



Homogenization of Portuguese long-term temperature data series: Lisbon, Coimbra and Porto

A. L. Morozova¹ and M. A. Valente²

¹Centro de Geofísica da Universidade de Coimbra, University of Coimbra, Portugal

²Instituto Dom Luiz, University of Lisbon, Portugal

Correspondence to: A. L. Morozova (anna_m@teor.fis.uc.pt)

Received: 21 June 2012 – Published in Earth Syst. Sci. Data Discuss.: 25 July 2012

Revised: 13 November 2012 – Accepted: 14 November 2012 – Published: 14 December 2012

Abstract. Three long-term temperature data series measured in Portugal were studied to detect and correct non-climatic homogeneity breaks and are now available for future studies of climate variability.

Series of monthly minimum (T_{\min}) and maximum (T_{\max}) temperatures measured in the three Portuguese meteorological stations of Lisbon (from 1856 to 2008), Coimbra (from 1865 to 2005) and Porto (from 1888 to 2001) were studied to detect and correct non-climatic breaks. These series, together with monthly series of average temperature (T_{aver}) and temperature range (DTR) derived from them, were tested in order to detect breaks, using firstly metadata, secondly a visual analysis, and thirdly four widely used homogeneity tests: von Neumann ratio test, Buishand test, standard normal homogeneity test, and Pettitt test. The homogeneity tests were used in absolute (using temperature series themselves) and relative (using sea-surface temperature anomalies series obtained from HadISST2.0.0.0 close to the Portuguese coast or already corrected temperature series as reference series) modes. We considered the T_{\min} , T_{\max} and DTR series as most informative for the detection of breaks due to the fact that T_{\min} and T_{\max} could respond differently to changes in position of a thermometer or other changes in the instrument's environment; T_{aver} series have been used mainly as control.

The homogeneity tests showed strong inhomogeneity of the original data series, which could have both internal climatic and non-climatic origins. Breaks that were identified by the last three mentioned homogeneity tests were compared with available metadata containing data such as instrument changes, changes in station location and environment, observation procedures, etc. Significant breaks (significance 95 % or more) that coincided with known dates of instrumental changes were corrected using standard procedures. It was also noted that some significant breaks, which could not be connected to known dates of any changes in the park of instruments or stations location and environment, were probably caused by large volcanic eruptions. The corrected series were again tested for homogeneity; the corrected series were considered free of non-climatic breaks when the tests of most of monthly series showed no significant (significance 95 % or more) breaks that coincide with dates of known instrument changes. Corrected series are now available within the framework of ERA-CLIM FP7 project for future studies of climate variability (doi:10.1594/PANGAEA.785377).

1 Introduction

Long instrumental climatological records assume a paramount role in the studies of variation of the atmospheric conditions. They provide vital information about climate variability, trends and cycles. Unfortunately, long-term series often contain inhomogeneities caused by a

number of non-climatic factors that could provide unrealistic trends, shifts and jumps (Peterson et al., 1998; Aguilar et al., 2003). These inhomogeneities are originated by changes in instruments, station locations and surrounding environment, observation routines and methods of preliminary data treatment. Undoubtedly, such inhomogeneities have to be

detected and corrected beforehand, and only after that could the data series be used in any kind of climate studies.

The problem of identification and correction of non-climatic inhomogeneities has been studied thoroughly (see, e.g. review in Peterson et al., 1998). The simplest way to detect the shift-like inhomogeneities is a visual analysis, preferably by an experienced meteorologist (Peterson et al., 1998). It is clear that this method is very subjective and could be used as an initial part of the analysis, providing information about “doubtful” periods that have to be studied thoroughly with other objective methods.

At the moment, there exist a lot of objective statistical methods accepted by the scientific community that can detect the presence and probable date of inhomogeneities, and new methods continue to be developed (see e.g. Venema et al., 2012). Most of these methods belong to one of three groups: likelihood-based methods, linear-regression based methods, and non-parametric methods (Wang et al., 2007). In climate studies, the most commonly used methods are the standard normal homogeneity test (SNHT; Alexandersson and Moberg, 1997) and its variations, the Buishand cumulative deviation test (Buishand, 1982), the non-parametric rank Pettitt test (Pettitt, 1979), the two-phase regression methods (e.g. Solow, 1987) and others. These methods estimate not only the level of inhomogeneity of the tested series, but also detect the highly probable homogeneity break points (hereafter: breaks). The other tests, like the von Neumann ratio test (von Neumann, 1941), do not give any information about the date of the break, but estimate the overall level of inhomogeneity in the data.

The tasks of non-climatic breaks correction are complicated by the fact that not all inhomogeneities existing in data series are of non-climatic origin. There are breaks that originate from “real” climate changes, like volcanic aerosol ejections or abrupt changes of atmospheric and/or oceanic circulation. The non-climatic inhomogeneities have to be somehow separated from the others, and this task could be done using the metadata – a record of station relocations, changes in station environment, changes in the instrument park, observation routines, applications of new formulae to calculate means, etc. The metadata could provide precise information about the dates and reasons for non-climatic changes and consequently ideal for use in any homogenization procedure. Moreover, all available information about stations’ history should be preferred over statistical methods, especially in the tasks of detection of the breaks dates (Venema et al., 2012). Any break detected by statistical methods had to be checked against metadata, and if there is a written note that some intervention took place in the station setup at the break date, this break should be considered as non-climatic and (in most cases) be corrected (Peterson et al., 1998; Aguilar et al., 2003).

The analysis of the separate monthly series could provide different break points for each month, both due to the randomness of the meteorological time series and to the fact

that some inhomogeneities could have larger effect during the warm part of the year than during the cold part. Therefore, not only annual but also monthly (or seasonal) means have to be analysed in the process of homogenization (Aguilar et al., 2003).

The detected non-climatic inhomogeneities required correction. The correction procedure was constructed so that all data were corrected in line with the conditions of the last homogeneous part of the data series: a period ranging from the last break to the end of the series. In this case, all future periods of the incoming data would not damage the homogeneity of the data series (Aguilar et al., 2003). The procedure of correction is applied to the data series backward in time, starting from the most recent break. The most usual way to correct non-climatic breaks is to calculate the means of the studied parameter during some time before and after each of the breaks. The adjustment value is then a difference (or ratio in case of parameters like precipitation) between these means. Accordingly, the adjustment value is applied to the inhomogeneous part (part before the break) of the series (Aguilar et al., 2003).

2 Methods for detection and correction of non-climatic breaks

2.1 Homogeneity tests

Despite the fact that the main role in the detection of the breaks was assigned in this study to the metadata, four simple, widely used statistical homogeneity tests were applied to the data (Klein Tank, 2007): von Neumann ratio test, Buishand test, standard normal homogeneity test (SNHT), and Pettitt test. The first test allows one to estimate only the presence of breaks in the dataset, whereas the last three tests also give information about the possible dates of such breaks. The use of tests of different types (parametric, non-parametric, likelihood), which also have different sensitivities in different parts of the series, could help to obtain more significant results.

Three of these tests (Buishand test, SNHT and Pettitt test) were used both in absolute mode – statistical analysis of the temperature series themselves and in relative mode – statistical analysis of the temperature series using (for coastal stations) monthly anomalies (relative to the 1961–1990 period) of sea-surface temperature (SST) obtained with HadISST2.0.0.0 in the 2 grid points nearest the Portuguese coastal stations (Rayner et al., 2012) or already corrected temperature series for the non-coastal station (Coimbra). To perform the relative homogeneity tests, the temperature and SST series were standardised (transformed to series with mean of 0 and standard deviation of 1); afterwards, the differences between temperature and SST anomalies were calculated and subjected to homogeneity tests.

Table 1. The 90 %, 95 % and 99 % critical values for the following homogeneity test statistics: N of the von Neumann ratio test, T_0 of the SNHT, X_K of the Pettitt test and Q of the Buishand partial sum tests, for data sets with different lengths (114, 141 and 153 elements).

Test	114			141			153		
	90 %	95 %	99 %	90 %	95 %	99 %	90 %	95 %	99 %
von Neumann (N)	1.76	1.7	1.57	1.79	1.72	1.61	1.79	1.74	1.63
SNHT (T_0)	7.9	9.3	12.4	8.0	9.5	12.6	8.1	9.5	12.7
Pettitt (X_K)	757	864	1071	1041	1187	1472	1176	1341	1663
Buishand (Q)	12.5	13.8	16.5	14.0	15.3	18.3	14.6	15.9	19.0

2.1.1 Von Neumann ratio test – non-parametric test

In this test the null hypothesis is that the data are independent identically distributed random values; the alternative hypothesis is that the values in the series are not randomly distributed. The von Neumann ratio N is defined as the ratio of the mean square successive (year to year) difference to the variance (von Neumann, 1941):

$$N = \sum_{i=1}^{n-1} (Y_i - Y_{i+1})^2 / \sum_{i=1}^n (Y_i - \bar{Y})^2. \quad (1)$$

Hereafter, for each of the test descriptions, n is the data set length, Y_i is i -th element of the data set, \bar{Y} is the mean value of the data set. When the sample is homogeneous the expected value is $N = 2$. If the sample contains a break, then the value of N tends to be lower than this expected value (Buishand, 1981). If the sample has rapid variations in the mean, then values of N may rise above two (Klein Tank, 2007). This test gives no information about the location of the shift. The critical values for N (for $n \geq 20$), with probability level α , are defined as

$$N_\alpha \approx 2 - 2u_\alpha \sqrt{\frac{n-2}{(n-1)(n+1)}}, \quad (2)$$

where u_α is the α -th percentile of a standard normal variate from the standard normal table (Buishand, 1981). Critical values for N for different data set lengths are given in Table 1. It should be mentioned that in case of a number of data sets with similar breaks and similar level of variations of the mean, the data set with smaller standard deviation has smaller N as well (see eq. 3 in Buishand, 1981). This means that annually averaged parameters should have smaller N values than monthly averaged ones.

2.1.2 Buishand test – parametric test

This test supposes that tested values are independent and identically normally distributed (null hypothesis). The alternative hypothesis assumes that the series has a jump-like shift (break). This test is more sensitive to breaks in the middle of time series (Costa and Soares, 2009). The test statistics, which are the adjusted partial sums (Buishand, 1982), are

defined as

$$S_k^* = n \sum_{i=1}^k (Y_i - \bar{Y}) / \sum_{i=1}^n (Y_i - \bar{Y})^2, \quad k = 1 \dots n \quad (3)$$

$$S_0^* = 0 \quad (4)$$

When series are homogeneous, the values of S_k^* will fluctuate around zero because no systematic deviations of the Y_i values with respect to their mean will appear.

Q-statistics: if a break is present in year K , then S_k^* reaches a maximum (negative shift) or minimum (positive shift) near the year $k = K$.

$$Q = \max_{0 \leq k \leq n} S_k^* \quad (5)$$

R-statistics (range statistics) are

$$R = \max_{0 \leq k \leq n} S_k^* - \min_{0 \leq k \leq n} S_k^*. \quad (6)$$

Buishand (1982) gives critical values for Q and R for different data set lengths (see Table 1).

2.1.3 Standard normal homogeneity test – likelihood ratio test

SNHT is one of the most popular homogeneity tests in climate studies. The null and alternative hypotheses in this test are the same as in the Buishand test; however, unlike the Buishand test, SNHT is more sensitive to the breaks near the beginning and the end of the series (Costa and Soares, 2009). Alexandersson and Moberg (1997) proposed a statistic $T(k)$ to compare the mean of the first k years of the record with that of the last $(n-k)$ years:

$$T(k) = k\bar{z}_1^2 + (n-k)\bar{z}_2^2, \quad k = 1 \dots n \quad (7)$$

where

$$\bar{z}_1 = \frac{1}{k} \frac{\sum_{i=1}^k (Y_i - \bar{Y})}{s} \quad (8)$$

$$\bar{z}_2 = \frac{1}{n-k} \frac{\sum_{i=k+1}^n (Y_i - \bar{Y})}{s} \quad (9)$$

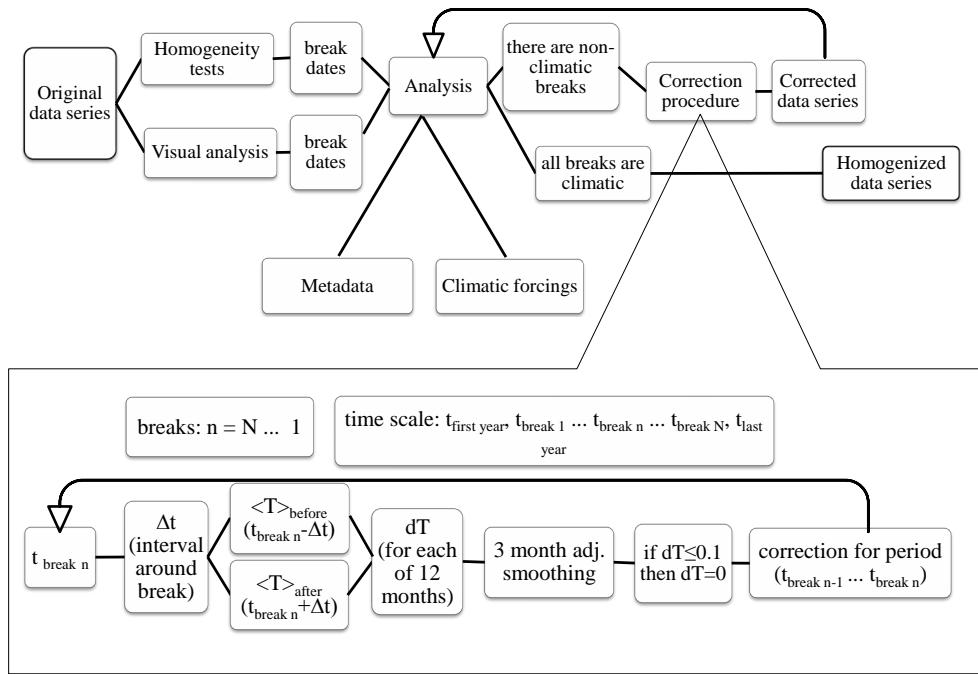


Figure 1. Homogenization procedure. Top – main procedure. Bottom – correction procedure for known non-climatic breaks.

$$s = \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y})^2 \quad (10)$$

If a break is located at the year K , then $T(k)$ reaches a maximum near the year $k = K$. The test statistic T_0 is defined as

$$T_0 = \max_{1 \leq k \leq n} T(k). \quad (11)$$

The null hypothesis is rejected if T_0 is above a certain level, which is dependent on the sample size. Critical values for different data set lengths are given in Khaliq and Ouarda (2007) – see Table 1.

2.1.4 Pettitt test – non-parametric rank test

The null and alternative hypotheses in this test are the same as in the Buishand test, and this test is also more sensitive to the breaks in the middle of the series (Costa and Soares, 2009). The ranks $r_1 \dots r_n$ of the $Y_1 \dots Y_n$ are used to calculate the statistics (Pettitt, 1979):

$$X_k = 2 \sum_{i=1}^k r_i - k(n+1), \quad k = 1 \dots n. \quad (12)$$

If a break occurs in year K , then the statistic is maximal or minimal near the year $k = K$:

$$X_K = \max_{1 \leq k \leq n} |X_k|. \quad (13)$$

The statistical significance (for probability level α) is given as

$$X_{K\alpha} = \left[-\ln \alpha (n^3 + n^2) / 6 \right]^{1/2}. \quad (14)$$

Critical values for X_K for different data set lengths are given in Table 1.

2.2 Homogenization procedure

At first, the series were inspected for outliers that could appear due to typing and/or OCR procedures. This manual and visual inspection was applied to the data both in tabular and in graphical form.

Afterward the following procedure was used for homogenizing the temperature data (see also Fig. 1):

1. Detection of possible breaks in the original data series using visual analysis and aforementioned homogeneity tests (absolute and relative).
2. Comparison of the break dates with available metadata and climatic forcing data (like volcanoes eruptions, anthropogenic landscape changes, etc.). It is possible that metadata do not list all changes in the stations' environments that occurred during the measurements periods; however, in this study we found no significant (as estimated by the statistical tests) breaks that could not be associated to metadata records or other sources.
3. Selection of non-climatic breaks in the data series for correction.

4. Correction of non-climatic breaks:

1. For each break (t_{break}) starting from the latest in time to the first,
 - a. selection of a time interval (Δt) around the current break taking in consideration the length of homogeneous periods before and after the current break;
 - b. calculation of the mean values of the temperature parameters ($\langle T \rangle$ (time period)) for each month separately during two time intervals, before the break (time period = $t_{\text{break}} - \Delta t$) and after the break (time period = $t_{\text{break}} + \Delta t$);
 - c. calculation of the corrections (dT) for each month separately as the difference of these means, $dT = \langle T \rangle(t_{\text{break}} + \Delta t) - \langle T \rangle(t_{\text{break}} - \Delta t)$;
 - d. smoothing of 12 monthly correction values dT by 3-month adjacent averaging to achieve a reasonable variation of dT throughout the year;
 - e. ignoring all dT that are smaller than instrumental errors (0.1°C);
 - f. correction of the data for the periods before current breaks using dT for each month.
2. Proceed to the previous (earlier) break (starting from step 4.1).
3. Visual analysis and homogeneity tests of the corrected data sets (see step 1).
4. In addition, to estimate the “quality” of the correction (Venema et al., 2012), the centered root mean square errors (CRMSE, see e.g. Taylor, 2001 and Gleckler et al., 2008) were calculated as well, using SST (see Sect. 2.4) series or already corrected temperature series for other stations as reference series. The final number of corrected breaks and time intervals for corrections (Δt) were chosen in a way that minimizes not only breaks detected by homogeneity tests statistics but also minimizes the number of months (for each station and each temperature parameter) for which CRMSEs of the corrected data are greater than corresponding CRMSEs for original series.

In case the analysis of corrected series shows the absence of non-climatic breaks (with 95 % significance), the corrected data series are considered to be homogenized for non-climatic breaks with significance of at least 95 %.

2.3 Volcanic eruptions and their effect on temperature variations

Some inhomogeneities detected in the meteorological data do not correspond to known dates of the instrumental or environmental changes. It is possible that these breaks could

Table 2. Major volcanic eruptions from 1850 to 2000. DVI values taken from the NCDC database.

Year	annual DVI	Volcanos	Regions
1855	155	Cotopaxi	Ecuador
1861	164	Makjan/Makian	Indonesia
1875	139	Aksja	Iceland
1883	209	Krakatoa	Indonesia
1888	182	Ritter Island, Bandai-san	Papua New Guinea, Japan
1902	201	St. Maria, Saufriere, Pelee	Guatemala, St. Vincent, Martinique
1963	166.2	Agung	Indonesia
1982	366.1	El Chichon	Mexico
1991	500	Pinatubo	Philippine

be caused by some sudden but natural forcings, e.g. volcanic eruptions (Martínez et al., 2010). The eruptions are accompanied by the injection of SO_2 and dust into the stratosphere. The increase of the dust and aerosol load in the stratosphere causes a reduction of the solar radiation in the lower atmosphere and leads to changes in the lower atmosphere circulation patterns during 2–4 yr after the eruptions (Robock, 2000). Table 2 shows major volcanic eruptions with the dust volcanic index (DVI) reaching more than 100 (from Mann et al., 2000 and NCDC database) from 1850 to 2000. The inhomogeneities that coincide with periods of strong eruptions (1855–1856, 1861–1862, 1875, 1883–1904, 1963–1964, 1982–1984, and 1991–1994) could be of natural (volcanic) origin, provided there were no records of instrumental changes for such epochs. In case some instrumental changes did take place during these periods, it would be difficult to make reasonable corrections only for the non-climatic part of these particular breaks.

2.4 Sea-surface temperature anomalies series

Monthly SST anomalies series (relative to the 1961–1990 period), obtained with HadISST2.0.0.0 (Rayner et al., 2012) in 2 grid points near the Portuguese coast during the period from 1899 to 2010, have been used as reference series to perform relative homogeneity tests and calculate CRMSE values. These series (comprised of a combination of 10 ensemble members) were extracted for the grid cells located between 8–9° W and 41–42° N in the case of the Porto nearest grid point and between 9–10° W and 38–39° N for Lisbon. HadISST2.0.0.0 is based on version 2.5 of the International Comprehensive Ocean Atmosphere Data Set (ICOADS) and includes updated ocean satellite data, among other components. Also, homogeneity adjustments have been applied by Rayner et al. (2012) to the SST data to correct for known bias in the data.

Table 3. Known dates of changes in thermometer heights (h_t) and locations for the three Portuguese stations of Lisbon, Coimbra and Porto.

Station	years	Character of changes	Correction
Lisbon	1864	Moved to a new building (distance is about 1 km)	yes
Lat. 38°43' N	1918	Thermometer height change (+4.1 m)	no
Long. 9°09' W	1920	Thermometer height change (-4.1 m)	no
Alt. 77 m	1929	Thermometer height change (+4.1 m)	no
$h_t = 1.6$ m	1937	Thermometer height change (+0.7 m)	no
	1941	Thermometer height change (-22.2 m)	yes
	1977	Changes in the observation periodicity	no
	1980	Minor changes in the location (within the same garden)	no
Coimbra	1922	Relocation of instrument park; installation of standard shelter	yes
Lat. 40°12' N	1933	Minor relocation	yes
Long. 8°25' W	1950	Thermometer height change (from 1.15 m to 1.45 m)	no
Alt. 141 m			
$h_t = 1.5$ m			
Porto	1916	Moved to new location; thermometer height change (from 10.3 m to 1.3 m above the ground)	yes
Lat. 41°08' N			
Long. 8°36' W	Sep 1920–Feb 1922	No measurements, Probable change in instrument	no
Alt. 93 m			
$h_t = 1.3$ m	1947	Changes in the measurement times	no
	1984	Changes in the measurement times	no

Table 4. Correlation coefficients between the temperature series from Porto and Lisbon (r_L) and Coimbra (r_C) calculated for the period 1910–1932 (± 10 yr around the gap). Significances of the correlation coefficients (p) are smaller than 0.02 with only one exception: $p = 0.37$ for correlation coefficients between T_{\min} of Porto and Lisbon in June (m6). Regression coefficients (A, L, C) for regression models (Porto $T_{\min/\max} = A + L \times T_{\min/\max}(\text{Lisbon}) + C \times T_{\min/\max}(\text{Coimbra})$) are chosen using the best subset procedure with maximization of adj. R^2 parameter and calculated using data for the period 1910–1932 (± 10 yr around the gap). (adj. $R^2 \times 100$) values show the percent of the variability of the dependent variables (Porto T_{\min} and T_{\max} series) that has been accounted for by the model under consideration.

	T_{\min}						T_{\max}					
	r_L	r_C	A	L	C	adj. $R^2 \times 100$	r_L	r_C	A	L	C	adj. $R^2 \times 100$
m1	0.81	0.82	-1.84	0.54	0.53	67.2	0.66	0.86	1.04	0	0.92	70.7
m2	0.90	0.91	-3.89	0.73	0.56	83.7	0.86	0.97	0.36	-0.2	1.09	94.5
m3	0.82	0.89	1.31	0	0.78	79	0.89	0.96	-2.72	0.15	0.88	92
m4	0.56	0.72	0.9	0	0.87	52.9	0.88	0.92	0.07	0.1	0.81	84.3
m5	0.75	0.72	-2.55	1	0	59.6	0.92	0.85	-3.85	0.67	0.34	91
m6	0.21	0.58	9.14	-0.22	0.61	31.6	0.87	0.96	1.67	-0.17	0.97	90.8
m7	0.57	0.75	1.4	0	0.91	58.9	0.92	0.95	-0.67	0.33	0.54	90.2
m8	0.51	0.66	4.79	0	0.66	41.9	0.82	0.93	-1.28	-0.1	0.95	80.8
m9	0.75	0.83	0.55	0	0.94	68.6	0.84	0.92	0.15	0.19	0.67	80.7
m10	0.52	0.58	3.65	0	0.66	30	0.90	0.91	-2	0.63	0.33	83.3
m11	0.74	0.83	-0.13	0	1	67.2	0.76	0.56	-1.85	0.76	0.2	57.1
m12	0.76	0.77	1.14	0	0.85	58.2	0.51	0.54	8.17	0.12	0.26	21.8

In the current analysis the SST series measured near Porto (mean of 10 ensemble members, later on “SST Porto”) have been used as reference series for Porto temperature series homogenization, and SST series measured near Lisbon (mean of 10 ensemble members, later on “SST Lisbon”) have been used in the homogenization of Lisbon temperature series.

3 Porto (Serra do Pilar) temperature series

3.1 Data description and metadata

The original data set contains monthly averages of daily minimum (T_{\min}) and maximum (T_{\max}) temperature and their annual means measured by Instituto Geofísico (Observatório Meteorológico da Serra do Pilar) da Universidade do Porto (IGUP), Porto, from 1888 to 2001. The data set length is

114 yr. Measurement errors are $\pm 0.2^\circ\text{C}$ (valid for all observed temperature series presented here).

The meteorological station of Porto has been in regular operation since 1888 when it was put under the jurisdiction of the Observatório Meteorológico da Princesa D. Amélia, now IGUP, on the south part of the river Douro. In 1916 the station location was changed slightly and the thermometer was moved from the tower (10.3 m above the ground) to the ground level (1.3 m above the ground). The data set has a gap from September 1920 to February 1922; there is also a possibility that the thermometer was replaced in March 1922. In 1947 and again in 1984/1985, changes in the observation times were made. Table 3 shows known dates of possible non-climatic breaks due to instrument changes (Pinhal, 2008).

Changes in the location of the instruments could result in sudden jumps of the measured parameter values. T_{\min} and T_{\max} could respond differently to changes in position of the respective thermometers, depending on the character of the changes in the instrument's environment (Aguilar et al., 2003). Therefore, the variations of DTR could be more important for the detection of the breaks; breaks could be weak in the T_{\min} or T_{\max} series, but clearly seen in DTR series (Wijngaard et al., 2003). The following parameters have been analysed (valid for all temperature series presented here):

1. minimum temperature (T_{\min});
2. maximum temperature (T_{\max});
3. temperature range ($\text{DTR} = T_{\max} - T_{\min}$);
4. monthly average temperature ($T_{\text{aver}} = (T_{\max} + T_{\min})/2$).

All series contain monthly and annual means; T_{\min} and T_{\max} are measured values, DTR and T_{aver} are calculated values. To perform relative homogeneity tests, the differences between standardized T_{\min} , T_{\max} and T_{aver} series and standardized SST Porto series were calculated.

3.1.1 Interpolation of the gap from September 1920 to February 1922

The gap in the data from September 1920 to February 1922 should be filled before the data are subjected to the homogeneity analysis. It is possible to fill the gap using the simple linear interpolation for the absent one or two values for each of the monthly data series. On the other hand, it is possible to build a mathematical regression model for a more realistic interpolation, using data from nearby meteorological stations, namely, Coimbra and Lisbon data series. All 12 monthly series of T_{\min} and T_{\max} were interpolated separately. The time period used for the regression models is 10 yr before the gap (1910–1919) plus 10 yr after the gap (1923–1932).

First, the correlation coefficients (r) between temperature parameters measured in Porto and Coimbra and Lisbon were

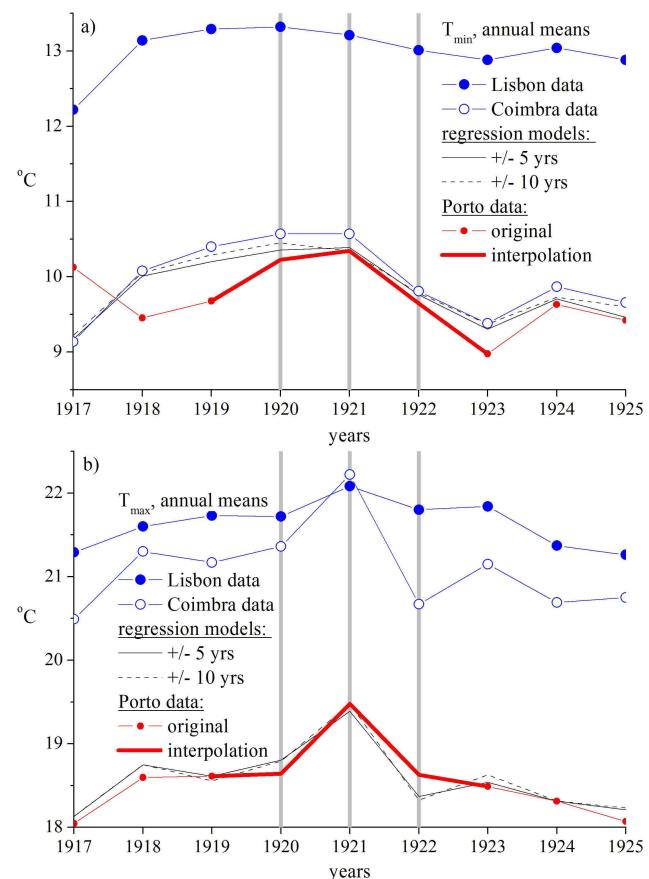


Figure 2. Variations of T_{\min} (a) and T_{\max} (b) measured in Lisbon, Coimbra and Porto-Serra do Pilar from 1917 to 1925 (annual means) and approximations by multiple regression models for time periods of ± 5 and ± 10 yr around the gap – annual sums. Grey vertical lines mark the period of absent data. The bold red lines show the accepted interpolation.

calculated (see Table 4). The significances (p) of the correlation coefficients are smaller than 0.02 with a single exception. There are strong correlations ($r = 0.51\ldots 0.92$) between the temperature variations in Porto and Lisbon and Coimbra for almost all months with only one exception – the correlation between Porto and Lisbon series of T_{\min} (June, m6): $r_L = 0.21$, $p = 0.37$. Nevertheless, it is still possible to use the data from Lisbon and Coimbra as regressors for Porto data in multiple regression models.

Multiple regression models for Porto T_{\min} and T_{\max} series were built using the Coimbra and Lisbon data as regressors. The models have been built using the “best subset” method, maximizing the adj. R^2 parameter. The regression coefficients are shown in Table 4 alongside with (adj. $R^2 \times 100$) values that show the percentage of the variability of the dependent variable (Porto series) that has been accounted for by the model under consideration. As one can see, the multiple regression models are good approximations of the real data and can be used for the gap interpolation in the Porto

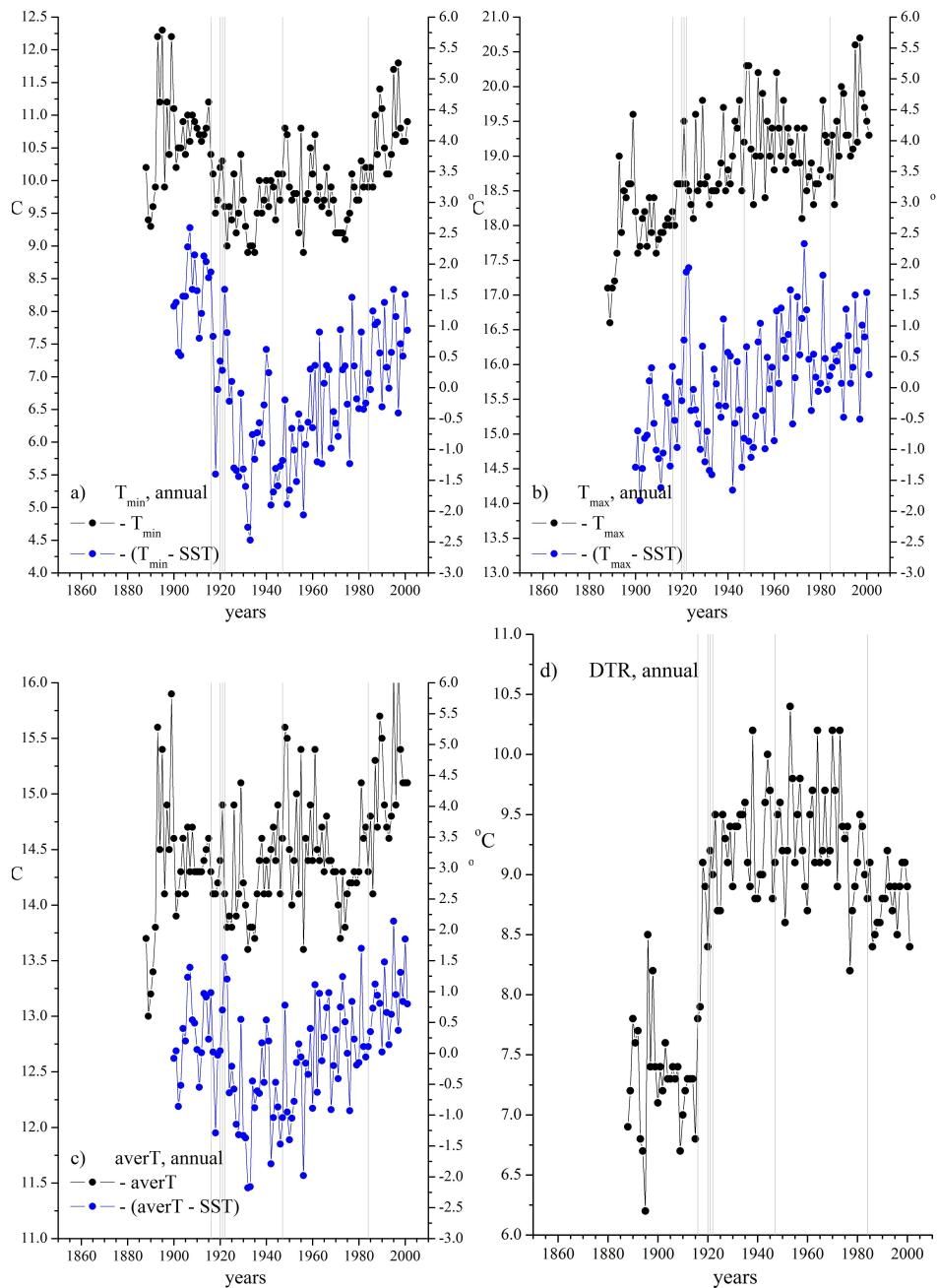


Figure 3. Porto: annual variations of T_{\min} (a), T_{\max} (b), averT (c) and DTR (d); temperature series are shown in black, differences between temperature and SST Porto series are shown in blue. Grey vertical bands show dates of known instruments relocation (see Table 3).

data. Similar regression models have been calculated using a smaller time period: ± 5 yr around the gap (1915–1919 plus 1923–1927). However, the 5-yr-around-gap models give, in general, worse approximations for the real data than the 10-yr-around-gap models. Finally, the gap from September 1920 to February 1922 was interpolated using the 10-yr-around-gap multiple regression models for each parameter and for each month separately. Annual values of T_{\min} and T_{\max} for 1920–1922 have been calculated using both measured and

interpolated monthly data. Both regression models (only for annual means) alongside with original annual means for all three meteorological stations are shown in Fig. 2a–b. Please note that the annual interpolation values for 5- and 10-yr-around-gap models shown in Fig. 2 are just average values calculated on the basis of monthly interpolation values of each model for presentation purposes only and were not used for interpolation. As can be seen, the interpolation using the multiple regression models instead of simple linear

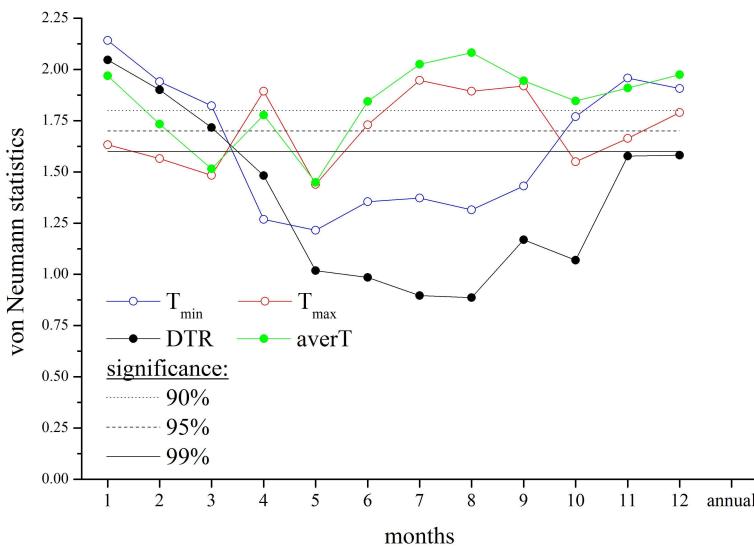


Figure 4. Porto: von Neumann ratio statistics for monthly series of T_{\min} , T_{\max} , DTR and T_{aver} . Black straight lines show probability levels.

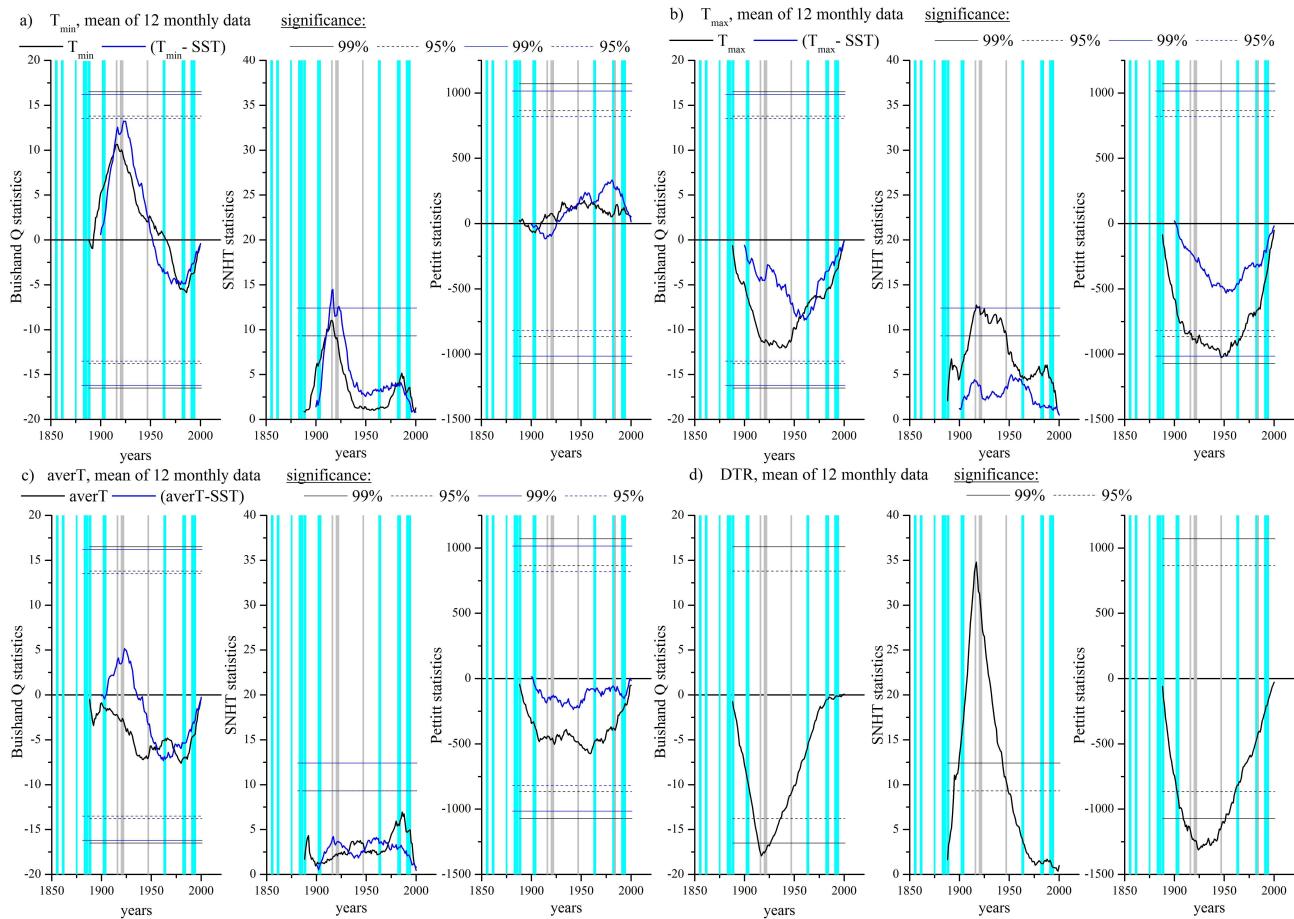


Figure 5. Porto: average of 12 monthly series of Buishand Q test (left panels), SNHT (middle panels) and Pettitt tests (right panels) statistics of T_{\min} (a), T_{\max} (b), T_{aver} (c) and DTR (d). Statistics of temperature series are shown in black, statistics of differences between temperature and SST series are shown in blue. Solid and dashed horizontal lines show probability levels of 99 % and 95 %, respectively. Grey vertical lines show known dates of instrumental changes, cyan broad vertical lines show periods of strong volcanic influence.

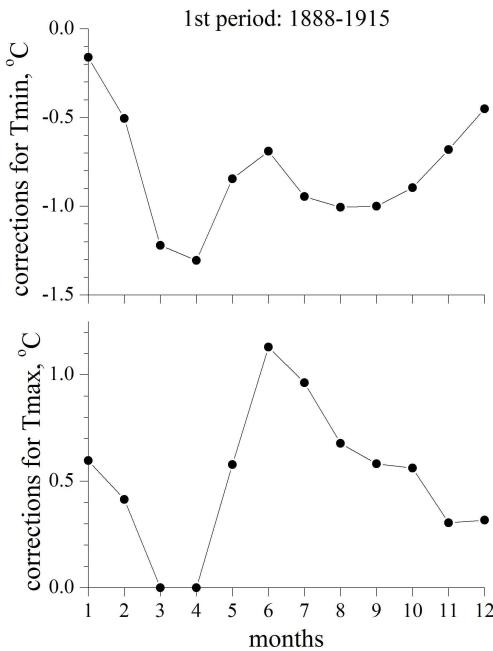


Figure 6. Porto: corrections (in °C) for T_{\min} (top panel) and T_{\max} (bottom panel) series for the period before the break, 1888–1915.

interpolation does not ignore the real variations of the temperature parameters that took place (according to the records from Lisbon and Coimbra) from September 1920 to February 1922. In particular, the interpolation using multiple regression models allowed us to preserve the following features that were observed in monthly series variations (not shown):

- higher than estimated by the linear interpolations values of T_{\min} during the periods of November to December 1920 and July to October 1921;
- lower than estimated by the linear interpolations values of T_{\min} in May 1921;
- higher than estimated by the linear interpolations values of T_{\max} in December 1920, January 1921, and during the period from March to December 1921;
- lower than estimated by the linear interpolations values of T_{\max} in October 1920 and in February and May 1921.

3.2 Homogenization

3.2.1 Visual analysis

Figure 3a–d show time variations of the annual series of T_{\min} , T_{\max} , T_{aver} and DTR, respectively. The grey vertical lines mark the dates of known changes in thermometer position. DTR variations (Fig. 3d) show at least one easily detectable break in 1916. This break corresponds to the most significant change in the instruments location: movement to a new

place and relocation of the instruments from the top of the tower to the ground level. The changes in the measurements times (in 1947 and 1984) and possible changes in the instrument park after the gap in 1922 could not be easily detected by the visual analysis. The break in 1916 has different effects on the monthly T_{\min} and T_{\max} variations (not shown). There are significant jumps in the monthly T_{\min} variations clearly seen during warm months (from April to September). However, there are no jumps in the monthly T_{\max} variations that could be easily detected by the visual analysis. This difference could be explained by the different sensitivity of the T_{\min} and T_{\max} to the change in location and in the instrument height.

3.2.2 Homogeneity tests results

Figure 4 shows the von Neumann ratio for 12 monthly series of T_{\min} , T_{\max} , DTR and T_{aver} . This test shows strong inhomogeneities in all four series and DTR in particular. As one can see, variations of the homogeneity of the data series strongly depend on the temperature parameter:

- T_{\min} – data series of warm months (from April to September) are more inhomogeneous than of cold ones;
- T_{\max} – data series of warm months (from April to September) are less inhomogeneous than of cold ones with one exception – May;
- DTR – data series of two months only (January and February) are apparently homogeneous;
- T_{aver} – these data are more homogeneous than T_{\min} and T_{\max} . They could be labelled as inhomogeneous with a probability of 95 % only in March and May.

Figure 5a–d show test statistics (absolute and relative) for Buishand, SNHT and Pettitt test for T_{\min} , T_{\max} , T_{aver} and DTR, respectively. The average of 12 monthly statistics series is plotted in these figures to emphasize the main features of each homogeneity test statistics and for better visualisation. From Fig. 5a–d it is possible to detect the strongest break in data homogeneity around 1916 – date of movement to a new location and change in the thermometers height. Also, for some months (not shown) there are breaks in the homogeneity around 1920s (gap and probable change of the thermometer), 1930s (unknown origin), 1947 (changes in the measurements time), 1963 (volcanic eruption), and 1984 (changes in the measurements time, which coincide with volcanic eruptions).

However, these breaks are seen only for some months and not by all three homogeneity tests at the same time. Therefore, the only break that has to be corrected is the 1916 break. Other non-climatic breaks have no statistically significant effect or could not be corrected due to the coincident influence of other (climatic) forces like the volcanic eruption around

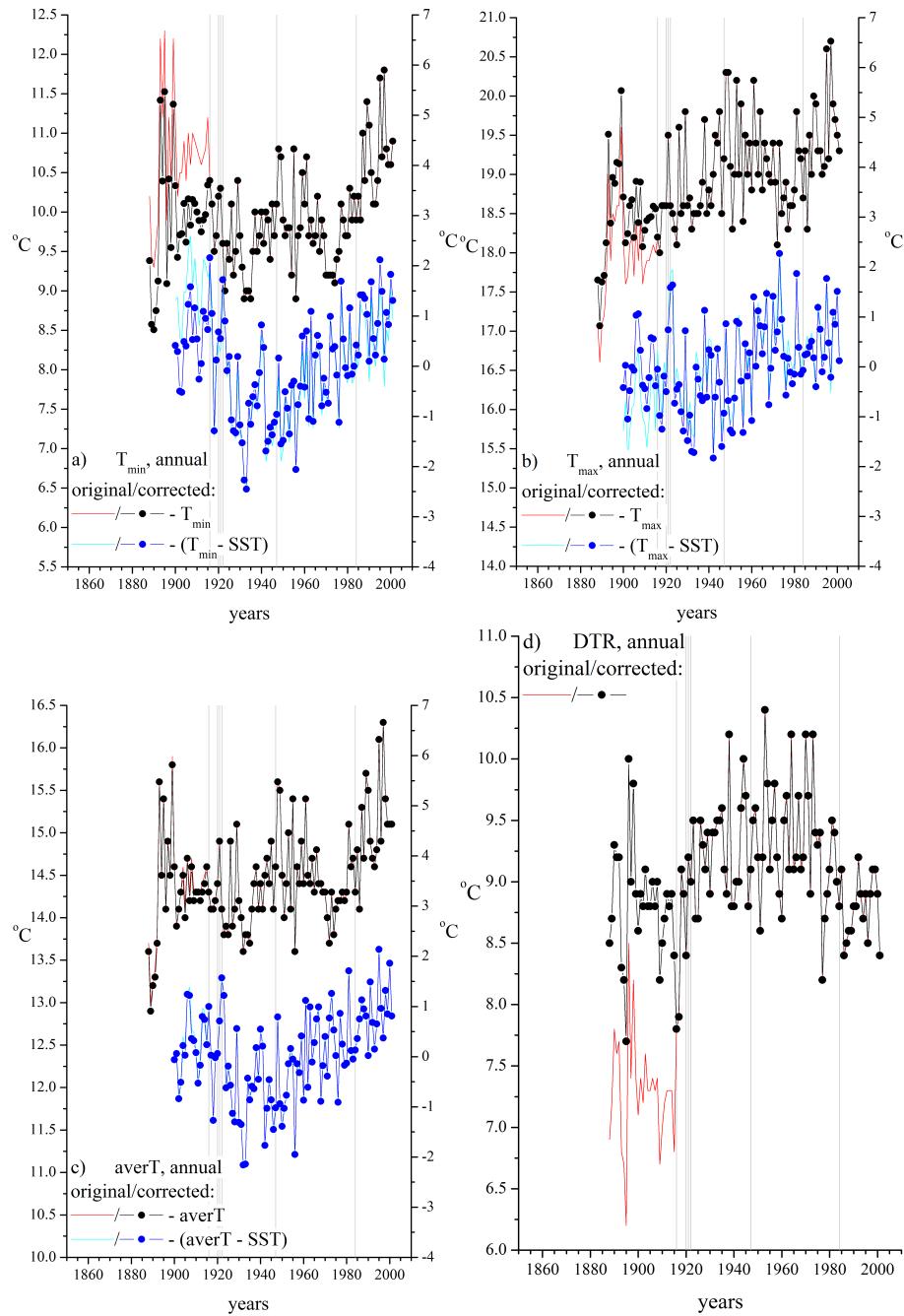


Figure 7. Porto: original and corrected annual series of T_{\min} (a), T_{\max} (b), T_{aver} (c) and DTR (d); temperature series are shown in black and red, differences between temperature and SST Porto series are shown in blue and cyan. Grey vertical bands show dates of known instruments relocation (see Table 3).

1984. This conclusion was confirmed later during the correction procedure (Sect. 3.3.1) – correction of possible breaks in 1922 and 1947 makes the corrected series even more inhomogeneous.

3.2.3 Preliminary conclusions

T_{\min} data set shows inhomogeneities in 1916, near the 1920s, 1947, 1963 and 1984; T_{\max} data set showed inhomogeneities in 1916, near the 1920s, 1930s, 1947, 1963 and, 1984; DTR data set showed strong inhomogeneity during the period 1916–1922 and, probably, a weak break in the 1940s; T_{aver} showed weak breaks around the 1920s, 1930s, 1947 and

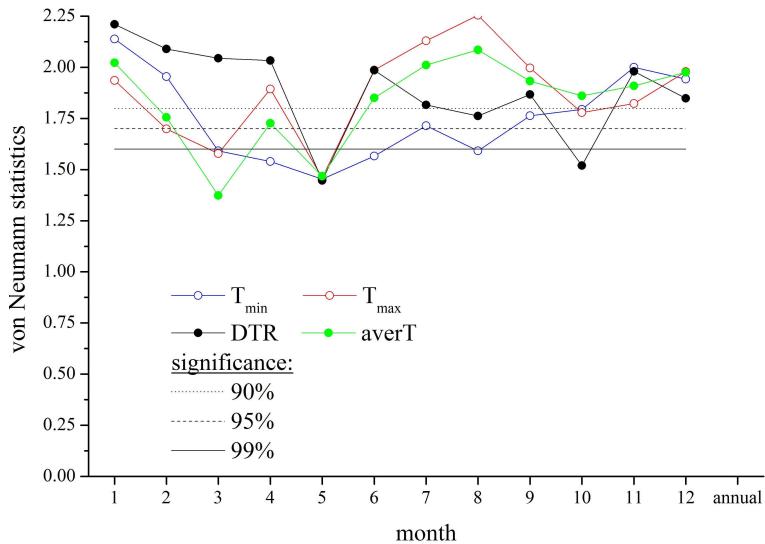


Figure 8. Porto: same as Fig. 4 but for the corrected series.

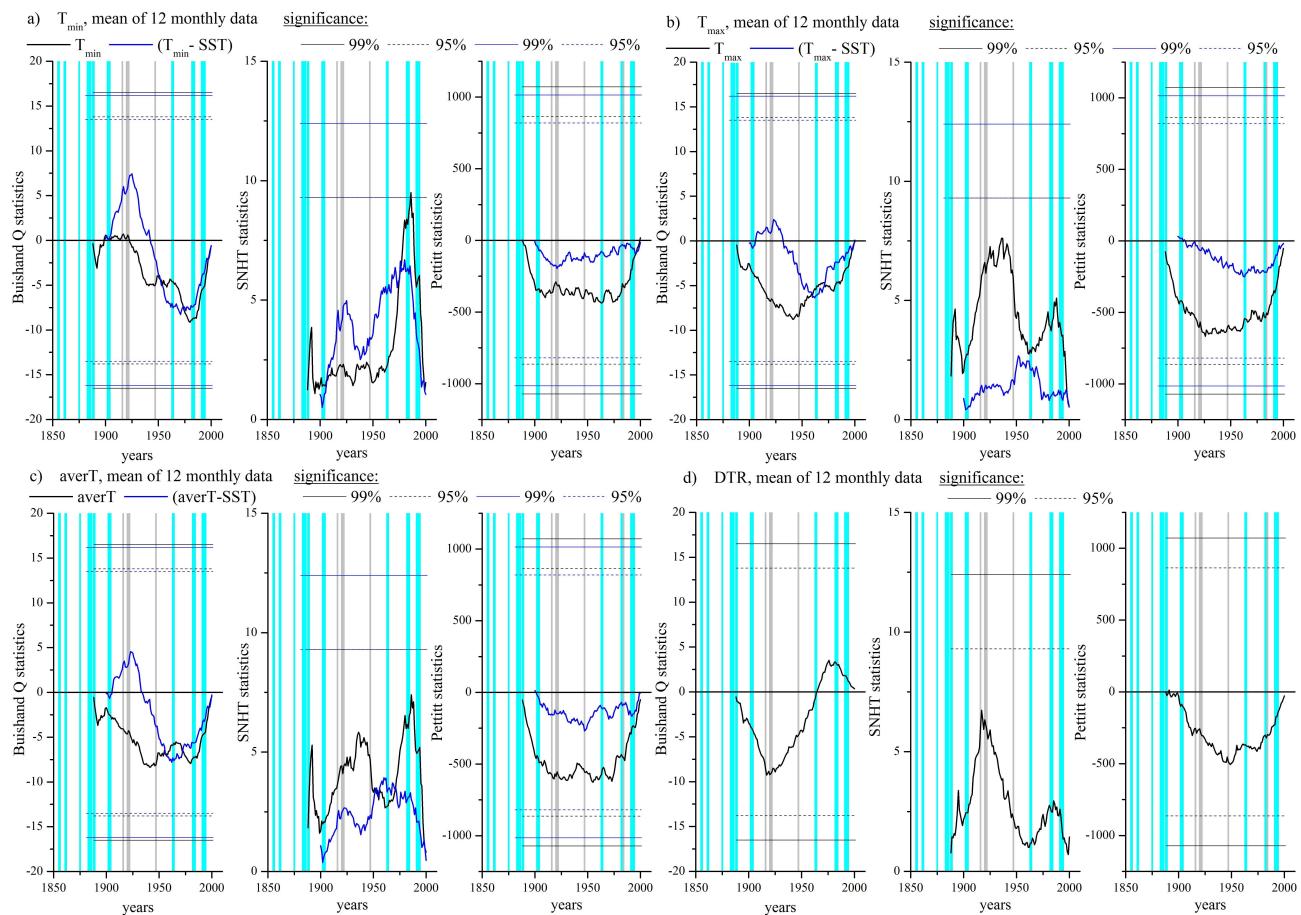


Figure 9. Porto: same as Fig. 5a-d but for the corrected series.

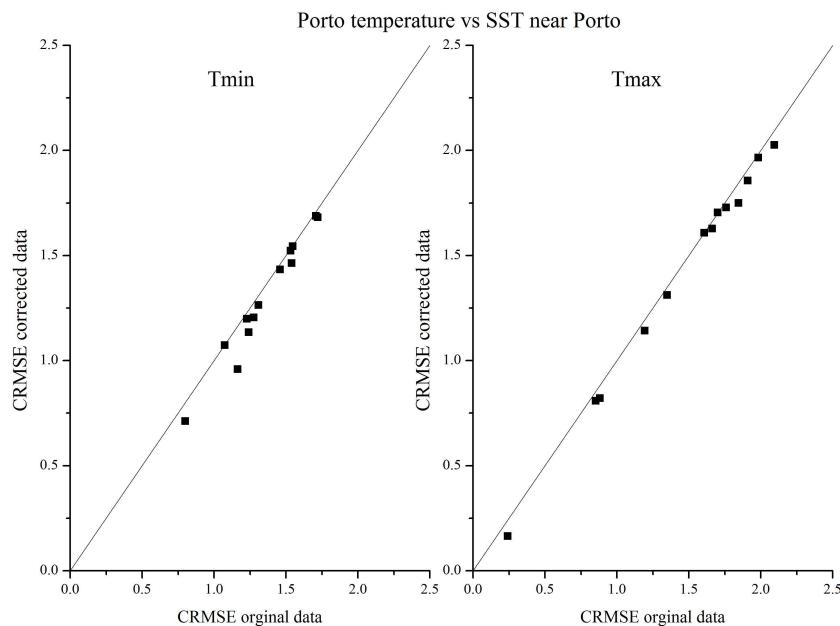


Figure 10. Porto: scatter plots of CRMSE before and after homogenization of 12 monthly and the annual series for T_{\min} (left) and T_{\max} (right). SST Porto anomalies data are used as reference series. Dots on or below the bisect indicate data sets with unchanged or improved (increased) homogeneity, while dots above the bisect indicate data sets with increased inhomogeneity.

1984; T_{\min} during warm months and T_{\max} during cold months are more inhomogeneous than in other months. The inhomogeneities that are not associated to known dates of instrumental changes may be due to the internal climatic variations caused, for example, by major volcanic eruptions. The most significant non-climatic break occurred in 1916 due to changes in the instruments location and height. This break requires correction. Other breaks detected by homogeneity tests have no statistically significant effect or could not be corrected due to the coincident influence of other (climatic) forces.

3.3 Correction for non-climatic breaks

3.3.1 Correction procedure

To correct the non-climatic breaks we used the procedure described in Sect. 2.2. The best corrections were obtained when T_{\min} and T_{\max} data sets were divided into two periods: 1888–1915 and 1916–2001. For each month the means of temperature parameters for certain time intervals (± 5 yr for T_{\min} and ± 15 yr for T_{\max}) around 1916 were calculated. The data for the 1888–1915 period were corrected using the correction values calculated as described above. All correction values are shown in Fig. 6. The corrections were applied to T_{\min} and T_{\max} data sets. Afterwards, corrected values of DTR and T_{aver} were calculated. Results of the correction as well as original data are shown in Fig. 7a–d. Please note that due to the use of standardized values, the difference between temperature series and SST anomalies presented in Fig. 7 (and similar for

other stations) show differences between corrected and original series even for non-corrected periods – the series means and standard deviations that are used in the standardizing procedure change after correction.

3.3.2 Homogeneity of the corrected series

All four corrected data sets were subjected to the same homogeneity tests as the original data. The results of these tests for T_{\min} , T_{\max} , T_{aver} and DTR are shown in Figs. 8 and 9a–d (similarly to Figs. 4–5). As one can see from the comparison of similar statistics for the original (Figs. 4–5) and corrected (Figs. 8–9) data, the latter data sets are less inhomogeneous but still contain inhomogeneities coinciding with the volcanic eruptions that occurred in the end of the 19th and 20th centuries. Some absolute tests for some months (not shown) still show breaks of homogeneity in a period lasting from 1922 to 1947, although there is no consistency between the three homogeneity tests (Buishand and Pettitt tests and SNHT) in relation to the dates of the breaks. The relative homogeneity tests showed an almost total absence of the breaks around dates of known instrument changes. Therefore, the corrections for these breaks were not necessary. The homogeneity level given by the von Neumann ratio of the corrected data series (Fig. 8) still depends on the temperature parameter; among all parameters T_{\min} is the least homogeneous. One of the possible reasons for the remaining inhomogeneities in the T_{\min} data series is the volcanic effect. Figure 10 shows CRMSEs of corrected series (SST Porto are reference series) plotted versus corresponding CRMSEs of the original series.

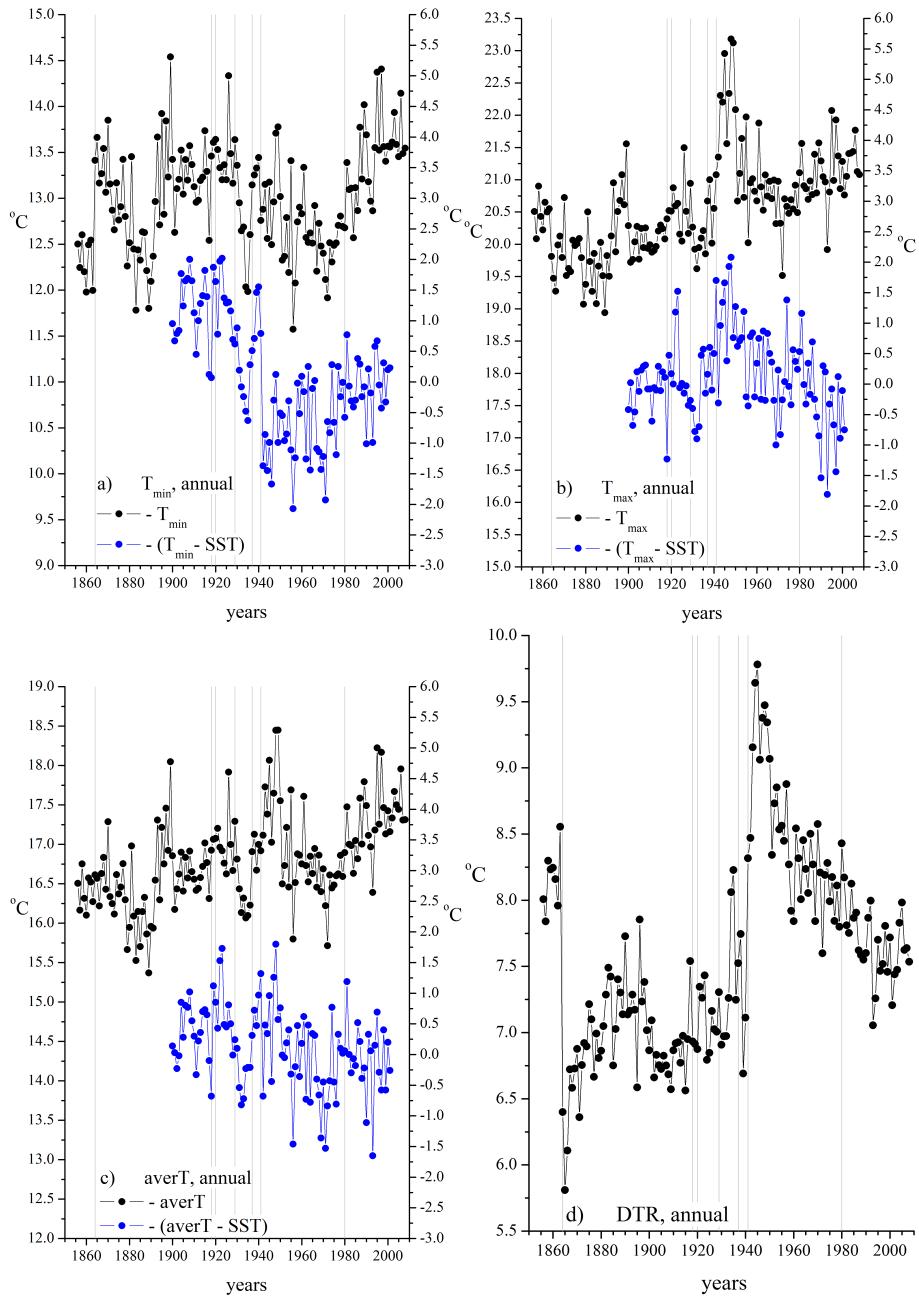


Figure 11. Lisbon: annual variations of T_{\min} (a), T_{\max} (b), averT (c) and DTR (d); temperature series are shown in black, differences between temperature and SST Lisbon series are shown in blue. Grey vertical bands show dates of known instruments relocation (see Table 3).

As one can see, the inhomogeneity level of both T_{\min} (left panel) and T_{\max} (right panel) decreases or stays unchanged for all monthly series.

Breaks detected in the corrected data sets by the different homogeneity tests are rarely coincident, except for the end of the 20th century (an epoch of El Chichon and Pinatubo eruptions – see Table 2). Sometimes the tests still show breaks of homogeneity during different periods but there is no consistency between the three homogeneity tests (Buishand and

Pettitt tests and SNHT) in the dates of the breaks. In our opinion, these inhomogeneities are caused by the application of the correction values which are already smoothed by a 3-months adjacent average to maintain the annual cycle (see Sect. 2.2) and we believe that in these cases additional corrections are not necessary. Thus, we consider the data sets of T_{\min} and T_{\max} corrected by the procedure described in the paper as free of non-climatic changes with a significance of at least 95 %.

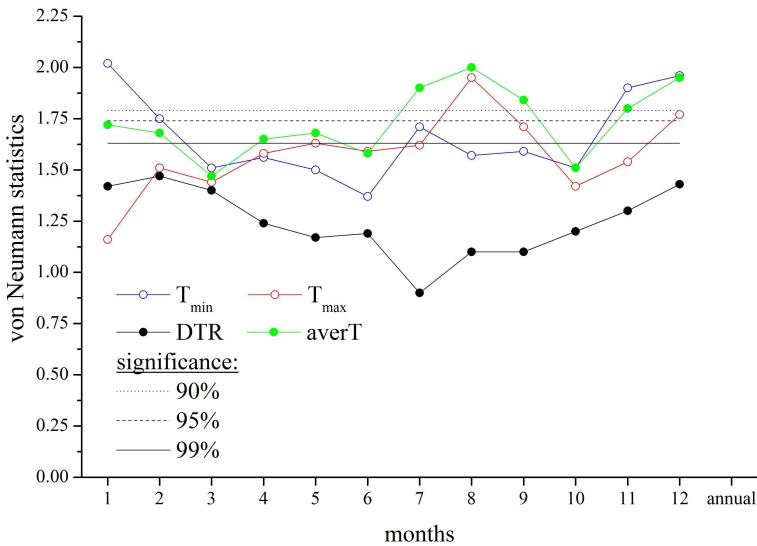


Figure 12. Lisbon: von Neumann ratio statistics for monthly series of T_{\min} , T_{\max} , DTR and T_{aver} . Black straight lines show probability levels.

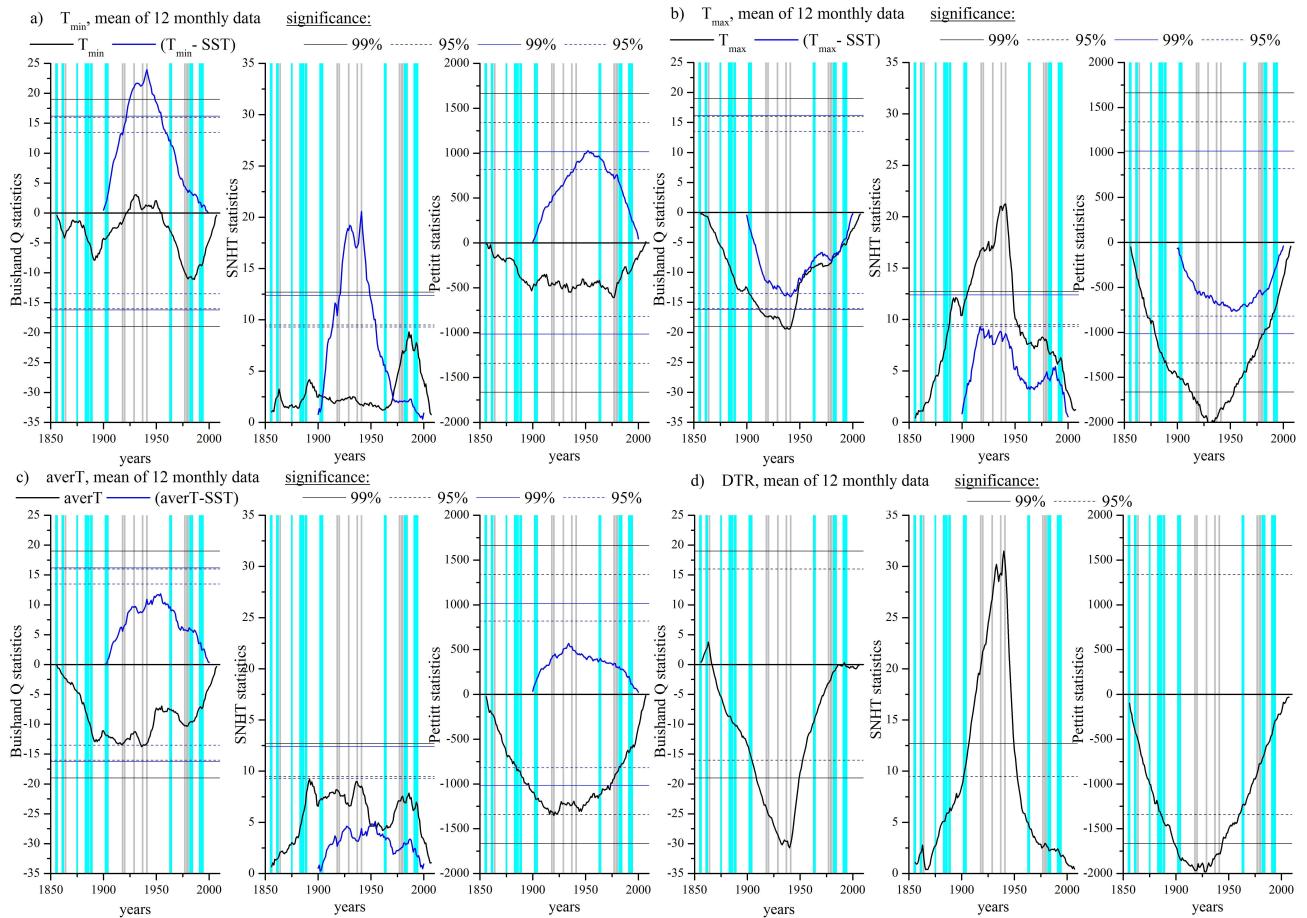


Figure 13. Lisbon: average of 12 monthly series of Buishand Q test (left panels), SNHT (middle panels) and Pettitt tests (right panels) statistics of T_{\min} (a), T_{\max} (b), T_{aver} (c) and DTR (d). Statistics of temperature series are shown in black, statistics of differences between temperature and SST series are shown in blue. Solid and dashed horizontal lines show probability levels of 99 % and 95 %, respectively. Grey vertical lines show known dates of instrumental changes, cyan broad vertical lines show periods of strong volcanic influence.

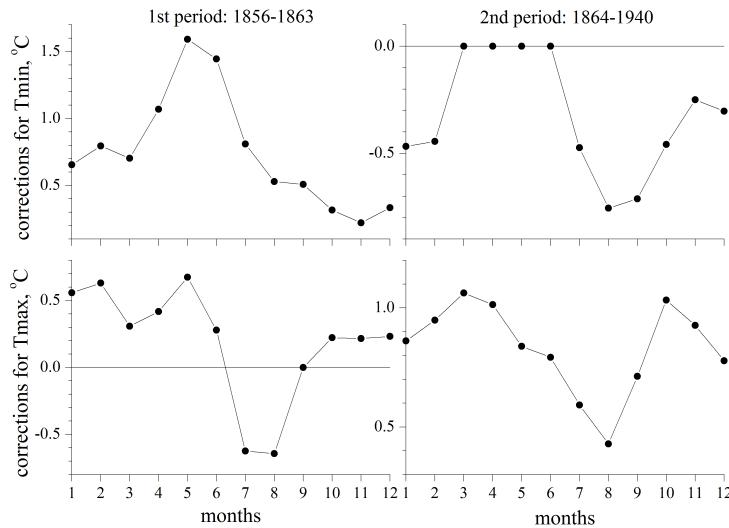


Figure 14. Lisbon: corrections (in °C) for T_{\min} (top panels) and T_{\max} (bottom panels) series for two periods between the breaks 1856–1863 (left panels) and 1864–1940 (right panels).

4 Lisbon IGIDL temperature series

4.1 Data description and metadata

The original data sets contain monthly averages of daily minimum (T_{\min}) and maximum (T_{\max}) temperature and their annual means measured at Instituto Geofísico do Infante D. Luís (IGIDL), Lisbon from 1856 to 2008. The data sets length is 153 yr.

The meteorological station of Lisbon/Geofísico has been in regular operation since October 1854. During the first ten years the thermometers were positioned in the terrace of the Observatory Tower of the old Escola Politécnica, located in the Jardim Botânico. This three-storied tower was built in 1854, thus leading to the foundation of the Infante D. Luiz Observatory (now IGIDL). This building proved inadequate for systematic observations and a new 4-storied tower was inaugurated in October 1863 in the main central edifice of Escola Politécnica, with the thermometers being reinstalled in the new terrace. This building still houses the IGIDL today and some of its meteorological instruments, but the park of instruments containing the thermometers (the Stephenson shelter), initially installed on the platform of the new observatory tower, was transferred to the grounds in Jardim Botânico in 1941 (the distance between the two locations is about 120 m). In 1979 the Jardim Botânico's instrumental park location was slightly changed. Additionally, in January 1977 changes in the times of observation have been made (Carvalho, 2001). Table 3 summarizes the information about possible non-climatic breaks that could appear in the Lisbon/Geofísico temperature series.

To perform relative homogeneity tests, the differences between standardized T_{\min} , T_{\max} and T_{aver} series and standardized SST Lisbon series were calculated.

4.2 Homogenization

4.2.1 Visual analysis

Figure 11a–d show the time variations of the annual series of T_{\min} , T_{\max} , T_{aver} and DTR, respectively. As one can see, DTR variations (Fig. 11d) show two easily visible breaks in 1863/1864 and 1940/1941. These breaks correspond to the two most significant changes in the instruments location: movement to a new place in 1864 and relocation of the instruments from the top of the tower to the ground level in 1941. At first sight, it seems that the minor changes in the thermometers height that took place from 1917 to 1937 and minor changes in the instruments location in 1979 were too small to have a significant influence on the data homogeneity.

These two breaks have different influences on the T_{\min} and T_{\max} variations (see Fig. 11a and b). As one can see, during the first break (1864) there are significant jumps both in T_{\min} and T_{\max} . However, during the second break in 1941 there is a significant jump in T_{\max} but a very small one (if any) in T_{\min} . On the contrary, the difference between T_{\min} and SST (Fig. 11a, blue line) has a significant jump in 1941, whereas the difference between T_{\max} and SST (Fig. 11a, blue line) shows no visible jumps. This dissimilarity could be explained by the different character of the changes in the instruments locations. In 1864 the instruments were moved to a new place with a new microclimate; in 1941 the change was mainly in the instruments height, not so much in location, causing a significant jump only in one of the extremes (see Aguilar et al., 2003). These conclusions were also derived from the visual analysis of the 12 monthly data series for each of the four temperature parameters (not shown) – jumps around 1864 and 1941 are seen for almost all months.

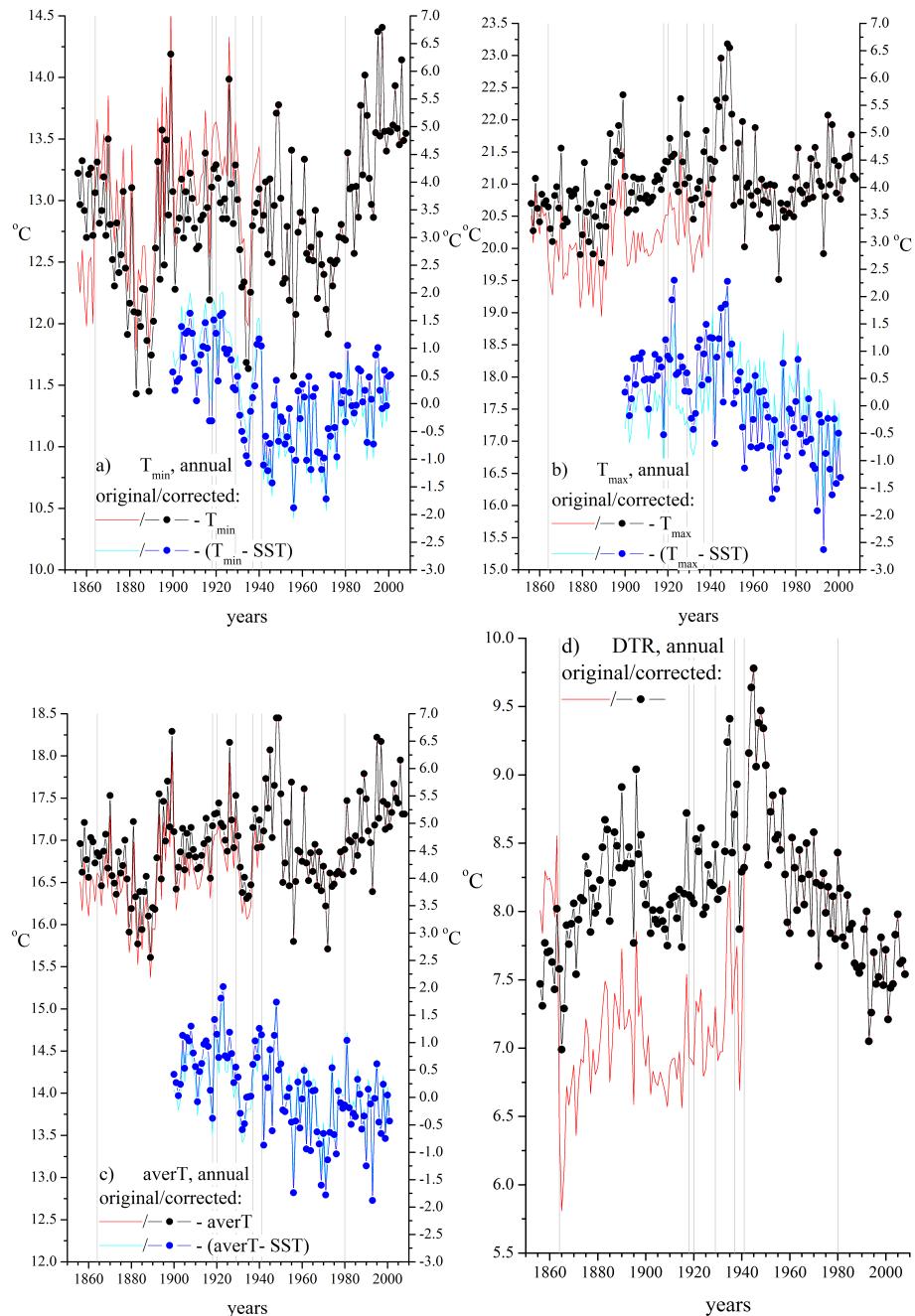


Figure 15. Lisbon: original and corrected annual series of T_{\min} (a), T_{\max} (b), averT (c) and DTR (d); temperature series are shown in black and red, differences between temperature and SST Lisbon series are shown in blue and cyan. Grey vertical bands show dates of known instruments relocation (see Table 3).

4.2.2 Homogeneity tests results

The von Neumann test statistics for the 12 monthly series of T_{\min} , T_{\max} , averT and DTR are shown in Fig. 12. The variations of the homogeneity of monthly data series (given by the von Neumann ratio) strongly depend on the temperature parameter:

- T_{\min} – data series of warm months are more inhomogeneous than of cold ones;
- T_{\max} – all months show strong inhomogeneity except August (m8);
- averT – data series of months from January to June and October (from m1 to m6 and m10) are inhomogeneous;

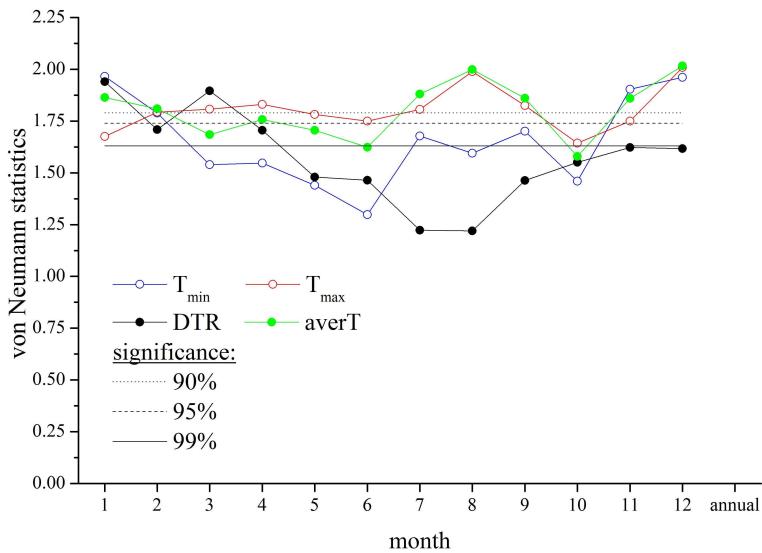


Figure 16. Lisbon: same as Fig. 12 but for the corrected series.

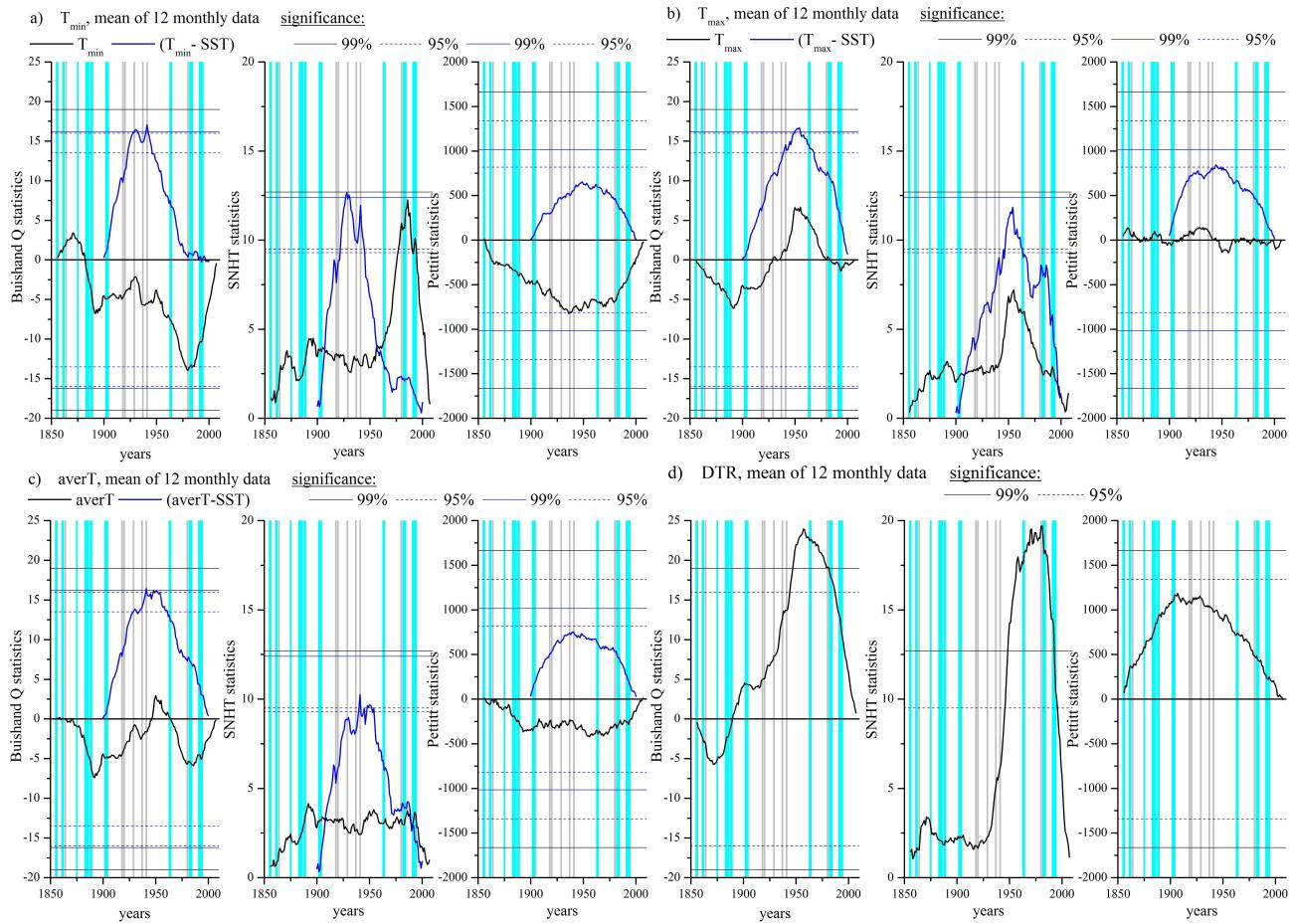


Figure 17. Lisbon: same as Fig. 13a–d but for the corrected series.

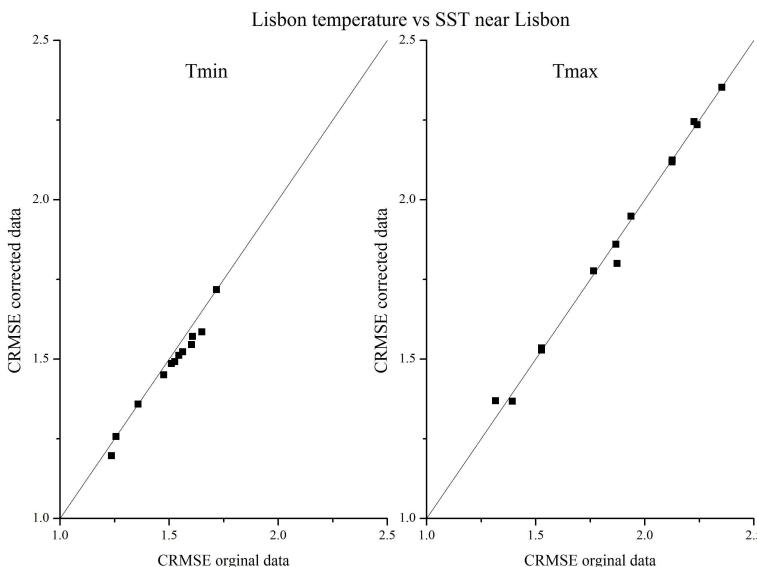


Figure 18. Lisbon: scatter plots of CRMSE before and after homogenization of 12 monthly and the annual series for T_{\min} (left) and T_{\max} (right). SST Lisbon anomalies data are used as reference series. Dots on or below the bisect indicate data sets with unchanged or improved (increased) homogeneity, while dots above the bisect indicate data sets with increased inhomogeneity.

- DTR – data series of warm months are more inhomogeneous than of cold ones.

The average of 12 monthly test statistics series (absolute and relative) for Buishand, SNHT and Pettitt test for T_{\min} , T_{\max} , T_{aver} and DTR for annual series are shown in Fig. 13a–d, respectively. The grey vertical lines mark the dates of known changes in thermometer position. As one can see, some of these dates (namely, 1864 and the period from 1916 to 1941) coincide with significant breaks depicted by the maxima (or minima) of the curves. It should be mentioned that for T_{\min} the coincidences between the known instrumental changes dates and break years detected by the absolute tests are rare, whereas for T_{\max} , DTR and T_{aver} these coincidences are very frequent. Also, there are two periods of possible break years detected by the homogeneity tests that do not coincide with known dates of instrument changes: one is at the end of the 19th century/beginning of 20th century (approx. from 1880–1890 to 1900) and the second is at the end of the 20th century (approx. from 1970 to 1990). Relative homogeneity tests (blue lines) of T_{\min} show significant breaks around 1937 (small changes in the thermometer height) and homogeneity tests of T_{\max} show significant breaks around 1941.

4.2.3 Preliminary conclusions

T_{\min} data sets show inhomogeneities in the 1860s, near 1970s–1980s and, possibly, near 1880s–1890s; T_{\max} is more sensitive than T_{\min} to the changes of the thermometer height that took place in 1864 and from 1916 to 1941. T_{\max} data sets show strong inhomogeneity during this period. There are also some inhomogeneities near 1880s–1890s and 1970s–

1980s; DTR and T_{aver} show the same three periods of inhomogeneity: near 1880s–1890s, 1910s–1940s and 1970s–1980s. Temperature series of warm months contain more inhomogeneities than those of cold months. The inhomogeneities that are not associated to known dates of instrumental changes could appear due to the internal climatic variations caused by, e.g. major volcanic eruptions. The most significant non-climatic breaks have occurred in 1864 and 1941 due to changes in the instruments location (1864) and height (1941). These breaks have to be corrected.

Small changes in the thermometer height took place from 1917 to 1936 and the short periods between the changes do not allow us to estimate statistically significant corrections. The dislocation of the station in 1979 does not significantly (with significance 95 % or more) affect the homogeneity of the data – the means of the temperature parameters for 1941–1978 and 1979–2008 are the same within the instrumental and statistical errors.

4.3 Correction for non-climatic breaks

4.3.1 Correction procedure

The T_{\min} and T_{\max} data sets were divided into three periods: 1856–1863, 1864–1940, and 1941–2008. We started from the most recent break – 1940/1941. For each month the means of temperature parameters for certain time intervals (± 20 yr for T_{\min} and ± 45 yr for T_{\max}) around 1941 were calculated. The second break (1863/1864) was corrected using means calculated for time intervals 1864 ± 8 yr both for T_{\min} and T_{\max} . All correction values are shown in Fig. 14. As one can see, the corrections for the second period (1864–1940) for

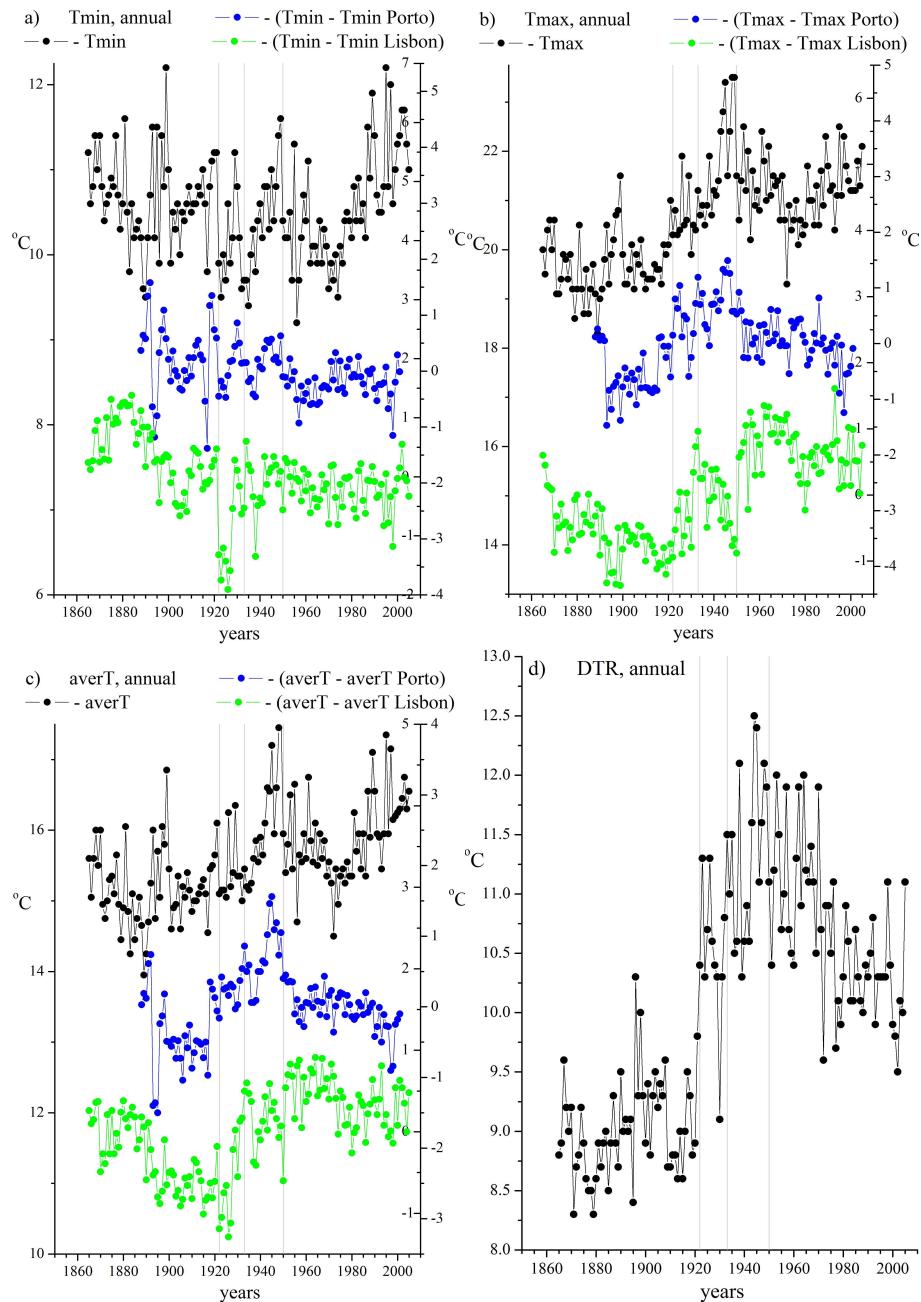


Figure 19. Coimbra: annual variations of T_{\min} (a), T_{\max} (b), T_{aver} (c) and DTR (d); Coimbra temperature series (black) and differences between Coimbra and Porto (blue) and Lisbon (green) temperature series. Grey vertical bands show dates of known instruments relocation (see Table 3).

T_{\max} are non-zero for all months whereas the corrections for T_{\min} are equal to zero for months from March to June (m3–m6). Results of the correction as well as of the original data are shown in Fig. 15a–d for annual series.

4.3.2 Homogeneity of the corrected series

All four corrected data sets were subjected to the same homogeneity tests as the original data. The results of these tests

of T_{\min} , T_{\max} , T_{aver} and DTR are shown in Figs. 16 and 17a–d (similarly to Fig. 13). As one can see from the comparison of the test statistics between the original (Figs. 12–13) and corrected (Figs. 16–17) data series, the last ones are less inhomogeneous but still contain inhomogeneities coinciding with the volcanic eruptions. Figure 16 (comparable with Fig. 12) shows the von Neumann ratio statistics of the corrected series. As can be seen, T_{\min} is less homogeneous than T_{\max}

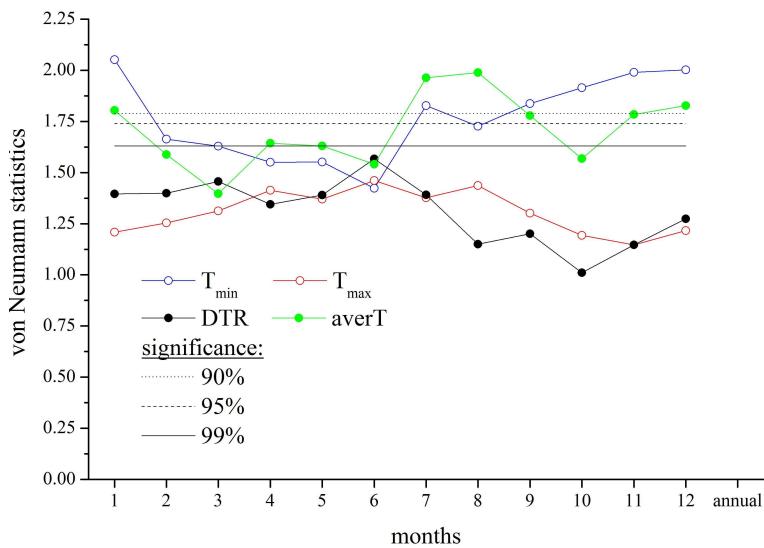


Figure 20. Coimbra: von Neumann ratio statistics for monthly series of T_{\min} , T_{\max} , DTR and T_{aver} . Black straight lines show probability levels.

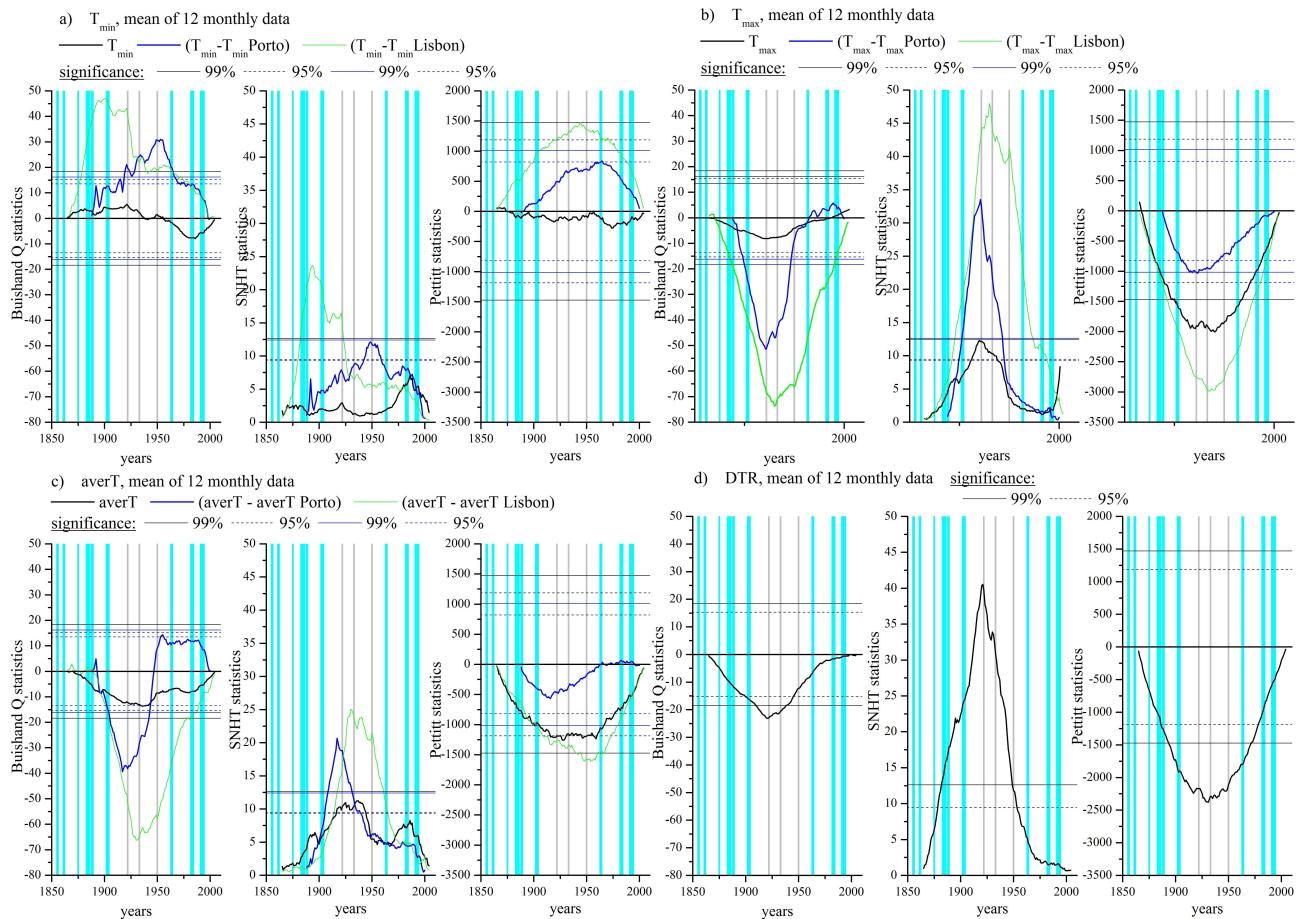


Figure 21. Coimbra: average of 12 monthly series of Buishand Q test (left panels), SNHT (middle panels) and Pettitt tests (right panels) statistics of T_{\min} (a), T_{\max} (b), T_{aver} (c) and DTR (d). Statistics of temperature series are shown in black, statistics of differences between Coimbra and Porto are shown in blue and between Coimbra and Lisbon are shown in green. Solid and dashed horizontal lines show probability levels of 99 % and 95 %, respectively. Grey vertical lines show known dates of instrumental changes, cyan broad vertical lines show periods of strong volcanic influence.

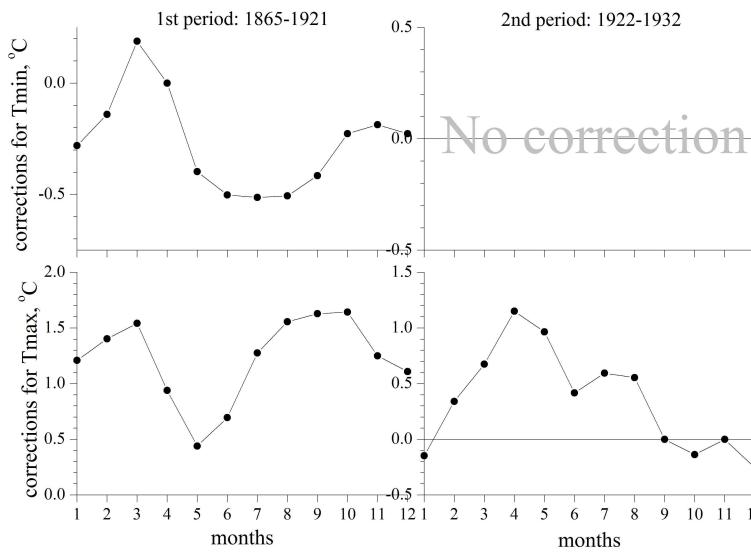


Figure 22. Coimbra: corrections (in $^{\circ}\text{C}$) for T_{\min} (top panels) and T_{\max} (bottom panels) series for two periods between the breaks; 1865–1921 (left panels) and 1922–1932 (right panels).

almost in all months. The possible reason for the remaining inhomogeneities in the T_{\min} series is the volcanic effect clearly seen in Fig. 17a–d.

Figure 18 shows CRMSEs of corrected series (SST Lisbon are used as reference series) plotted versus corresponding CRMSEs of the original series. As one can see, the inhomogeneity level of T_{\min} (left panel) slightly decreases – dots are lower than the bisect; on the contrary, the inhomogeneity level of T_{\max} (right panel) stays almost the same for 10 out of 12 monthly series but CRMSE of two monthly series slightly increases.

Sometimes the tests still show breaks of homogeneity in the period from 1917 to 1936 but there is no consistency between the three homogeneity tests (Buishand and Pettitt tests and SNHT) in the dates of the breaks. In our opinion, these inhomogeneities are caused again by the application of the smoothed correction values and we believe that in these cases additional corrections are not necessary. Thus, we consider the data sets of T_{\min} and T_{\max} corrected by the procedure described in the paper as free of non-climatic changes with a significance of at least 95 %.

5 Coimbra IGUC temperature series

5.1 Data description and metadata

The original data set contains monthly averages of daily minimum (T_{\min}) and maximum (T_{\max}) temperature and their annual means measured at Instituto Geofísico da Universidade de Coimbra (IGUC), Coimbra from 1865 to 2005. The data set length is 141 yr.

Accordingly to IGUC logbooks, during the entire period (1865–2005) the meteorological station remained in the same

location – the park of IGUC. However, the park of instruments has undergone some changes in position and environment described in Table 3. There were two more or less significant changes in the instruments location in 1922 and 1933; besides that, the standard (Stephenson's) shelter was installed in 1922 and in 1950 the thermometer height increased slightly (from 1.15 m to 1.45 m). Since Coimbra is not a coastal station, the already corrected temperature series for Porto and Lisbon were used as reference series; the differences between T_{\min} , T_{\max} and T_{aver} series and corresponding series for Porto and Lisbon were calculated to perform relative homogeneity tests.

5.2 Homogenization

5.2.1 Visual analysis

Figure 19a–d show time variations of the annual series of T_{\min} , T_{\max} , T_{aver} and DTR, respectively. The DTR variations show easily a visible break in 1921/1922 (relocation of the instruments and installation of the shelter) coinciding with a significant jump in T_{\min} (Fig. 19a), but not in T_{\max} (Fig. 19b). Another break probably appears in 1949/1950 (changes in thermometer height); it can be seen both in T_{\min} and T_{\max} data. This break is however absent in DTR data (probably, due to almost equal shifts in T_{\min} and T_{\max}). There is also a small break in 1932/1933 (small relocation).

5.2.2 Homogeneity tests results

The statistics of four homogeneity tests applied to this data set are shown in Figs. 20 and 21a–d. The von Neumann ratio test (Fig. 20) shows strong inhomogeneity of monthly values, which strongly depends on the temperature parameter:

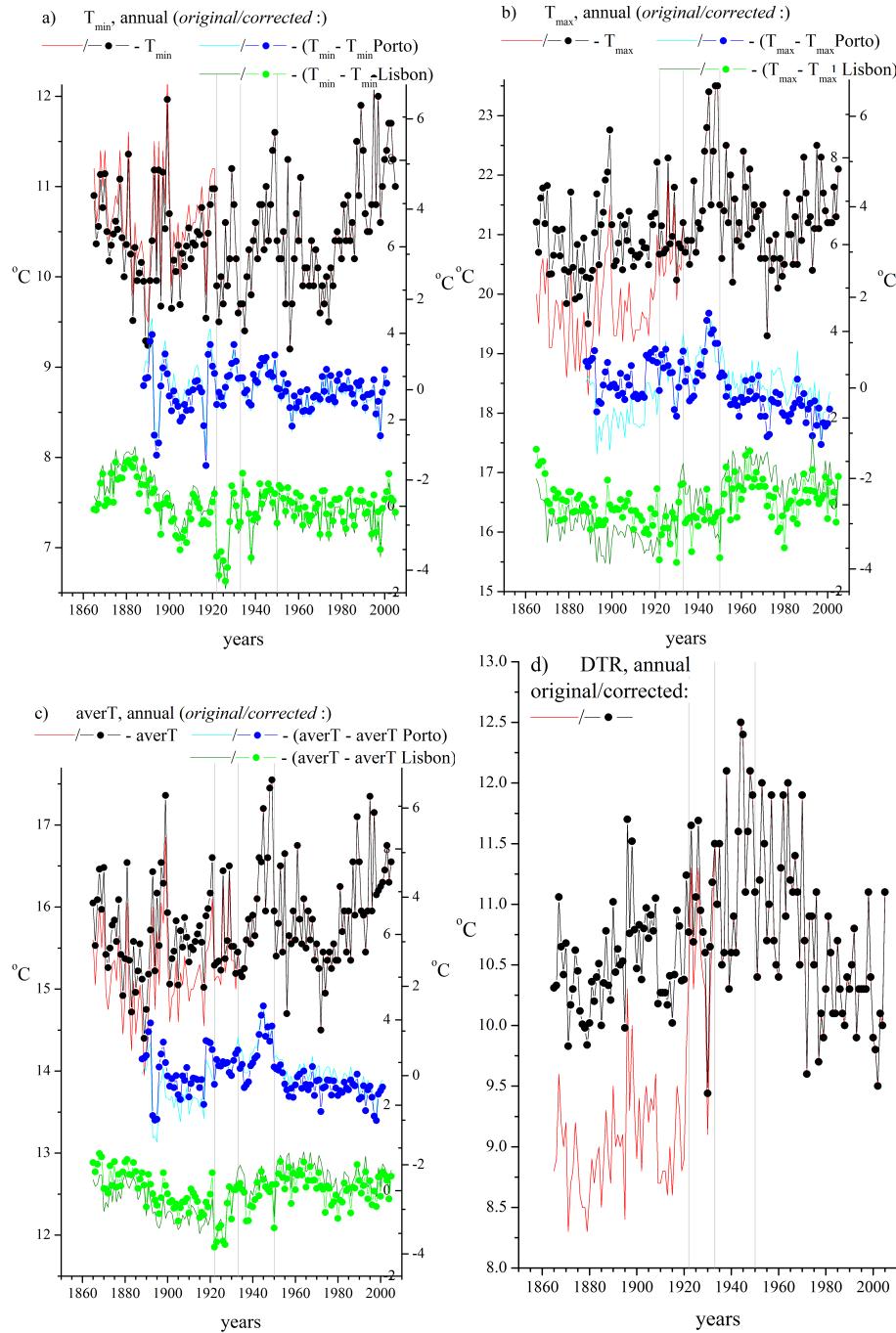


Figure 23. Coimbra: original and corrected annual series of T_{\min} (a), T_{\max} (b), averT (c) and DTR (d); Coimbra temperature series (red and black) and differences between Coimbra and Porto (cyan and blue) and Lisbon (dark green and green) temperature series. Grey vertical bands show dates of known instruments relocation (see Table 3).

- T_{\min} – data for months from February to June (m2 to m6) are more inhomogeneous than others;
- T_{\max} – all months show strong inhomogeneity;
- DTR – data for months of the second half of the year are more inhomogeneous than others;

– T_{aver} – data from February to June and October (from m2 to m6 and m10) are inhomogeneous (temperature data from Lisbon discussed in Sect. 4 show similar characteristics of the annual inhomogeneities variations).

The average of 12 monthly statistic series for other three homogeneity tests (in absolute and relative mode) applied to

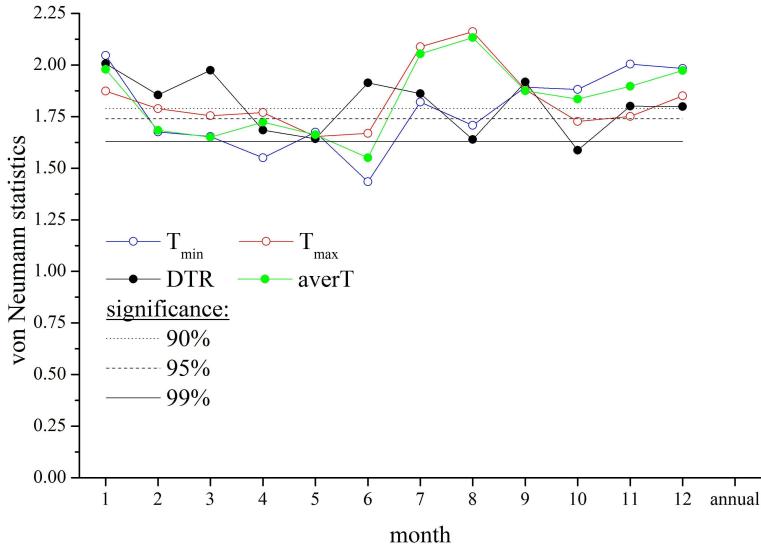


Figure 24. Coimbra: same as Fig. 20 but for the corrected series.

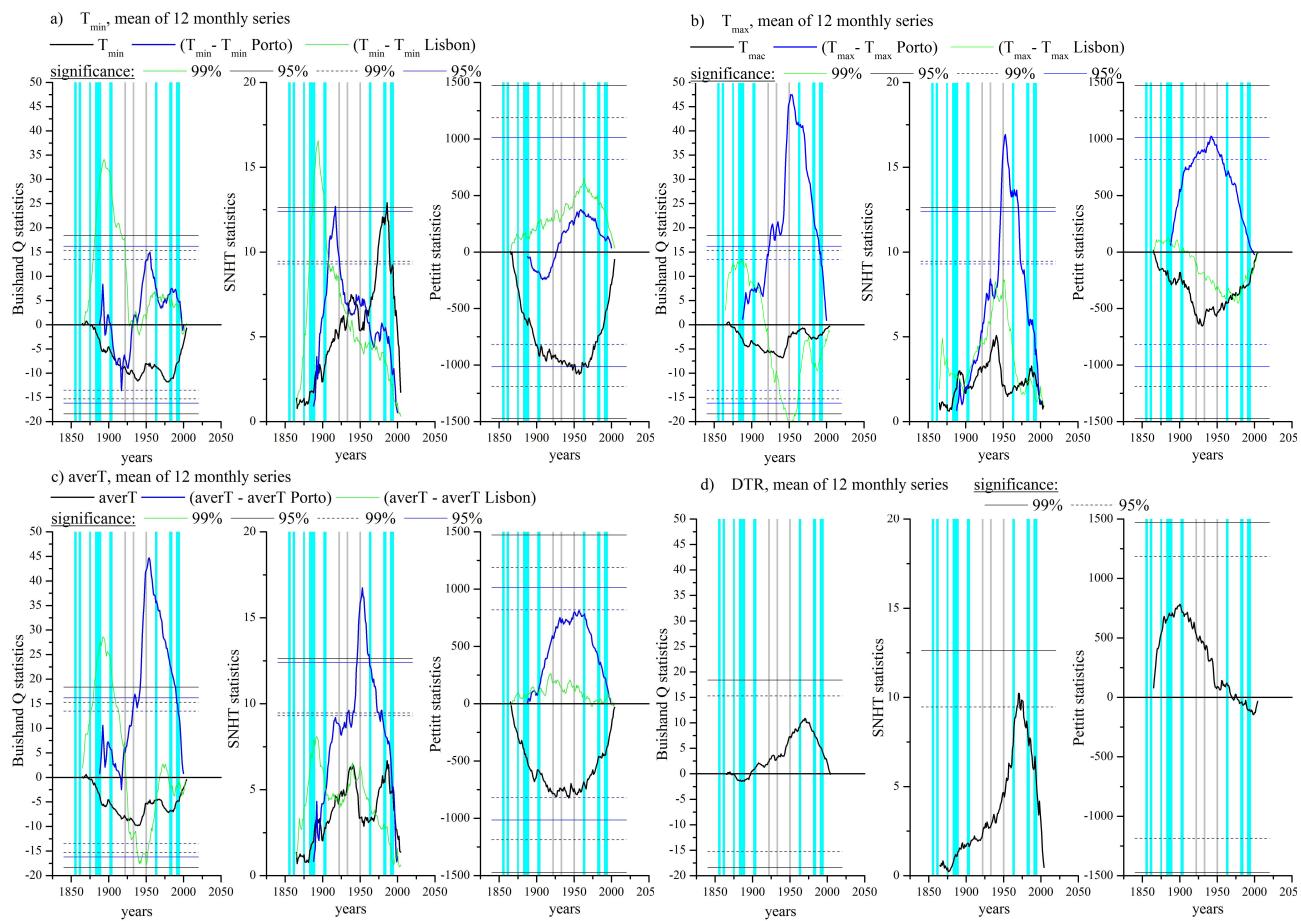


Figure 25. Coimbra: same as Fig. 21 but for the corrected series.

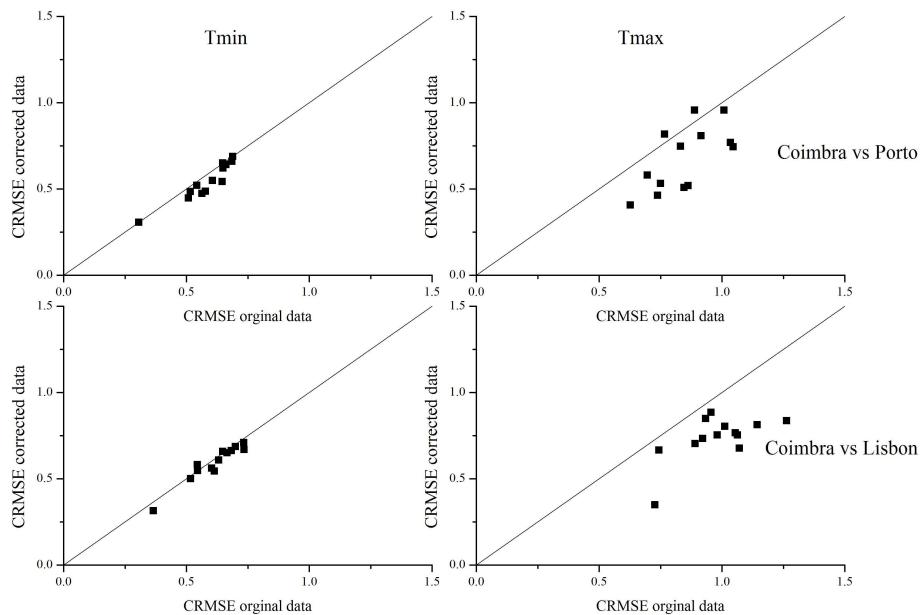


Figure 26. Coimbra: scatter plots of CRMSE before and after homogenization of 12 monthly and the annual series for T_{\min} (left) and T_{\max} (right). Corrected Porto (top panels) and Lisbon (bottom panels) temperature series are used as reference series. Dots on or below the bisect indicate data sets with unchanged or improved (increased) homogeneity, while dots above the bisect indicate data sets with increased inhomogeneity.

series of T_{\min} , T_{\max} , T_{aver} and DTR are shown in Fig. 21a–d, respectively. It should be mentioned that SNHT statistics, both for annual and monthly T_{\max} , show an unexpected behaviour: despite the absence of any jumps in the temperature data, the SNHT statistic shows strong inhomogeneities at the end of the data set (2002–2005). These inhomogeneities do not correlate with inhomogeneities detected on the same data by other tests. This unexpected behaviour could be explained by the known tendency of the SNHT to generate false alarm results close to the start and the end of data sets (Wang, et al., 2007). Therefore, to disambiguate the interpretation, Fig. 21b does not show SNHT statistics for T_{\max} during 2002–2005 yr.

The analysis of the homogeneity tests statistics provides the most probable time periods of the breaks in the data homogeneity: around 1885–1890, around 1905, around 1916, around 1920, around 1930–1936, in the 1940s, 1960s and 1980s. Many inhomogeneities, which are detected by the tests but could not be associated with known instrumental changes, correspond to volcanic effects.

The comparison between homogeneity test statistics of Coimbra and Lisbon data shows more or less a similar character of the annual inhomogeneities variations for both places. These similarities arise from the relative proximity of Lisbon and Coimbra and likeliness in the character of their climatic variation as well as from the volcanic origin of a number of inhomogeneities of the data.

5.2.3 Preliminary conclusions

T_{\max} data showed more inhomogeneities than other temperature parameters; T_{\min} data sets showed inhomogeneities near 1880s, 1900s, 1920s, 1960s and 1980–1990s; DTR and T_{aver} showed strong inhomogeneities around 1885–1890, around 1905, around 1916, 1922, around 1930–1936, around 1941, in the 1960s and 1980s; and T_{\min} and T_{aver} data had more inhomogeneities during warm months. The inhomogeneity levels of T_{\max} and DTR data were more or less constant throughout the year. The most significant non-climatic break occurred in 1922 due to changes in the instruments location. This break is clearly seen in relative homogeneity tests statistics both for T_{\min} and T_{\max} . Another break was associated with the small relocation of the instruments park in 1933. This break is seen only in relative homogeneity tests statistics for T_{\max} . These two breaks required correction. The change in the thermometer height in 1950 showed no significant (significance 95 % or more) effect on the homogeneity of the temperature data.

5.3 Correction for non-climatic breaks

5.3.1 Correction procedure

To correct the non-climatic breaks, T_{\min} and T_{\max} data sets were divided into three periods: 1865–1921, 1922–1932, and 1933–2004. We started from the most recent break – 1932/1933. This break was corrected only in T_{\max} series. For each month the means of T_{\max} for time intervals of ± 10 yr

around 1933 were calculated. The second break (1921/1922) was corrected both in T_{\max} and T_{\min} series using intervals of ± 10 yr for T_{\max} and ± 40 yr for T_{\min} . All correction values are shown in Fig. 22. Results of the corrections as well as original data are shown in Fig. 23a–d for annual series.

5.3.2 Homogeneity of the corrected series

All four corrected data sets were subjected to the same homogeneity tests as the original data. The results of these tests for T_{\min} , T_{\max} , T_{aver} and DTR are shown in Figs. 24 and 25a–d (similarly to Fig. 21). As one can see from the comparison of homogeneity test statistics of original (Figs. 20–21) and corrected (Figs. 24–25) series, the corrected data sets are less inhomogeneous. The statistics of the relative homogeneity tests show much less inhomogeneities in the corrected series than statistics of absolute homogeneity tests. The corrected series still contain inhomogeneities caused (probably) by the volcanic eruptions. It can be seen that, as a whole, the annual variations of the corrected data series homogeneity given by the von Neumann ratio still depends on the temperature parameter; T_{\min} is the less homogeneous among all parameters. Despite the fact that annual series still contain non-climatic inhomogeneities, monthly series, in most cases, are free of them. For a couple of months homogeneity tests still show breaks in homogeneity in the period from 1922 to 1933, but there is no consistency between the three homogeneity tests (Buishand and Pettitt tests and SNHT) in the dates of breaks. Figure 26 shows CRMSEs of corrected series (corrected Porto and Lisbon temperature series are used as reference series), plotted versus CRMSEs of the original series. As one can see, the inhomogeneity level of T_{\min} (left panels) decreases slightly – dots are close to the bisect, whereas on the contrary the inhomogeneity level of T_{\max} (right panels) significantly decreases for all monthly series when compared to Lisbon temperature series (low panel) and for 10 monthly series when compared to Porto temperature series (top panels). These homogeneity tests allow one to consider the corrected series of T_{\min} and T_{\max} as free of non-climatic changes with a significance of at least 95 %.

6 Conclusions

Homogeneity tests show the presence of strong non-climatic breaks in all temperature series. Most of the detected breaks were corrected and the homogeneity tests of the corrected series show no significant (significance 95 % or more) breaks around dates of instrumental changes.

6.1 Porto

One strong non-climatic break was detected in the temperature series of Porto Serra do Pilar, IGUP. This break was caused by the changes in the instruments location and

height (1916). This break did not coincide with known volcanic eruptions of significant strength and required correction. Other breaks detected by the homogeneity tests either had low levels of significance (lower than 95 %) or coincided with (probably caused by) strong volcanic eruptions. Such was the case of the possible non-climatic break in 1984, which could not be corrected due to the aforementioned coincidence. The break that took place in 1916 was corrected.

6.2 Lisbon

Two strong non-climatic breaks were detected in the temperature series of Lisbon, IGIDL. These breaks were caused by the changes in the instruments location (1864) and height (1941). These breaks were corrected. Other breaks detected by the homogeneity tests had low levels of significance (lower than 95 %).

6.3 Coimbra

Two strong non-climatic breaks were detected in the temperature series of Coimbra, IGUC. These breaks were caused by the changes in the instruments location (1922 and 1933). These breaks were corrected.

Acknowledgements. The authoresses would like to thank IGIDL, IGUC and IGUP for supplying the temperature data up to 1940. Also Instituto de Meteorologia, I.P. for supplying the post 1940 data and the UK Meteorological Office Hadley Centre for giving us the HadISST2.0.0.0 data.

We would like to thank personally Alexandra Pais, João Fernandes and Ivo Alves from Centro de Geofísica da Universidade de Coimbra and Maria José Neves Chorro currently from Universidade de Aveiro for supplying the data and for their helpful scientific discussions. We are indebted to Leopold Haimberger of the University of Vienna for reading the first draft and suggesting the use of reference series.

Anna Morozova was supported by a Post-Doc FCT scholarship (ref.: SFRH/BPD/74812/2010). This work was developed in the context of FP7 project ERA-CLIM (Grant Agreement Nr. 265229).

Edited by: G. König-Langlo

References

- Aguilar, E., Auer, I., Brunet, M., Peterson, T. C., and Wieringa, J.: Guidance on metadata and homogenization, WMO-TD No. M.1186, (WCDMP-No. 53), 1–55, 2003.
- Alexandersson, H. and Moberg, A: Homogenization of swedish temperature data. Part I: Homogeneity test for linear trends, Int. J. Climatol., 17, 25–34, 1997.
- Buishand, T. A.: The analysis of homogeneity of long-term rainfall records in The Netherlands, R. Neth. Meteorol. Inst. (K.N.M.I.), De Bilt, Sci. Rep. No. 81-7, 77 pp., 1981.
- Buishand, T.: Some methods for testing the homogeneity of rainfall records, J. Hydrol., 58, 11–27, 1982.

- Carvalho, A. C. F. J.: Tendências da média e variabilidade climática em algumas estações meteorológicas de Portugal continental, Licenciatura Dissertation, Faculdade de Ciências da Universidade de Lisboa, 59 pp., 2001.
- Costa, A. C. and Soares, A.: Homogenization of climate data: review and new perspectives using geostatistics, *Math. Geosci.*, 41, 291–305, 2009.
- Gleckler, P. J., Taylor, K. E., and Doutriaux, C.: Performance metrics for climate models, *J. Geophys. Res.*, 113, D06104, doi:10.1029/2007JD008972, 2008.
- Khaliq, M. N. and Ouarda, T. B. M. J.: On the critical values of the standard normal homogeneity test (SNHT), *Int. J. Climatol.*, 27, 681–687, 2007.
- Klein Tank, A. M. G.: Algorithm Theoretical Basis Document (ATBD), European Climate Assessment & Dataset (ECA&D) project document, KNMI, 38 pp., 2007.
- Mann, M. E., Gille, E. P., Bradley, R. S., Hughes, M. K., Overpeck, J. T., Keimig, F. T., and Gross, W. S.: Global Temperature Patterns in Past Centuries: An Interactive Presentation, *Earth Interact.*, 4, 1–29, 2000.
- Martínez, M. D., Serra, C., Burgueño, A., and Lana, X.: Time trends of daily maximum and minimum temperatures in Catalonia (ne Spain) for the period 1975–2004, *Int. J. Climatol.*, 30, 267–290, doi:10.1002/joc.1884, 2010.
- NCDC database: http://www.ncdc.noaa.gov/paleo/ei/ei_data/volcanic.dat, last access: 19 September 2011.
- Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullette, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Førland, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., and Parker, D.: Homogeneity adjustments of in situ atmospheric climate data: a review, *Int. J. Climatol.*, 18, 1493–1517, 1998.
- Pettitt, A. N.: A Non-Parametric Approach to the Change-Point Problem, *Appl. Stat.-J. Roy. St. C.*, 28, 126–135, 1979.
- Pinhal, E. O. S.: Variabilidade interdecadal da temperatura, pressão e precipitação do Observatório da Serra do Pilar, Teses de mestrado, Universidade de Lisboa, Lisboa, 96 pp., 2008.
- Rayner, N. A., Kennedy, J. J., Titchner, H. A., Saunby, M., and Saunders, R. W.: The development of the new Hadley Centre Sea Ice and Sea-surface temperature dataset, HadISST2, to explore uncertainty in boundary forcing for reanalysis, Poster Presentation, 4th World Climate Research Programme International Conference on Reanalyses, Silver Spring, Maryland, USA, 7–11 May 2012.
- Robock, A.: Volcanic eruptions and climate, *Rev. Geophys.*, 38, 191–220, 2000.
- Solow, A.: Testing for climatic change: an application of the two-phase regression model, *J. Clim. Appl. Meteorol.*, 26, 1401–1405, 1987.
- Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, 106, 7183–7192, 2001.
- Venema, V. K. C., Mestre, O., Aguilar, E., Auer, I., Guijarro, J. A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P., Viarre, J., Müller-Westermeier, G., Lakatos, M., Williams, C. N., Menne, M. J., Lindau, R., Rasol, D., Rustemeier, E., Kolokythas, K., Marinova, T., Andresen, L., Acquaotta, F., Fratianni, S., Cheval, S., Klancar, M., Brunetti, M., Gruber, C., Prohom Duran, M., Likso, T., Esteban, P., and Brandsma, T.: Benchmarking homogenization algorithms for monthly data, *Clim. Past*, 8, 89–115, doi:10.5194/cp-8-89-2012, 2012.
- von Neumann, J.: Distribution of the Ratio of the Mean Square Successive Difference to the Variance, *Ann. Math. Stat.*, 12, 367–395, 1941.
- Wang, X. L., Wen, Q. H., and Wu, Y.: Penalized Maximal t Test for Detecting Undocumented Mean Change in Climate Data Series, *J. Appl. Meteorol. Clim.*, 46, 916–931, 2007.
- Wijngaard, J. B., Klein Tank, A. M. G., and Können, G. P.: Homogeneity of 20th century European daily temperature and precipitation series, *Int. J. Climatol.*, 23, 679–692, 2003.