. 1998a).

The last issue can be approached in several ways. The most objective approach would be always to use techniques that are resistant (Lanzante 1996). Unfortunately this is not the case for many of the available methods. An alternative and rather radical solution would be removing all outliers. This would allow a safe use of nonresistant statistics at the expense of losing interesting climatological information. Another alternative that would combine keeping the information of extreme data and showing consideration for nonresistant methods would be to censor outliers by means of replacing them by some threshold value that keeps the information of an extreme event and yet does not have such an important influence on nonresistant statistics (Barnett and Lewis 1994). The original value of the outliers can be later restored for specific studies concerning extreme values. This last subjective approach has been the one used herein as a preprocessing for the nonresistant homogenization techniques employed in this work.

A homogeneous climate series is defined as one where variations are caused only by changes in weather and climate (Conrad and Pollak 1950). This would mean that the measurements have been consistently registered by the same method, with the same instrumentation, at the same time of day and place and in the same environment. Unfortunately, this is seldom the case and at most old observatories, instruments, location, observer, or other environmental factors have been altered (Jones et al. 1985; Karl and Williams 1987; Peterson et al. 1998a). In the case of precipitation, progressive improvements of instrumentation can also introduce artificial systematic increases (Hanssen-Bauer and Førland 1994), and thus long-term variations and trends should be interpreted cautiously (Houghton et al. 1992). Accurate long-term climate analysis requires homogeneous data, thus, inhomogeneities must be detected and inhomogeneous series be excluded or adjusted.

Several statistical tests have been developed that allow detection of inhomogeneities in time series. Some depend on the metadata (Karl and Williams 1987; Rhoades and Salinger 1993) and others use them as additional information when the station history is poorly documented or missing at all (Buisband 1982; Alexanderson 1986). Some tests are univariate, while others use information from reference time series that are compared with the test series to decide upon its quality. A comprehensive review on this topic has been presented by Peterson et al. (1998b).

Considerable research has been accomplished on the characteristics of low-frequency climate variability of precipitation in the western Mediterranean (Maheras 1988; Valero et al. 1993; Zorita et al. 1992; Fernández-Mills 1995; Valero et al. 1996b; Rodríguez-Puebla et al. 1998; Serrano et al. 1999; González-Rouco et al. 2000). However, homogeneity assessments of historical

precipitation data during the twentieth century have seldom been performed (Almarza et al. 1996; Garrido et al. 1996; Esteban-Parra et al. 1998). Thus, a detailed description of this problem from the perspective of an objective methodology using all the information from a high number of stations is still to come.

In this work, we focus on the detection and treatment of outliers and inhomogeneities in a dataset of precipitation in the southwest of Europe. The adjustment of outliers has been performed by taking into account dispersion in the time series by means of windsorization, that is, trimming the values that exceed a certain threshold (Barnett and Lewis 1994). For detection and processing of inhomogeneities we have used a procedure based upon the Standard Normal Homogeneity Test developed by Alexanderson (1986). This method assumes invariance of the ratio of the values of the test series and those of a reference series built using information from several neighboring stations. An inhomogeneity in the test series will be revealed by a systematic change in this ratio. However, if there exists no a priori knowledge of which ones are the homogeneous time series, the test has to be applied several times to decide which stations have good quality records and can be used as a reference. We have followed a procedure divided in six steps that allows identifying reference series and adjusting inhomogeneous ones. The first five steps follow those of Hanssen-Bauer and Førland (1994) and the sixth one has been added to account for series with two inhomogeneities.

It is worthwhile making a comment on the limits of the corrections applied in this work. The procedures herein are applied with the intention of improving the quality of the data, but not of providing an error-free dataset. The use of the term "correct" should not be taken to imply a level of perfection in the adjustments that does not exist in any dataset. Future inclusion of new data and/or the use of other methods will hopefully lead to further improvements.

Data are described in the next section. A description of the methodology is presented in section 3. Results concerning outliers and homogenization are presented in the following section as well as a description of their impact on trends. Finally conclusions are summed up and discussed.

## 2. Data

We have used long time series of monthly precipitation totals (mm) covering most of the Iberian Peninsula, south of France and north of Africa and spanning the period 1899–1989. Iberian rain gauge data were provided by the Instituto Nacional de Meteorología (Spain). The Carbon Dioxide Information Analysis Center (United States) and MetéoFrance (France) supplied African and French precipitation data, respectively. Figure 1 shows the spatial distribution of the 95 observatories; the corresponding names and lengths of the as-

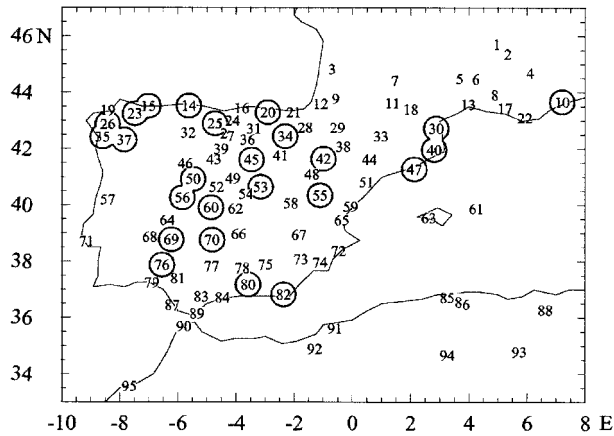


FIG. 1. Spatial distribution of precipitation stations. Numbers refer to those of Table 1. Circles point time series in which information from secondary sites was used.

sociated time series are listed in Table 1. We have selected the period 1899–1989 for this first version of the data.

Noncircled numbers in Fig. 1 stand for sites in which time series were used unchanged to apply the procedures described in the next section. In some cases (circled identifiers in Fig. 1) long data gaps were present and/or temporal extension of the time series was desirable. In order to improve the quality of these records and also the spatial coverage of the dataset additional information from 43 secondary sites (nonnumbered rows in Table 1) was used. Thus, initial corrections following the ratio method (Mitchell et al. 1966) were calculated and applied using data from the overlapping periods between the base stations and the additional observatories. The lower threshold value for correlation between the base station time series and the time series of a candidate secondary site to be used for data interpolation was 0.8.

### 3. Methodology

In this section a brief description is provided on the development of outliers and homogeneity corrections.

#### a. Outliers

Historically, the identification of outliers has been the primary emphasis of quality control work (Grant and Leavenworth 1972). Outliers can be erroneous measurements or correct extreme values. In this version of the dataset the approach has not been searching for erroneous measurements, but reducing the size of the distribution tails in order to make a safer use of the non-resistant homogenization techniques used later. This has been performed by trimming extreme data by a threshold value. This extreme value keeps the information of an extreme event, but its reduced magnitude in comparison with the original value has less impact on non-resistant statistics (Lanzante 1996). In the next section

the spatial distribution of values identified as outliers will be described showing that in most cases the trimmed values have a physical meaning.

Outliers were identified as those values trespassing a maximum threshold for each time series (Trenberth and Paolino 1980; Peterson et al. 1998a) defined by

$$P_{\text{out}} = q_{0.75} + 3\text{IQR}, \quad (1)$$

where  $q_{0.75}$  is the third quartile and IQR the interquartile range. The IQR has been used in quality control of climate data (Eischeid et al. 1995) because it is resistant to outliers. Values over  $P_{\text{out}}$  were substituted by this limit. This way of proceeding reduces the bias caused by outliers and yet keeps the information of extreme events (Barnett and Lewis 1994).

#### b. Homogenization

##### 1) THE STANDARD NORMAL HOMOGENEITY TEST

The homogenization procedure is based on the application of the Standard Normal Homogeneity Test (Alexanderson 1986). This test is based on the assumption that precipitation amounts at the station being tested (test station) and some regional average values are proportional to each other. This relationship is expressed in terms of the ratio  $q$  between the test station normalized precipitation values and those of a regional time series defined as a weighted average of several neighboring reference stations. Thus, the ratio  $q$  in a specific year may be denoted as

$$q_i = \frac{F_i}{G_i}, \quad i = 1, \dots, n, \quad (2)$$

$n$  being the number of time steps, and  $F_i$  and  $G_i$  functions of the precipitation at the test and reference stations that are defined as

$$F_i = \frac{P_i}{\bar{P}}, \quad G_i = \frac{\sum_{j=1}^{k_i} v_j \frac{Q_{ij}}{\bar{Q}_j}}{\sum_{j=1}^{k_i} v_j}, \quad (3)$$

$P_i$  being the precipitation at the test station,  $Q_{ij}$  the precipitation at the  $j$ th reference station, and  $v_j$  a weight factor for the  $j$ th reference station. In this work  $v_j$  was defined as the square of the correlation coefficient between the test series and the  $j$ th reference series. Here  $k_i$  is the number of reference sites used in time step  $i$  that varies as a result of different length of time series and the existence of missing data. Overbars denote averages throughout the observational period.

In order to apply the test the  $q_i$  values are standardized by their sample mean,  $\bar{q}$ , and standard deviation,  $s_q$ :

$$z_i = \frac{(q_i - \bar{q})}{s_q}.$$

This series has zero mean and unit standard deviation

and is assumed to be normally distributed,  $N(0, 1)$ . An inhomogeneity in the test series would affect the series of ratios and thus  $z_i$ . Therefore, the homogeneity test is based on the definition of the following hypotheses.

*Null hypothesis, ( $H_0$ ):* The test series is homogeneous. Any subset of  $z_i$  is distributed as  $N(0, 1)$ .

*Alternative hypothesis, ( $H_1$ ):* The test series is inhomogeneous. There is an inhomogeneity in year  $m$  such that the  $m$  first years of the  $z_i$  have mean value  $\mu_1$  and the  $(n - m)$  last years have mean value  $\mu_2$ , with the standard deviation 1 in both temporal segments.

These definitions enable the test to detect and adjust only one inhomogeneity in the original test series. A test statistic,  $T_m$ , is computed for each of the  $n - 1$  possible change points:

$$T_m = m\bar{z}_1^2 + (n - m)\bar{z}_2^2, \quad m = 1, \dots, n - 1, \quad (4)$$

where  $\bar{z}_1$  and  $\bar{z}_2$  are the mean values of  $z_i$  during the  $m$  first and the  $(n - m)$  last years, respectively. A high  $T$  value in year  $m$  implies that  $\mu_1$  and  $\mu_2$  depart significantly from zero, making  $H_0$  unlikely.

The maximum  $T_m$  value in the time series will be denoted by  $T_x$ . The probability that  $T_x$  exceeds a certain value if  $H_0$  is correct depends only on the length of the time series. The critical  $T$  values for the 5% and 10% significance levels ( $T_{95}$  and  $T_{90}$ , respectively) are given by Alexanderson (1986). In practice, the test is applied making use also of the existing information about the history of the station, the metadata. A precipitation series is then classified as inhomogeneous when (i) it contains an inhomogeneity significant at the 5% level 5 yr or more from either end of the series, and (ii) the series contains an inhomogeneity significant at the 10% level, which is explainable by metadata.

The interaction with the metadata information at the 10% level reduces the probability of rejecting homogeneous series, thus increasing the power of the test. The reason for not accepting unexplained inhomogeneities close to the ends of the series is an increased probability for high  $T$  values near the ends (Hawkins 1977).

Once a series is identified as inhomogeneous, the data before the inhomogeneity date are corrected through multiplying them by

$$f = \frac{q_a}{q_b}, \quad (5)$$

where  $q_a$  and  $q_b$  are the mean values of  $q_i$  after and before the inhomogeneity. When the date of a change in the historical records was close to the year corresponding to  $T_x$ , the historical date was used for the correction. Otherwise, the year corresponding to  $T_z$  was assumed.

Equations (2) and (3) can also be used for interpolating missing data (Alexanderson 1986; Valero et al. 1997a). Assuming  $q(t) \approx \bar{q} = 1$  for missing data, the estimations of the missing  $F_i$  values,  $\hat{F}_i$ , can be obtained from

$$\hat{F}_i = q_i G_i \approx \bar{q} G_i = G_i, \quad (6)$$

and thus precipitation for the test station can be estimated from

$$\hat{P}_i = \bar{P} G_i. \quad (7)$$

## 2) SELECTION OF REFERENCE STATIONS

The above description states how factors can be calculated to produce an homogeneous test series by using information for several reference stations. However, usually no a priori knowledge exists on which ones should form the reference series and the test has to be run in an iterative way several times to decide upon which are the best quality series. We have followed the same procedure as Hanssen-Bauer and Forland (1994) for this purpose. We have added one step to their procedure in order to account also for series with two inhomogeneities. Figure 2 highlights the steps of the method. The three first stages are essentially devoted to improve the selection of reference stations:

*Step 1: A first homogeneity assessment.* In this step the test is applied to each time series. The reference series will be those that correlate best with the test station. Since reference stations were not necessarily homogeneous, the result is a first approach to make a classification of homogeneous (H1; see Fig. 2) and inhomogeneous (I1) series.

*Step 2: Using adjusted reference series.* Corrections are made to the I1 set following Eq. (5). Subsequently the test is applied to each one of the original time series again though using an improved dataset as reference series: the H1 group and the corrected I1 series. Since this correction reduces the amount of inhomogeneities in the reference set this produces a new, somewhat optimized, classification into homogeneous (H2) and inhomogeneous (I2) series. The only difference with step 1 is using adjusted reference series.

*Step 3: Applying the test to corrected series.* The I2 group is corrected and then the test is applied only to these adjusted time series; the pool of possible reference stations is built with the H2 set and the adjusted I2 stations. The result is a group of homogeneous time series that have been corrected once (HC3) and a subset of inhomogeneous series with more than one inhomogeneity (IC3).

At this stage of the procedure all time series can be grouped into an homogeneous group (H2 and HC3) and an inhomogeneous one (IC3). The next two steps revise this classification by using as a reference only time series that have proved to be homogeneous after applying the test.

*Step 4: Testing again using only homogeneous series.* All series are checked again: the homogeneous H2 group and the corrected I2 group. The reference series are only those that have proved their homogeneity so far, namely, the H2 and HC3 sets. The H2 set provides two new subsets after applying the test, a group of homogeneous

TABLE 1. Names and periods of the time series used in the compilation of the precipitation dataset. Numbers match those in Fig. 1. Correlation values between main sites and additional secondary observatories are also shown.

No.	Name	Period	Correlation values
1	Lyon	1841–1994	
2	Grenoble	1833–1976	
3	Bourdeaux	1899–1994	
4	Gap	1899–1992	
5	Bagnols Les Bains	1899–1993	
6	Joyeuse	1900–93	
	Service des crues	1900–30	0.98
7	Cahors	1899–1993	
8	Avignon	1899–1993	
9	Saint Server	1899–1993	
10	Nice	1899–1993	
	Rue Gioffredo	1891–1920	0.90
11	Toulouse	1899–1994	
12	Peirehorade	1851–1992	
13	Montpellier	1899–1993	
14	Gijon	1913–94	
	Cabo Peñas	1914–22	0.76
15	Castropol	1907–94	
	Tapia de Casariego	1907–25	0.80
	Castropol	1923–72	0.95
16	Santander	1912–91	
17	Marselle	1899–1994	
18	Revel	1899–1973	
19	Coruña	1899–1993	
	Bilbao sondica	1899–1993	
	Guernica	1925–30	0.86
	Larrasquitu	1925–46	0.86
	Machichaco	1913–30	0.77
	Basauri	1940–50	0.93
21	San sebastian	1899–1993	
22	Toulon	1899–1994	
23	Lugo	1913–94	
	Mundín	1913–21	0.92
	Brigós	1922–39	0.94
	Sarria	1952–66	0.91
24	Reinosa	1911–91	
25	Camporredondo	1917–92	
26	Santiago	1906–93	
	Universidad	1899–1950	0.89
27	Cerv. pisuerga	1912–86	
28	Pamplona	1899–1993	
29	Canfranc	1910–94	
30	Perpignan	1899–1993	
	Pont Rouge	1899–1923	0.98
31	Oña	1900–94	
32	Leon	1904–93	
33	Capdella	1915–91	
34	Logroño	1911–94	
	Instituto	1911–47	0.85
35	Pontevedra	1899–1994	
	Salcedo	1949–85	0.96
36	Burgos	1899–1991	
37	Orense (Granja)	1899–1994	
	Santuario Molas	1924–25	0.90
	Beariz	1926–35	0.87
	Villamayor	1937–50	0.90
	Ginzo	1969–72	0.91
38	Huesca	1899–1994	
39	Palencia	1913–88	
40	Gerona	1911–91	
	Instituto	1911–69	0.89
41	Soria	1899–1994	
42	Zaragoza	1899–1994	

TABLE 1. (Continued)

No.	Name	Period	Correlation values
	Facultad de ciencias	1899–1941	0.89
43	Valladolid	1899–1994	
44	Lerida	1913–91	
45	Vid de aranda	1899–1994	
	Col. Franciscanos	1914–39	0.85
	Aranda del Duero	1943–94	0.84
46	Zamora	1909–94	
47	Barcelona	1899–1994	
	Universidad	1899–1937	0.91
48	Daroca	1899–1993	
49	Segovia	1899–1994	
50	Salamanca	1899–1993	
	Observatorio	1899–1944	0.94
51	Tortosa	1899–1991	
52	Avila	1900–94	
53	Guadalajara	1912–91	
	Atienza	1937–45	0.81
54	Madrid	1899–1994	
55	Teruel	1899–1994	
	Ariño	1936–43	0.90
56	Hervas	1913–94	
	Bejar	1913–79	0.90
	Aldea del Camino	1951–54	0.90
57	Coimbra	1899–1989	
58	Cuenca	1908–90	
59	Castellón	1911–93	
60	Talavera	1913–90	
	Candeleda	1936–40	0.90
	Robledillo de la Vera	1940–42	0.91
	Marrupe	1944–47	0.87
	Alcañizo	1959–62	0.96
	Cabañuelas	1989–90	0.85
61	Mahón	1899–1994	
62	Toledo	1908–91	
63	Palma mallorca	1899–1993	
64	Caceres	1904–91	
65	Valencia	1899–1994	
66	Ciudad real	1904–91	
67	Albacete	1899–1991	
68	Badajoz	1899–1994	
69	Alange	1913–94	
	Valencia del Ventoso	1913–34	0.89
	Miajadas	1917–50	0.91
70	Almaden	1913–94	
	Pozoblanco	1936–94	0.84
71	Lisboa	1899–1989	
72	Alicante	1899–1993	
73	Cehegim	1913–94	
74	Murcia	1899–1993	
75	Cazorla	1911–94	
76	Aracena	1913–94	
	Galaroza	1960–94	0.95
	Valdeazufre	1964–94	0.96
77	Cordoba	1911–94	
78	Jaen	1899–1993	
79	Huelva	1903–94	
80	Granada	1899–1993	
	Alhendim	1977–90	0.93
81	Sevilla	1899–1994	
82	Almeria	1910–94	
	Estación sismológica	1909–32	0.89
	Ciudad Jardín	1933–67	0.89
83	Grazalema	1912–94	
84	Malaga	1900–95	
85	Dar-el-beida	1838–1994	
86	Algiers	1889–1994	

TABLE 1. (Continued)

No.	Name	Period	Correlation values
87	San Fernando	1899–1989	
88	Constantinopla	1899–1994	
89	Gibraltar	1899–1994	
90	Tanger	1899–1989	
91	Oran	1899–1994	
92	Tlemcem	1899–1989	
93	Biskra	1902–94	
94	Djelfa	1899–1994	
95	Casablanca	1902–89	

series (H4) and a group of series that contain at least one inhomogeneity (I4). The adjusted I2 is divided into a group of homogeneous series (HC4) and a set of series that have at least two inhomogeneities (IC4).

Step 5: *Last corrections for one inhomogeneity.* The inhomogeneous group I4 is corrected and checked for new inhomogeneities. The reference series are taken from the just updated homogeneous (H4) and homogeneous after one correction (HC4) sets.

In steps 3–5 corrections for series with one inho-

mogeneity are applied. Step 6 has been incorporated to correct two inhomogeneities.

Step 6: *Correcting two inhomogeneities.* The test is applied to parts of the series containing only one inhomogeneity. First, the data before the first inhomogeneity in series belonging to the groups IC4 and IC5 are omitted and the remaining data are corrected using as reference series those classified as homogeneous in the five preceding steps (H4, HC4, HC5); second, the test is run with all the data in order to correct the first inhomogeneity.

4. Results

a. Distribution of outliers

Figure 3a shows the distribution of  $P_{out}$  values [Eq. (1)]. This plot describes the variability of the data that reaches maximum values in the northwest and south of the Iberian Peninsula and most of the Pyrenees. Data dispersion achieves its lowest values in the central Plateau and along the Mediterranean coast where the frequency of extreme events is higher (Linés 1970; Valero

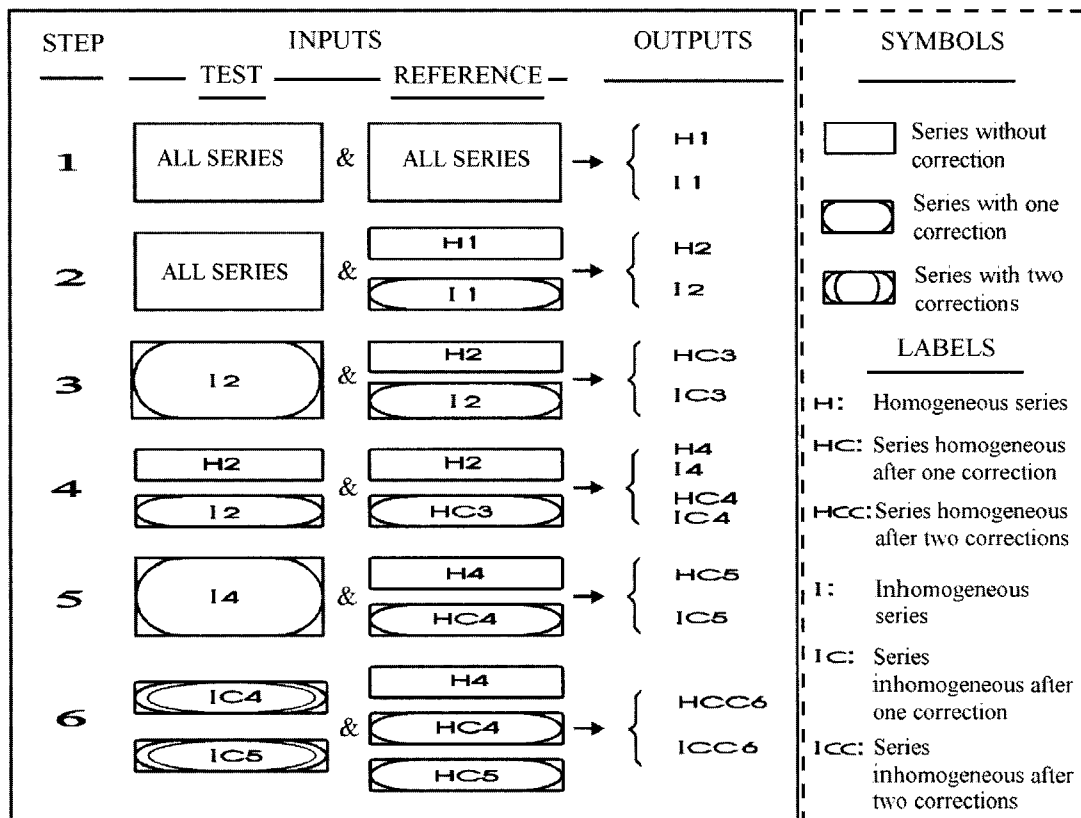


FIG. 2. Diagram showing the six steps of the homogenization procedure [after Hanssen-Bauer and Førlund (1994)]. Test series are tested by comparing with reference series and sorted into output groups at each step of the procedure for subsequent use. The output groups are identified by the result of the test (H for homogeneous and I for inhomogeneous), by the number of corrections applied to the input test series (C for one correction and CC for two corrections) and by the number of the step in which the test is performed. Independently of this convention symbols indicate the number of corrections applied to any input series.

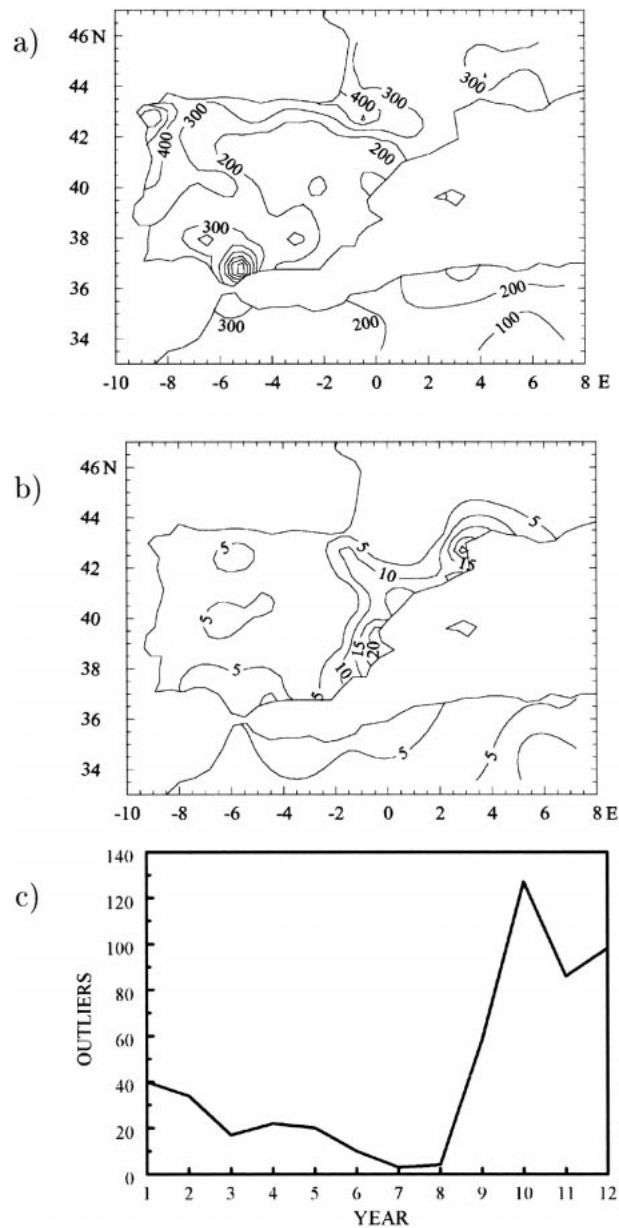


FIG. 3. (a) Spatial distribution of  $P_{out}$  (mm) values. (b) Spatial distribution of the absolute number of outliers in each station. Contour interval is 100 mm. (c) Time evolution of the monthly absolute number of outliers in the area of interest. Contour interval is 5 units.

et al. 1997b). This agrees with the distribution of outliers in Fig. 3b that shows a clear maximum in the Mediterranean area of influence.

The total number of corrected values was 519: 12% of them were produced in spring, 3% in summer, 52% in autumn, and 33% in winter. Figure 3c shows the corresponding monthly distribution, characterized by a maximum in October and a minimum in the summer.

The spatial distributions of the number of outliers for each season (not shown here) reveals that maxima are located in the east coast in spring and autumn in a similar

TABLE 2. Number and type of series involved in each step of the homogenization procedure.

Step	Input series		Output series					
	Test	Reference	H	HC	HCC	I	IC	ICC
1	95	95	43	—	—	52	—	—
2	95	95	42	—	—	53	—	—
3	53	95	—	45	—	—	8	—
4	95	87	38	44	—	4	9	—
5	4	82	—	3	—	—	1	—
6	10	85	—	—	9	—	—	1

fashion as Fig. 3b. In winter they spread over the west and south of the Iberian Peninsula and for the summer months there is no apparent structure but isolated cases in the central plateau.

The overall agreement of the areas and timing of outliers with those of extreme events supports the idea of physical mechanisms as causes for the detected anomalous data rather than human-induced errors.

### b. Homogenization

For this version of the dataset, homogeneity corrections were developed from the annually averaged time series and subsequently applied to monthly data. Monthly corrections can also be incorporated through aligning the whole series as a sequence of  $12n$  values,  $n$  being the number of years or through considering individual monthly series (Alexanderson 1986). In the former case, results support those obtained using annual values, although somewhat more significant breaks are obtained. Also, there are cases of dry months giving very uncertain ratios that distort the test. This is also the situation when treating individual monthly series.

This is relevant when applying the homogenization process described in the last section because it is a non-robust method (distribution dependent). The  $q_i$  ratios in Eq. (2) are assumed to follow a Gaussian distribution and deviations from it could damage the performance of the test.

The problem of very dry months, a frequent one in the southern stations, is reduced when using annual data. After applying a  $\chi^2$  test for the goodness of the fit (Snedecor and Cochran 1972) of the  $q_i$  to a normal distribution, it was found that for a significance level of 0.1, 88% of the  $q_i$  series in step 1 [see section 3b(2)] met the Gaussian hypothesis. Since the non-Gaussian 12% is in the range of the significance level, the approximation was made that the set of  $q_i$  were normally distributed.

Table 2 shows the number and type of series involved in each step (column 1) of the homogenization procedure. The number of time series tested in each step is shown in column 2 and the number of series making part of the reference group appears in column 3. The quantity of resulting homogeneous (H) or inhomoge-

neous (I) time series are indicated in columns 4–9, either after one (HC, IC) or after two (HCC, ICC) corrections.

In the two first steps almost the same results are obtained. The most relevant changes are shifts in the classification of some series that switch from homogeneous in step 1 (H1) to inhomogeneous in step 2 (I2; stations 42, 51, 72, and 90) or vice versa (stations 13, 23, 34, 79, and 94). In the third stage the 53 inhomogeneous series obtained in step 2 are corrected and tested producing a subset of 45 homogeneous after one correction (HC3) and 8 inhomogeneous (IC3) series.

Following the methodology previously described, in the fourth step the subset of 42 homogeneous series in step 2 (H2) was tested and provided 38 purely homogeneous series without any corrections (H4) and 4 series that contain at least one inhomogeneity (I4). The 53 I2 series were corrected and tested again and provided 44 series homogeneous after one correction (HC4) and 9 series with at least two inhomogeneities (IC4).

During the fifth stage of the method the 4 I4-type series obtained in step 4 were corrected and tested. The output was a set of 3 homogeneous series (HC5) and 1 with at least 2 inhomogeneities (IC5).

In the last step, the procedure was applied to the 10 series with more than one inhomogeneity (IC4 and IC5) being 9 of them successfully corrected (HCC6) and only one (station 69, Alange), which showed more than two inhomogeneities (ICC6). This last station was rejected from the dataset. Here 40% of the series were homogeneous, an additional 49.5% became homogeneous after one correction, and 9.5% more after two corrections.

Figure 4a describes the type (symbols) of inhomogeneity of each time series. It can be appreciated that there is no evident spatial structure in the distribution of inhomogeneities. The same figure shows the number of reference stations used for the analysis of each series (solid lines). This number was imposed to be always between 5 and 10 and was chosen as a function of the correlation (always higher than 0.5) between the test station and any candidate reference station. There is a clear structure, with the highest values in the west of the Iberian Peninsula and the south of France where correlations between pairs of time series are also higher than in any other area. This also corresponds with the areas of highest correlation values between the  $F$  and the  $G$  series [Eq. (3)] that are shown in Fig. 4b. The high correlations between  $F$  and  $G$  also support our criterion for selecting reference stations.

Figure 4c shows the temporal distribution of correction factors. Inhomogeneities that are not supported by metadata are identified with a cross sign and those that agree with the history of the station with a circle or a triangle corresponding to a relocation or a change in height, respectively. The histogram shows the distribution of the changes during the century, both for the detected inhomogeneities (black bars) and for the changes registered in the metadata (white bars).

More complete history files of stations would be de-

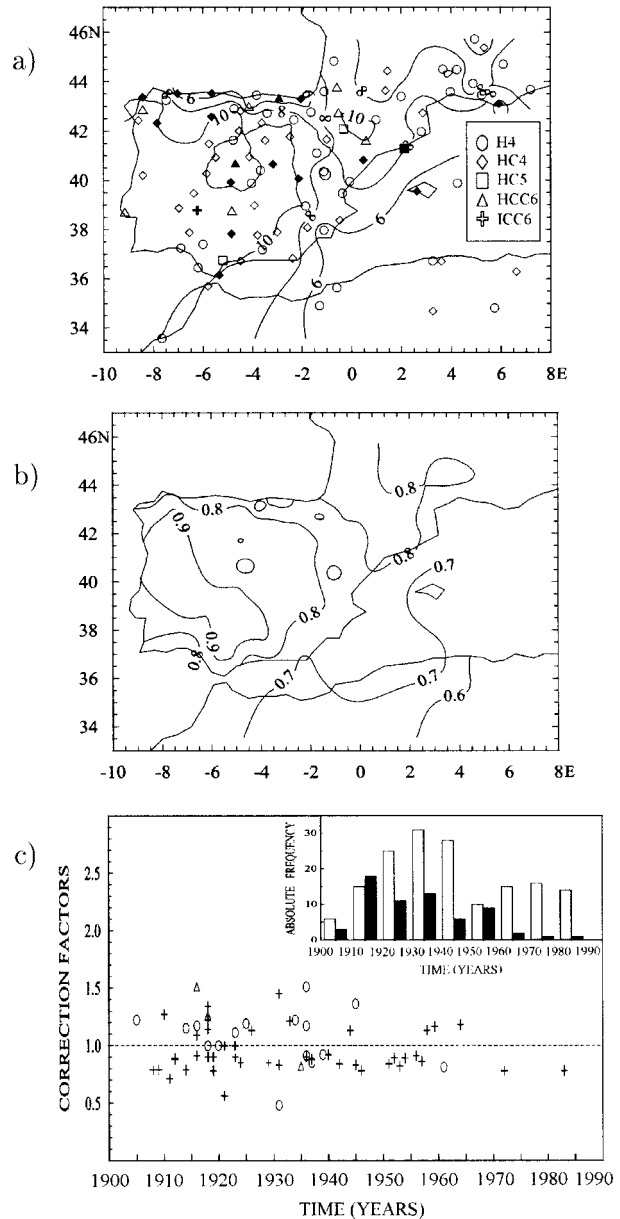


FIG. 4. (a) Type of time series after applying the homogenization procedure: homogeneous (H), homogeneous after one (HC) or two (HCC) corrections, and inhomogeneous after two corrections (ICC). Numbers in the legend stand for the step of the procedure. Solid symbols indicate agreement with metadata. Isolines indicate the number of reference stations used. Contour interval is 2 units. (b) Correlation values between  $F$  and  $G$  [see Eq. (3)]. Contour interval is 0.1. (c) Distribution of correction factors within the studied period. Crosses stand for inhomogeneities of unknown origin, circles correspond to relocations or interpolations of data from distant sites, and triangles to changes in height, the histogram of inhomogeneity dates both for the detected changes (black bars) and those registered in the metadata (white bars).



sirable to discuss the potential influence of changes in instrumentation, observer, and other factors upon the quality of the dataset. Unfortunately, only relocation and changes in the height of a few stations could be related to identified inhomogeneities (30% of the total). The cases in which data from external sites were used to fill gaps in the time series, both at the early stages of this work and by the original sources (see section 2) have also been considered as a type of relocation in the comparison of inhomogeneities and metadata. Relocation can lead to inhomogeneities as a result of the difference in annual precipitation in both the old and the new position as well as differences in sheltering conditions. In 70% of the inhomogeneities identified as relocations an increase of catchment (factor  $> 1$ ) has been produced. However, the mean of all correction factors in Fig. 4c does not depart significantly from one, therefore, there is no bias in the occurrence of inhomogeneities either to increase or decrease precipitation means on the whole dataset. This is due to the fact that unknown changes (crosses in Fig. 4c) have led to a decrease in catchment (factor  $< 1$ ) in 66% of the cases. Nevertheless, changes in specific sites have played a decisive role as will be shown later.

The three changes in height (triangles) seem to agree somehow in the type of inhomogeneity. The two first ones occur in La Coruña (1916) and San Sebastián, Spain (1918) and reveal an increase in catchment related to a new position of the pluviometer in a lower and probably more sheltered location. The third one (Gibraltar in 1935) corresponds to a change of the pluviometer to a higher elevation, in principle less sheltered, and thus favoring a decrease in catchment.

Some of the inhomogeneities lie on the 1.0 value line. All of them correspond to cases in which two inhomogeneities were detected, but only the data between them had to be corrected, thus in such cases the first inhomogeneity is identified in Fig. 4c with value 1.0.

Finally it is also worth pointing out that 80% of the inhomogeneities occur before 1950. Such a bias is also present in the histogram of the metadata (white bars). Though the metadata archives are incomplete they can be taken as an indication that there have been fewer changes in the more recent decades. This supports the smaller amount of inhomogeneities after 1950.

Figure 5 gives some typical examples of the application of the Standard Normal Homogeneity Test to several types of time series showing values for the  $T$  statistic [Eq. (4)] and the series of ratios  $q_i$  [Eq. (2)]. Figure 5a displays results for Murcia, Spain (station number 74), classified as homogeneous in the fourth step of the procedure (H4) and shows no significant changes in the  $T$  statistic; there is a secondary maximum in 1955 that seems to agree with a change in the position of the rain gauge registered in the metadata. The series of ratios,  $q$ , does not show apparent displacements from the mean value. Palma de Mallorca, Spain (Fig. 5b), is an example of homogeneous after one correction series

(HC4). In step 2 of the procedure (dots), the test reveals a wide  $T$  maximum centered in 1940, possibly related to a change in the position of the rain gauge in 1937. This maximum disappears after the correction factor is applied to the values before 1937. The  $q$  values also reveal this inhomogeneity: noncorrected values in dots cluster above (below) 1 for most years before (after) the change while for the corrected ones (continuous line) this steplike behavior is reduced.

Another example (Fig. 5c) of homogeneous series after applying one correction is La Coruña (number 19). In this case the test statistic shows several significant maxima. The most important one takes place in 1915 and agrees with a change in the height of the rain gauge; the second peak occurs in 1942 and does not match with any entry in the historical records. The first peak disappears and the second one is not significant after applying the correction factor to the data before 1915. The  $q$  series shows some displacement from the mean value before 1915 (dots in the oval), which also disappears after the correction is applied. Hervás (number 56), one of the series reconstructed for the dataset, is yet another example of HC4-type series. The  $T$  statistic shows (Fig. 5d) a low-amplitude maximum in 1983 not supported by the available metadata that drops out in step 4 of the procedure.

Bilbao, Spain (number 20), described by Fig. 5e is an example of a series that became homogeneous after two corrections. The values of the  $T$  statistic in the second step (hollow dots) show a maximum centered in 1946 that agrees with a reference in the metadata in 1945. After the correction of this inhomogeneity, a second maximum rises over the significance level in the fourth step (solid dots). This peak agrees with a change in the position of the rain gauge in 1920. In the sixth step both inhomogeneities are corrected and the  $T$  statistic shows no important peaks. The  $q$  values also show these changes: it can be noticed how noncorrected data (dots) depart more strongly from unity than the corrected ones (solid line), particularly during the period between both inhomogeneities (highlighted with the oval).

Finally, results for Alange, Spain (number 69), are shown also in Fig. 5f. This is the only station which remained inhomogeneous after two corrections (ICC). The  $T$  values exhibit significant maxima in years 1934, 1917, and 1955 corresponding to the steps 2, 4, and 6 of the procedure, respectively. The series of  $q$  ratios show an appreciable displacement from the mean value only in step 2 (hollow dots in the oval). The inhomogeneity in 1917 is supported by only a single peak value and the maximum in step 6 hardly reaches the  $T_{95}$  significance level. This accounts for the low magnitude of these inhomogeneities and their lack of evidence in the series of ratios for steps 4 and 6. The three inhomogeneities are probably related to data interpolations from the secondary sites.

The correction factors deduced from the analysis of

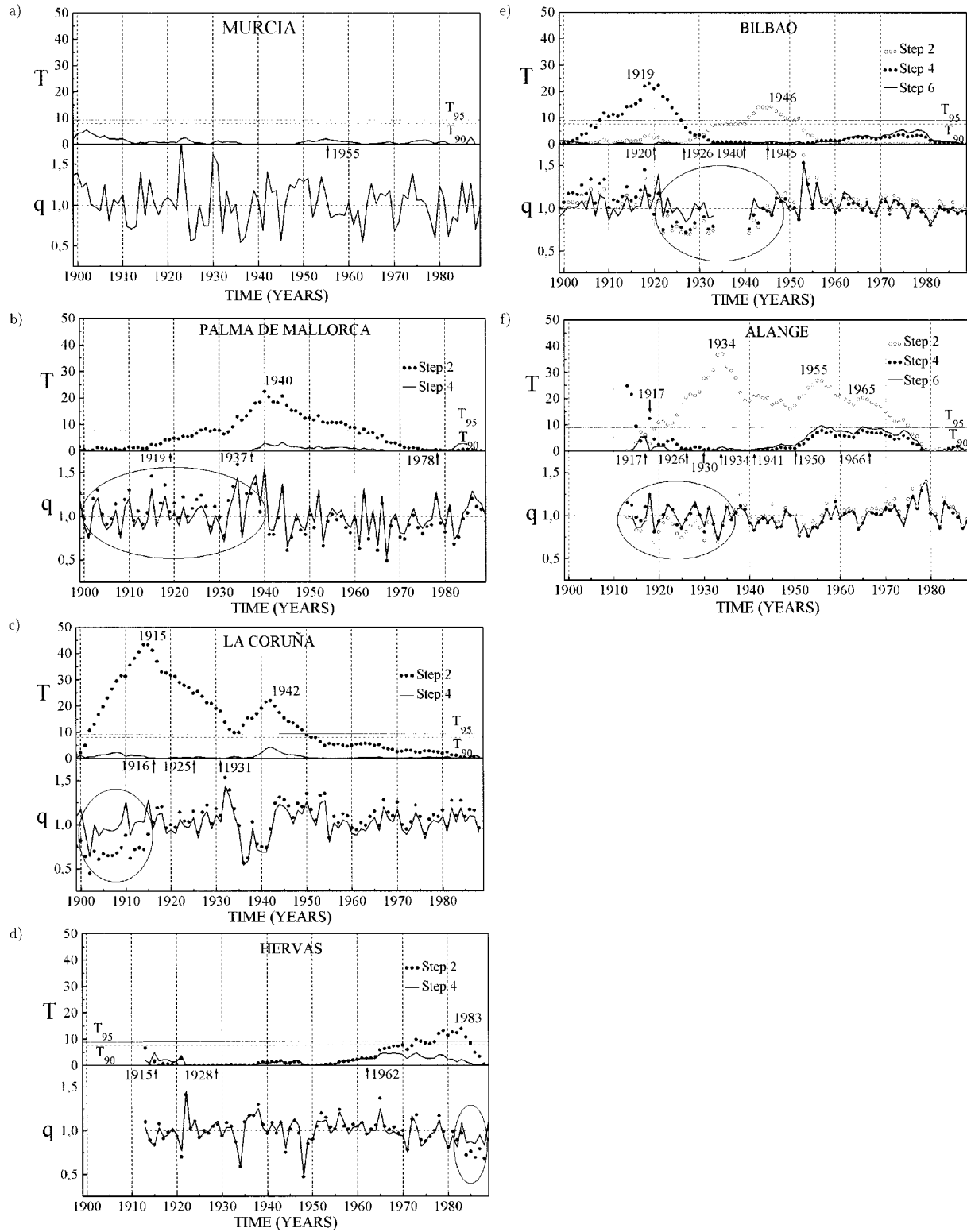


FIG. 5. Values of  $T$  and  $q$  for (a) Murcia, (b) Palma de Mallorca, (c) La Coruña, (d) Hervás, (e) Bilbao, and (f) Alange. Dates of metadata are indicated with arrows in the horizontal axis. Here  $q$  data before the inhomogeneities are highlighted in the oval.

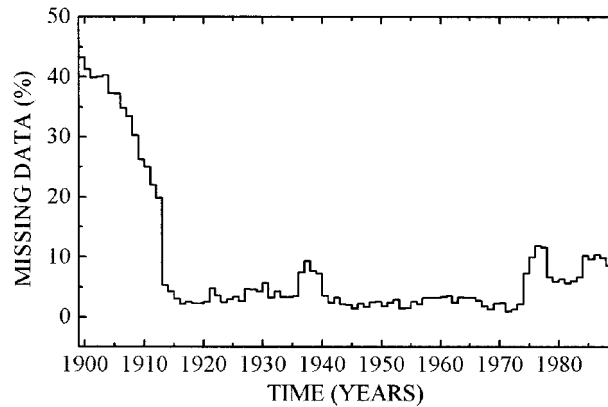


FIG. 6. Temporal distribution of the annual percentage of missing data.

the annual data were applied to the monthly data and interpolation for missing values was also accomplished following Eqs. (6) and (7). Figure 6 shows the temporal distribution of missing values in the dataset. The highest percentages are reached at the beginning of the century due to the lower amount of stations at this time. There is a secondary maximum in the 1970s and 1980s due to the shut down of measurements in some stations. A small maximum close to 1940 related to the Spanish Civil War (1936–39) can also be appreciated. After the interpolation was accomplished Camporredondo, Spain (number 25), and Algiers, North Africa (number 86), were eliminated from the dataset because of the long data gaps present and neighboring stations that can describe the local climate features were available in both cases.

### c. Trends

This section provides a description of some effects of the adjustments. The effects of outliers adjustments on trends and on the homogenization procedure are described as well as the effects of homogenization corrections on trends. A more complete description of trends, including a discussion of connections to general circulation changes, is beyond the scope of the present work. Some results with the same dataset are presented by González-Rouco et al. (2000).

Trends are obtained through the calculation of least square linear regressions of precipitation data against time. A significance  $t$  test (Snedecor and Cochran 1972) for constant precipitation is applied. The effective sample size and autocorrelation in the residuals are also taken into account. As precipitation time series often show high autocorrelation, values are not independent of each other and the effective number of data has also been used in the testing of significance (Trenberth 1984; Zwiers and von Storch 1995). Also, residuals resulting from the least square regression are usually assumed to be white noise, but this is not often the case. When

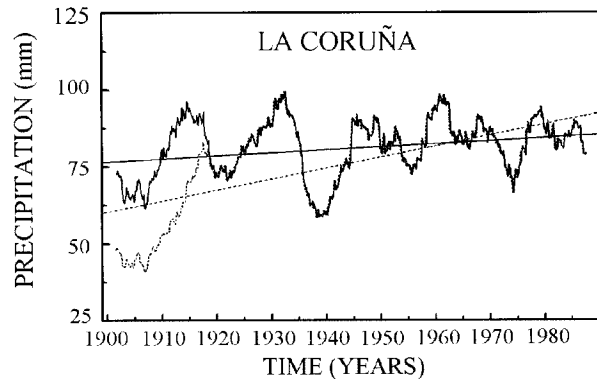


FIG. 7. Trends and outputs of a centered 5-yr moving average for monthly precipitation values in La Coruña before (dashed line) and after (solid line) the homogeneity correction.

residuals show autocorrelation both the slope and its error are badly estimated. We have used a Durbin–Watson method (Durbin and Watson 1971; Valero et al. 1996a) to account for the existence of autocorrelation in the residuals. An AR(1) model has been assumed for the errors of the least square regression whenever significant autocorrelations were detected.

Figure 7 shows the effect of the homogeneity correction for the time series from La Coruña. The solid lines correspond to a low-pass filter output and a linear fit to the corrected data; the dashed line is similarly obtained from the original series. The linear fit to noncorrected data shows a significant ( $\alpha = 0.05$ ) slope of  $3.6 \text{ mm decade}^{-1}$  that is reduced to a nonsignificant value of  $0.9 \text{ mm decade}^{-1}$  after the correction.

Figure 8 shows trends for the dataset in different stages: for the original data, without corrections (Fig. 8a), after applying outliers corrections (Fig. 8b), after applying outliers and homogeneity corrections (Fig. 8c), and after interpolating data in the homogenized dataset for the period 1899–1989. In Figs. 8a–c results for each time series correspond only to the period of existence of data within the time interval 1899–1989 (i.e., Grenoble, France: 1899–1976, see Table 1), therefore, interpreting spatial structures from these plots should be done with caution since the time span of the time series can change from site to site.

In the noncorrected data (Fig. 8a) 30% of the trends are significant. In general, positive trends seem to dominate, particularly in the north of the Iberian Peninsula. Figure 8b shows trends after the corrections of outliers were applied. Nearly the same pattern is obtained, except in some specific cases where changes in the magnitude and significance of slopes arise: Perpignan, France (number 30), Palencia (39) and Leon, Spain (32), as the only change in sign, and Castellón, Spain (50).

In order to check the influence of outliers corrections in the homogenization procedure, step 1 in Fig. 2 was rerun with the original data without outliers corrections. The test results were different in eight cases: stations

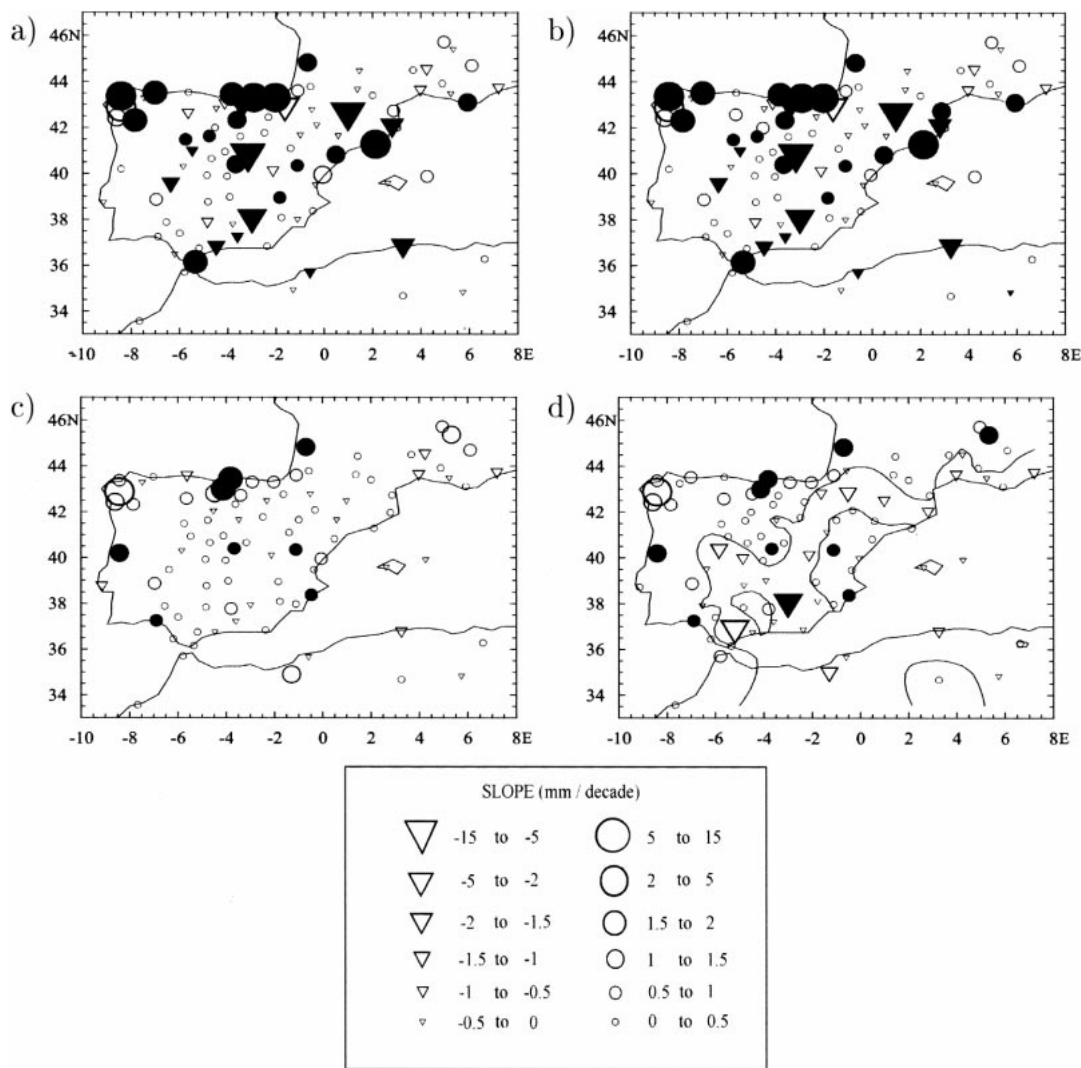


FIG. 8. Trends in the (a) original dataset, (b) after correcting outliers, (c) after correcting outliers and inhomogeneities. Annual means for the periods of existence of data in each station have been used for the calculations. (d) Trends in the corrected dataset after interpolating missing data for the period 1899–1989 (inputs were also annual means). A solid line for value zero is plotted for easy visualization of different regions.

38, 43, 47, 71, and 84 switched from inhomogeneous to homogeneous and 72, 87, and 94 from homogeneous to inhomogeneous. In all these cases  $T$  values were very close to the significance level. If outliers were not trimmed the series were pushed to the significance or nonsignificance side depending on the specific characteristics of the time series. Outliers cause an increase in the mean of the period in which they occur. If this rise in mean sharpens (shortens) the differences between the period before and after a given change, it will increase (decrease) the  $T$  values and the significance of the change. The authors consider that it is safer to take decisions upon the quality of the time series that depend mostly on its overall characteristics rather than on the extreme behavior of a few data. This is also valid for the calculation of trends, since we are interested in re-

sults that describe the long-term behavior of the time series and are not so sensitive to individual data. Therefore we think that the use of trimmed data is justified.

Figure 8c shows the trends after the data have been corrected for outliers and inhomogeneities. If we compare it to Fig. 8b we can describe the changes as dramatic. Significance is reached in only eight sites and the picture is more uniform. In most sites the value of trends is negligible and there are almost no negative trends left; positive trends in the north of the Iberian Peninsula still remain.

Figure 8d shows trends of the corrected and interpolated data for the period 1899–1989. A solid zero line is plotted for easier visualization of regions of different sign. Though significance is reached only at a few stations, a highly coherent spatial behavior arises in this

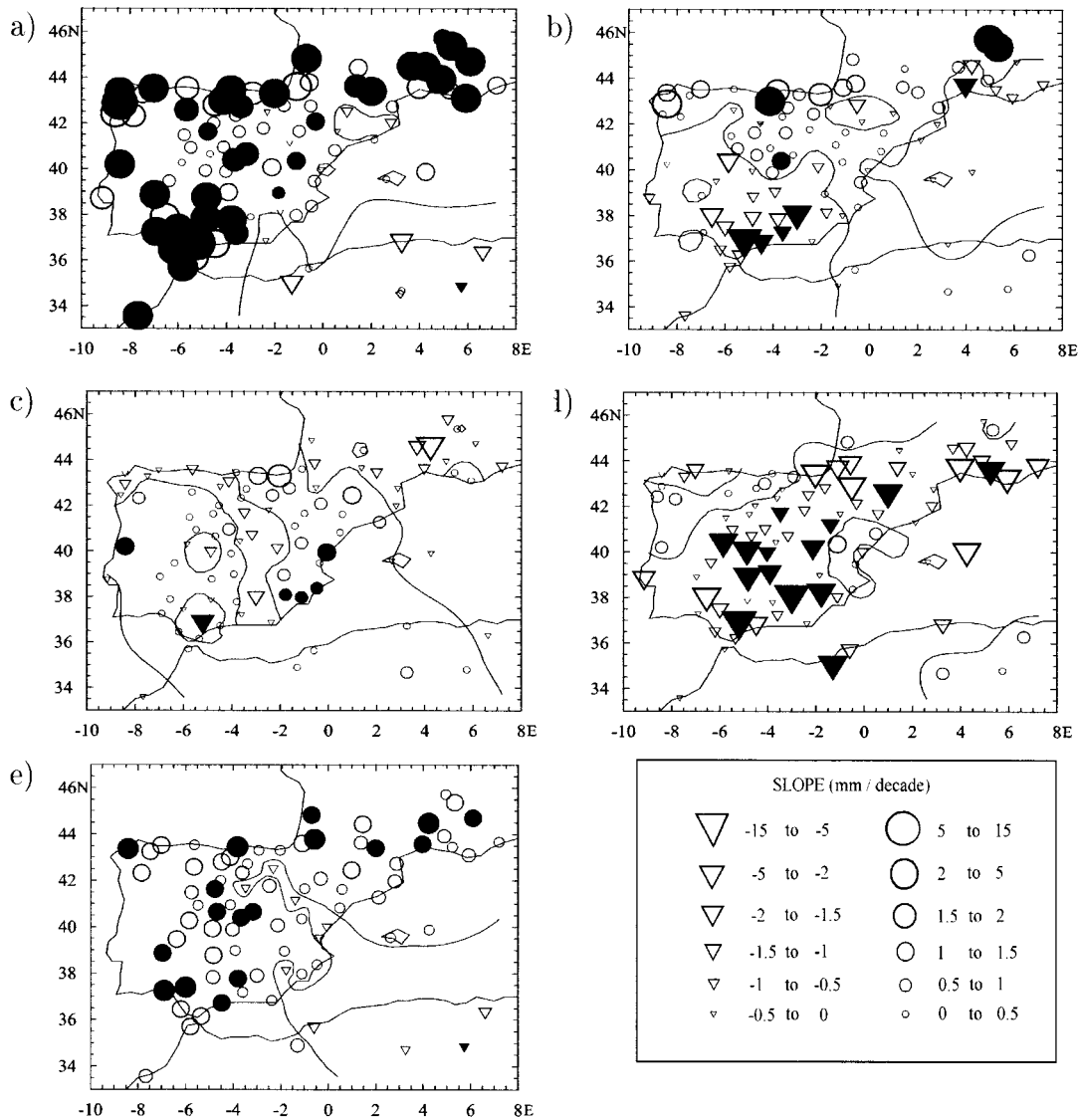


FIG. 9. Trends in the final dataset for (a) winter, (b) spring, (c) summer, and (d) fall. (e) Winter trends for the period 1920–89. Monthly data for each season were used as inputs. Full symbols highlight significant trends (95% significance level). A solid line for value zero is plotted for easy visualization of different regions.

plot. All stations close to the Atlantic coast show positive trends as do most French stations. On the east coast of the Iberian Peninsula, positive values also arise. The Pyrenees, the French Mediterranean coast, north of Africa, and quite a few stations in the middle and south of Spain show negative values.

These annual trends result from the cumulative effects of seasonal trends. Some insight into the latter is obtained from Fig. 9. Slopes have been obtained using monthly anomalies for each season (winter: December–February; spring: March–May; summer: June–August; autumn: September–November). Anomalies were calculated by subtracting the long-term monthly means to filter out the mean annual cycle.

Trends in winter (Fig. 9a) show positive values over

nearly all the area except for the north of Africa. Significance is reached in 40% of the trends. In spring (Fig. 9b) the positive trends lose significance in the northern half of the Iberian Peninsula and give way to negative trends in the southern half. In the summer (Fig. 9c) the pattern alternates areas of positive and negative values. In autumn (Fig. 9d), negative, often significant, precipitation trends are obtained nearly everywhere except for some sites on the east coast and in the northwest. The decrease of precipitation in autumn and spring and increase in winter indicates strengthening of the annual cycle. These results agree with those of other authors that report a significant downward trend during this century, both in the intensity of the zonal circulation in the North Atlantic and in the amplitude of its annual cycle,

these being ascribed by changes in the North Atlantic oscillation (NAO) index (Lamb and Pepler 1987; Hurrell 1995). When the NAO index annual amplitude is less than normal, that is, when smaller autumn–winter and/or greater NAO for summer is observed (Makrogiannis et al. 1991) the amplitude of the annual wave in precipitation in the Iberian Peninsula increases (Valero et al. 1996b).

It is worth remarking on the possible influence of missing data on trends. Figure 6b shows that changes in the early part of the century can have a considerable influence on trends. According to Fig. 6a, 40% of the data are missing during the first decade of the century and are interpolated using information from the existing 60%. There is a possibility that this interpolation biases the magnitude and sign of the trends toward spatial continuity. This has been explored through calculating trends for the period 1920–89, thus avoiding the problem of long missing data gaps in some stations. Figure 9e shows the result for the wintertime. Though the magnitude of trends is not so high as in Fig. 9a, regional agreement is still the general feature. The magnitude of trends is smaller than in Fig. 9a, but this could also be due to natural variability. In addition to the interpolation method we suggest that the homogenization procedure could influence some regional agreement. The fundamental hypothesis in the Standard Normal Homogeneity Test states that the ratios between a test station and some regional precipitation average should be constant. It is possible that this assumption leads to removing distinct local trends in stations imbedded in microclimates and where changes in long-term trends are not necessarily related to homogeneity problems.

## 5. Discussion and conclusions

Outliers and inhomogeneities have been taken into account as potential quality problems in a dataset of 95 precipitation time series in southwestern Europe.

Outliers adjustments have been carried out by trimming very extreme data with the aim of reducing large distribution tails as a means of preprocessing previous to homogeneity corrections. There are, however, caveats to the use of this subjective approach since it involves dismissing interesting climatological information about extreme events and their temporal and spatial distribution. Nevertheless, it is worth noting that this does not mean a “point of no return” since the original value of the extreme data can be returned to the dataset at any time, also incorporating the homogeneity corrections.

The outliers adjustments applied herein lead to changes of minor importance in trend assessment as shown in the previous section. Concerning the impact of outliers adjustments in the homogeneity test, this becomes evident in cases where the  $T$  values are very close to the significance level. In such situations the effect of a few extreme values can lead to significance or not significance of an inhomogeneity, depending on the case.

Though it might seem contradictory to support corrections that have such a limited effect, it is the authors' opinion that the use of trimmed data is legitimate because decisions concerning trends and especially homogeneity are then taken with regard to the long-term behavior of the series rather than the marginal extreme properties of a few values.

Analysis of the homogeneity of the time series found 40% were homogeneous, an additional 49.5% were homogeneous after applying one correction factor and 9.5% more after applying two correction factors. In 30% of the cases the inhomogeneities could be linked to some historical change in the rain gauge, in spite of the metadata being scarce. Examples have been presented that provide evidence that precipitation records from the area of study show severe homogeneity problems. There is considerable influence of homogeneity corrections in the estimation of trends, emphasizing that the results of any long-term variability study from this area are very dependent upon the use of homogenized or nonhomogenized data.

Trends for the corrected data present great regional agreement. Winter values show positive and usually significant trends over most of the study area. Significant negative trends show up mainly in autumn. These trends seem to agree with those described in the literature for the zonal circulation in the North Atlantic (Makrogiannis et al. 1991; Valero et al. 1996b). A brief description of trends in relationship to the outliers and homogeneity adjustments made on the dataset has been provided. A deeper insight into long-term trends and their relationship to general circulation is beyond the scope of this work.

The Standard Normal Homogeneity Test is based on the hypothesis that the ratios between the test-station values and some regional precipitation average should be constant. It is possible that this underlying assumption leads to removing distinct local trends in stations embedded in microclimates and not necessarily related to inhomogeneity problems. It should also be pointed that the output of such a quality control procedure is not expected to be an error-free dataset, but is one where the quality of the time series has been significantly improved. On the other hand, different methodologies can lead to partially different results, since not only the error associated with any homogeneity test, but also the various methods and criteria that can be chosen for the selection of reference stations add uncertainty. Different countries and institutions have chosen methodologies that specifically suit their purposes (Peterson et al. 1998b). In this sense, work intercomparing different methods is desired and still to come.

*Acknowledgments.* This work was supported by Project Number CLI97-0558, Comunidad de Madrid, and Fundación Ramón Areces. The authors also thank P. Brandt, P. Cuesta, M. Widmann, and E. Zorita for their comments on the original version of the manuscript. We

thank also F. Zwiers and the two anonymous reviewers for their valuable comments and suggestions.

## REFERENCES

- Alexanderson, H. A., 1986: A homogeneity test applied to precipitation data. *J. Climatol.*, **6**, 661–675.
- Almarza, C., J. A. López, and C. Flores, 1996: Homogeneidad y variabilidad de los registros históricos de precipitación de España. Tech. Rep. A-143, Instituto Nacional de Meteorología, Madrid, Spain, 318 pp. [Available from Instituto Nacional de Meteorología, Madrid, Spain.]
- Barnett, V., and T. Lewis, 1994: *Outliers in Statistical Data*. 3d ed. J. Wiley and Sons, 604 pp.
- Buishand, T. A., 1982: Some methods for testing the homogeneity of rainfall records. *J. Hydrol.*, **58**, 11–27.
- Conrad, V., and C. Pollak, 1950: *Methods in Climatology*. 2d ed. Harvard University Press, 459 pp.
- Durbin, J., and G. S. Watson, 1971: Testing for serial correlation in least squares regression. III. *Biometrika*, **37**, 1–19.
- Eischeid, J., C. B. Baker, T. Karl, and H. F. Díaz, 1995: The quality control of long-term climatological data using objective data analysis. *J. Appl. Meteor.*, **34**, 2787–2795.
- Esteban-Parra, M. J., F. S. Rodrigo, and J. Castro-Diez, 1998: Spatial and temporal patterns of precipitation in Spain for the period 1880–1992. *Int. J. Climatol.*, **18**, 1557–1574.
- Fernández-Mills, G., 1995: Principal component analysis of precipitation and rainfall regionalization in Spain. *Theor. Appl. Climatol.*, **50**, 117–131.
- Filippov, V. V., 1968: Quality control procedures for meteorological data. Tech. Rep. 26, WMO, Geneva, Switzerland, 38 pp.
- Garrido, J., J. A. García, and V. L. Mateos, 1996: Homogeneidad y variabilidad natural de series largas de precipitación. *An. Fis.*, **92**, 19–31.
- González-Rouco, J. F., H. Heyen, E. Zorita, and F. Valero, 2000: Agreement between observed rainfall trends and climate change simulations in the southwest of Europe. *J. Climate*, **13**, 3057–3065.
- Grant, E. L., and R. S. Leavenworth, 1972: *Statistical Quality Control*. McGraw-Hill, 764 pp.
- Hanssen-Bauer, I., and J. Förlund, 1994: Homogenizing long Norwegian precipitation series. *J. Climate*, **7**, 1001–1013.
- Hawkins, P. M., 1977: Testing a sequence of observations for a shift in location. *J. Amer. Stat. Assoc.*, **72**, 180–186.
- Hegerl, G. C., H. von Storch, K. Hasselmann, B. D. Santer, U. Cubasch, and P. D. Jones, 1996: Detecting anthropogenic climate change with an optimal fingerprint method. *J. Climate*, **9**, 2281–2306.
- Houghton, J., B. A. Callendar, and S. K. Varney, Eds., 1992: *Climate Change: The Supplementary Report to the IPCC Scientific Assessment*. Cambridge University Press, 205 pp.
- , L. G. Meira Filho, B. A. Callander, N. Harris, A. Kattenberg, and K. Maskell, Eds., 1995: *Climate Change 1995: The Science of Climate Change*. Cambridge University Press, 572 pp.
- Hurrell, J. W., 1995: Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, **269**, 676–679.
- Jones, P., and Coauthors, 1985: A grid point surface air temperature data set for the northern hemisphere. Tech. Rep. TRO22, U.S. Department of Energy, Carbon Dioxide Research Division, Washington, DC, 251 pp.
- Karl, T. R., and C. N. Williams, 1987: An approach to adjusting climatological time series for discontinuous inhomogeneities. *J. Climate Appl. Meteor.*, **26**, 1744–1763.
- Lamb, P. J., and R. Pepler, 1987: North Atlantic Oscillation: Concept and application. *Bull. Amer. Meteor. Soc.*, **68**, 1218–1225.
- 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. *Int. J. Climatol.*, **16**, 1197–1226.
- Linés, A., 1970: The Climate of the Iberian Peninsula. *Climates of Northern and Western Europe*, C. Wallen, Ed., *World Survey of Climatology*, Vol. 5, Elsevier, 195–239.
- Maheras, P., 1988: Changes in precipitation conditions in the western Mediterranean over the last century. *J. Climatol.*, **8**, 179–189.
- Makrogiannis, T. J., H. S. Sahsamanoglou, A. A. Flocas, and A. Bloutsos, 1991: Analysis of the monthly zonal index values and longterm changes in the North Atlantic and Europe. *Int. J. Climatol.*, **11**, 493–503.
- Mitchell, J. M., B. Dzerdzeevski, H. Flohn, W. L. Hofmeyr, H. H. Lamb, K. N. Rao, and C. C. Wallen, 1966: Climate change. Tech. Rep. 79, WMO, Geneva, Switzerland, 79 pp.
- Peterson, T. C., R. Vose, R. Schmoyer, and V. Razuvaev, 1998a: Global Historical Climatology Network (GHCN) quality control of monthly temperature data. *Int. J. Climatol.*, **18**, 1169–1179.
- , and Coauthors, 1998b: Homogeneity adjustments of *in situ* atmospheric climate data: A review. *Int. J. Climatol.*, **18**, 1493–1517.
- Rhoades, D. A., and M. J. Salinger, 1993: Adjustment of temperature and rainfall records for site changes. *Int. J. Climatol.*, **13**, 899–913.
- Rodriguez-Puebla, C., A. H. Encinas, S. Nieto, and J. Garmendia, 1998: Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula. *Int. J. Climatol.*, **18**, 299–316.
- Santer, B. D., K. E. Taylor, T. M. Wigley, J. E. Penner, P. D. Jones, and U. Cubasch, 1995: Towards the detection and attribution of an anthropogenic effect on climate. *Climate Dyn.*, **12**, 77–100.
- Serrano, A., J. A. García, V. L. Mateos, M. L. Cancillo, and J. Garrido, 1999: Monthly modes of variation of precipitation over the Iberian peninsula. *J. Climate*, **12**, 2894–2919.
- Snedecor, G. W., and W. G. Cochran, 1972: *Statistical Methods*. 6th ed. Iowa University Press, 593 pp.
- Trenberth, K. E., 1984: Some effects of finite sample size and persistence on meteorological statistics. Part I: Autocorrelations. *Mon. Wea. Rev.*, **112**, 2359–2368.
- , and D. A. Paolino, 1980: The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities. *Mon. Wea. Rev.*, **108**, 856–872.
- Valero, F., J. A. García-Miguel, M. L. Martín, and Y. Luna, 1993: Detecting long-term variations in precipitation in the southeastern Iberian Peninsula. *Ann. Geophys.*, **11**, 837–845.
- , J. F. González, F. J. Doblas, and J. A. García-Miguel, 1996a: A method for the reconstruction and temporal extension of climatological time series. *Int. J. Climatol.*, **16**, 213–227.
- , F. J. Doblas, J. F. González, and J. A. García-Miguel, 1996b: On long-term evolution of seasonal precipitation in southwestern Europe. *Ann. Geophys.*, **14**, 976–985.
- , R. M. Rodríguez, and J. A. García-Miguel, 1997a: Filling data algorithms in urban air pollution monitoring networks. *J. Environ. Eng. Div. ASCE*, **103**, 395–403.
- , M. Y. Luna, and M. Martín, 1997b: An overview of a heavy rain event in Southeastern Iberia: The role of large-scale meteorological conditions. *Ann. Geophys.*, **15**, 494–502.
- Zorita, E., V. Kharin, and H. von Storch, 1992: The atmospheric circulation and sea surface temperature in the North Atlantic area in winter: Their interaction and relevance for Iberian precipitation. *J. Climate*, **5**, 1097–1108.
- Zwiers, F. W., and H. von Storch, 1995: Taking serial correlations into account in tests of the mean. *J. Climate*, **8**, 336–351.