HOMOGENIZATION OF DAILY AIR PRESSURE AND TEMPERATURE SERIES FOR BRNO (CZECH REPUBLIC) IN THE PERIOD 1848–2005

P. Štěpánek[1], L. Řezníčková[2], R. Brázdil[2]

[1] Czech Hydrometeorological Institute, Regional Office Brno, Czech Republic
[2] Institute of Geography, Masaryk University, Brno, Czech Republic
petr.stepanek@chmi.cz

ABSTRACT

Homogenization of daily meteorological series is a difficult task. Several kinds of problem have to be taken into consideration in the course of homogenization: selection of a proper homogenization method with regard to the data used, creation of reference series, completion of missing values, annual course of adjustments, and others. This paper presents an attempt to create a homogeneous series of daily air pressure and temperature readings in the city of Brno (Czech Republic). Two basic approaches were adopted: (i) homogenization of monthly series and projection of estimated smoothed monthly adjustments in annual variation of daily adjustments and (ii) homogenization of daily values in individual months and direct estimation of daily adjustments, again smoothed by low-pass filter. Differences in the results obtained from these two approaches are further discussed.

INTRODUCTION

In the recent years considerably more attention has been devoted to the analysis of the daily data widely recorded and stored in databases. Prior to analysis, the need to homogenize the data and check their quality arises. There is no widely accepted homogenization approach that could be generalized and applied to various meteorological elements, different climatic patterns, etc., and this will probably never be possible. This is, for example, due to the fact that the statistical properties of daily data and regional differences between them make general homogenization of daily values difficult, as well as involving more demanding data handling. During data processing, several kinds of problem have to be taken into consideration. These involve selection of a proper method for homogenization with regard to the data used, i.e. fulfilling all the conditions necessary to applying selected tests of relative homogeneity (e.g. normal distribution), creation of reference series (defining selection criteria), completion of missing values, annual course of adjustments, and others.

Only a few studies, in comparison with monthly or annual data series, have been devoted to techniques addressing daily values. For example, Brandsma (2000) compared monthly adjustments, daily adjustments derived from monthly adjustments (using iterative cubic spline interpolation to preserve monthly adjustments) and daily adjustments derived from weather types. Wijngaard et al. (2003) did not use measured values, but their characteristics, such as diurnal temperature range and its annual mean, as well as the annual mean of the absolute day-to-day differences for temperature, and the annual number of wet days for precipitation. Following various homogeneity tests, these series were labelled as recommendations for further analysis.

Mekis and Vincent (2004) derived daily adjustments from monthly adjustments. These were obtained using linear interpolation between mid-month “target” values objectively chosen so that the average of the daily adjustments over a given month is equal to the monthly adjustment. This approach does not require the creation of a daily reference series or the identification of inhomogeneities in daily temperatures. Moreover, finally homogenized series of daily temperatures are compatible with homogenized monthly data-sets.

The present paper is dedicated to the search for a proper methodology for daily data-set handling and its subsequent application to daily air pressure and temperature series for Brno in the period 1848–2005, with the aim of creating a homogeneous series for Brno with regard to both elements. Although there are, in general, no gaps in the Brno measurements, data are unfortunately not available from a single site, so it becomes necessary to combine different series to get one Brno series suitable for further analysis. The basic Brno stations were tested separately for relative homogeneity and, after homogenization, they were combined using overlap periods. All calculation was performed using AnClim and ProClimDB software (Štěpánek, 2006a, 2006b).

1. A BRIEF HISTORY OF METEOROLOGICAL OBSERVATIONS IN BRNO

Meteorological observations in Brno began in 1799, the work of Captain Emeritus Ferdinand Knittelmayer, but his observations for the period 1799–1812 are preserved only in the daily averages. For the subsequent years 1813–1819, the observations exist only in the form of monthly averages. On the basis of several daily readings, meteorological observations were published in the daily newspaper “Mährisch-Ständische Brünner Zeitung” from January 1820 to December 1847. In some years, parallel observations from two places in Brno were also made. Although monthly value series for air pressure, air temperature and precipitation totals have been homogenised and analysed (Brázdil et al., 2005), work with daily readings or daily averages requires further research. For this reason, the analysis provided in this paper works only with data from 1848 onwards.

Meteorological observations after 1848 come from Dr. Paul Olexik (1800–1878), a physician from St. Anne’s hospital (Fig. 1). He was probably making observations from as early as the end of 1845, but it was only from 1848 that his measurements started to be published regularly in the Austrian Meteorological Yearbooks, i.e. when his station became part of the network of the Central Meteorological Institute in Vienna. He observed at 0600, 1400 and 2200 hours. On 3 December 1853 he moved the point of his meteorological observations from the hospital (204 m.a.s.l.) a short distance, to his new flat at Pekařská Street 100 (219 m.a.s.l.). Meteorological observations at this new site continued until 30 June 1878. By this time, Gregor Johan Mendel (1822–1884), abbot of the Augustinian Monastery and a pioneer geneticist, was helping to complement Olexik’s measurements, something he continued alone from 1 July 1878 in the monastery garden (204 m a.s.l.) until 30 November 1883. He began with standard readings at 0700, 1400 and 2100 hours. Alfred Lorenz (1825–1890), a professor at the I. R. Technical University, continued meteorological observations in Brno from the university building (225 m a.s.l.), located close to the city centre, from 1 January 1884 until his death in June 1890. Upon his death, air temperature and pressure measurements definitely stopped and no new place for observation was to be found (Brázdil, 1979).

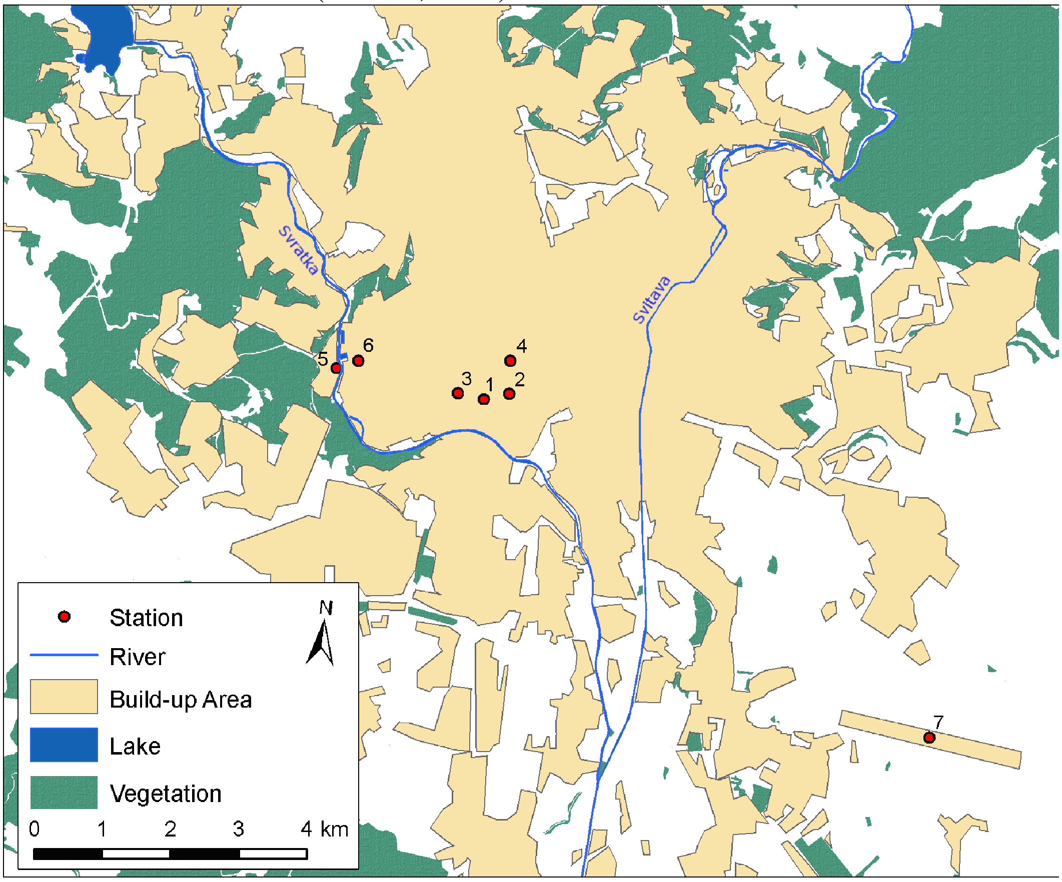


Fig. 1. Location of meteorological stations in Brno: 1 – St. Anne's hospital; 2 – Pekařská Street 100; 3 – Augustinian monastery; 4 – I. R. Technical University; 5 – Pisárky, waterworks; 6 – Květná Street; 7 – Tuřany Airport

However, from 1 June 1890 a meteorological station at the city waterworks in Pisárky (204 m a.s.l.) (further as Brno-Pisárky) began operations, with a full observation programme up to 1937 and with air temperature measurements continuing up to 1962. Further meteorological stations in different parts of Brno were established later, of which only the two used in this paper are mentioned. The first of them was located close to the previous station on Květná Street (further as Brno-Květná), in the garden of the research agricultural institute (223 m.a.s.l.), with observations from 1 August 1922 to 31 December 1970. The second station (Brno-Tuřany) is located at the Brno airport, south-east of the city of Brno (238 m.a.s.l.), i.e. opposite to all the previously mentioned stations, which are concentrated in its western part. Observations started there on 14 April 1958. For this reason, compilation of the Brno daily temperature and pressure series is made with respect to this station.

In summary, addressing knowledge of the history of temperature and pressure measurements in Brno from 1848 onwards, with respect to homogenization, it should be stressed that measurements

- were provided from different parts of Brno, at different altitudes
- were provided by different types of instruments
- were provided in different observation terms before and after 1878
- are limited by lack of available overlap for observations predating 1890.

2. DATA USED

For outlier identification as well as for relative homogeneity testing, other stations with long-term series in the broad surroundings of Brno were also used (Fig. 2). A list of them, with basic characteristics, is given in Table 1. We have used all the measurements, i.e. not only daily averages but also separate series from individual observation hours. As has already been mentioned, as well as standard observations times at 0700, 1400 and 2100 hours local mean time, observations were also carried out at 0600, 1300 and 2200 hours. Finally it was decided, that all the terms should further be treated as if they were 0700, 1400 and 2100 hours in the hope of disclosing possible inhomogeneities arising out of various observing times during homogeneity testing. The original observing hours were taken into consideration during decision-making about adjustments of inhomogeneities found.

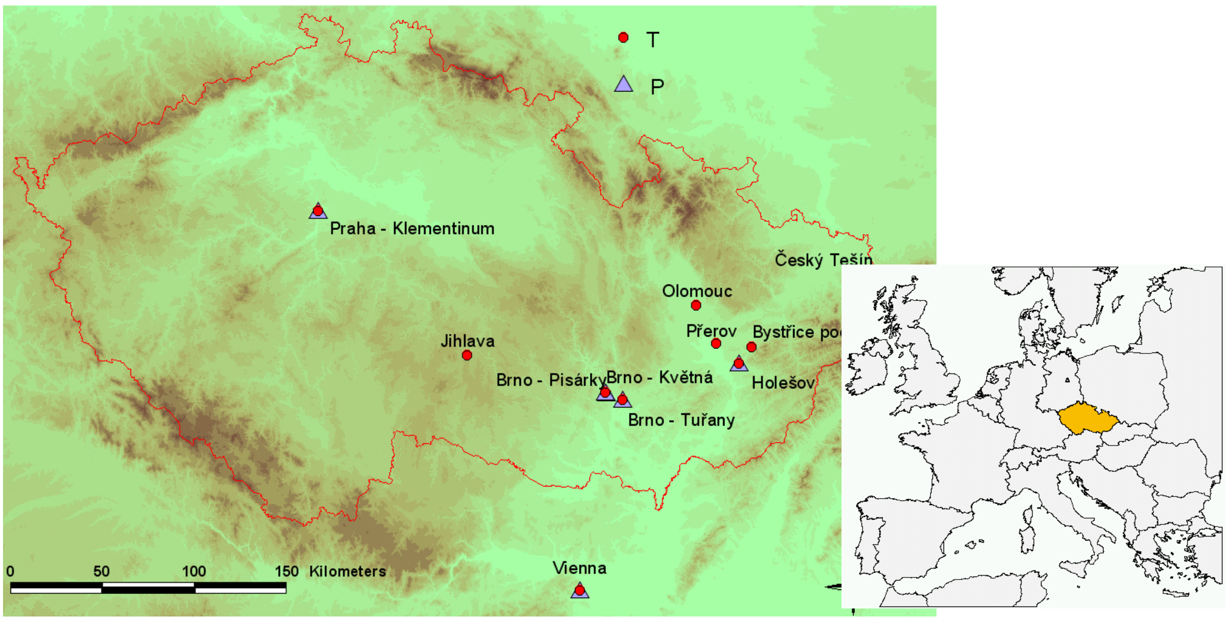


Fig. 2. Geographical distribution of stations used for homogenization of the Brno series (T – air temperature, P – air pressure)

Table 1. Basic information about stations used for homogenization of the Brno series (station coordinates are given for their last or recent locations)

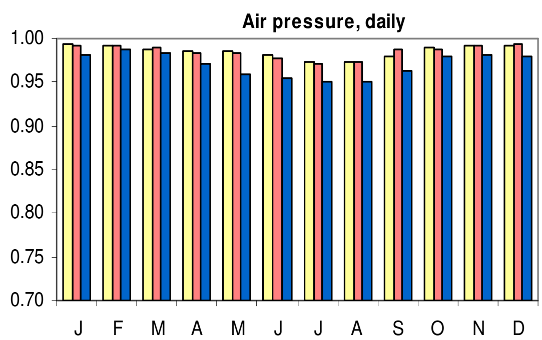
Air temperature

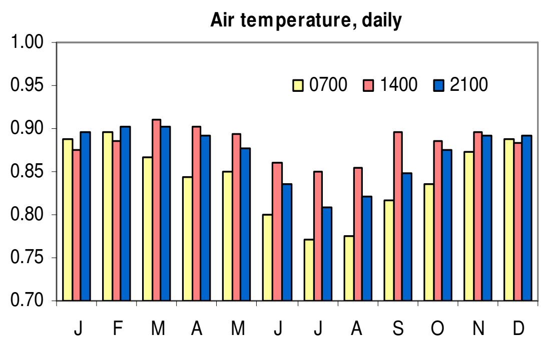
Station name	Latitude (N)	Longitude (E)	Altitude (m.a.s.l.)	Beginning	End	Observing hours
Brno (various places)	49°12´	16°37´	225	1 Jan. 1848	31 Dec. 1889	07 (06), 14 (13), 21 (22)
Brno-Pisárky	49°12´	16°34´	203	1 June 1890	31 May 1962	07, 14, 21
Brno-Květná	49°12´	16°34´	223	1 Aug. 1922	31 Mar. 1970	07, 14, 21
Brno-Tuřany	49°09´	16°ˇ42´	241	14 Apr. 1958	31 Dec. 2005	07, 14, 21
Bystřice pod Hostýnem	49°24´	17°40´	315	1 Sep. 1865	31 Dec. 2005	07 (06), 14, 21 (22)
Český Těšín	49°44´	18°37´	280	1 Jan. 1885	31 Oct. 1938	07, 14, 21
Holešov	49°19´	17°34´	224	1 July 1895	31 Dec. 2005	07, 14, 21 (22)
Jihlava	49°23´	15°32´	560	27 July 1873	31 Dec. 1934	07 (08), 14, 21 (22)
Olomouc	49°36´	17°15´	215	1 Jan. 1876	31 Dec. 1960	07 (08), 14, 21 (20)
Prague-Klementinum	50°05´	14°25´	191	1 Jan. 1775	31 Dec. 2005	07, 14, 21
Přerov	49°25´	17°24´	203	1 Apr. 1874	31 Dec. 1979	07, 14, 21
Vienna-Hohe Warte	48°13´	16°21´	199	1 Jan. 1872	31 Dec. 2005	07, 14, 19

Air pressure

Station name	Latitude (N)	Longitude (E)	Altitude (m.a.s.l.)	Beginning	End	Observing hours
Brno (various places)	49°12´	16°37´	225	1 Jan. 1848	31 Dec. 1889	07 (06), 14 (13), 21 (22)
Brno-Pisárky	49°12´	16°34´	203	1 June 1890	31 Dec. 1937	07, 14, 21
Brno-Květná	49°12´	16°34´	223	1 Aug. 1922	31 Dec. 1962	07, 14, 21
Brno-Tuřany	49°09´	16°42´	241	14 Apr. 1958	31 Dec. 2005	07, 14, 21
Holešov	49°19´	17°34´	224	1 Jan. 1961	31 Dec. 2005	07, 14, 21
Prague-Klementinum	50°05´	14°25´	191	1 Aug. 1787	31 Jan. 2002	14
Vienna-Hohe Warte	48°13´	16°21´	199	1 Jan. 1872	31 Dec. 2005	07, 14, 19

The correlation coefficients for both elements analyzed are high enough for all stations involved (Fig. 3). Their values were calculated from original data (not from series of first differences), so they are biased by inhomogeneities in a shift (the values would otherwise be higher) and also by trends (the values would be lower if the trend were removed from the series). Correlations of monthly averages are higher than those of daily averages during the winter months, while the opposite holds in summer, i.e. the correlations of monthly averages drop below the values of daily data. From this it follows that both monthly and daily data should be used for data homogenization; daily data are more sensitive to inhomogeneity detection, especially during the summer months.





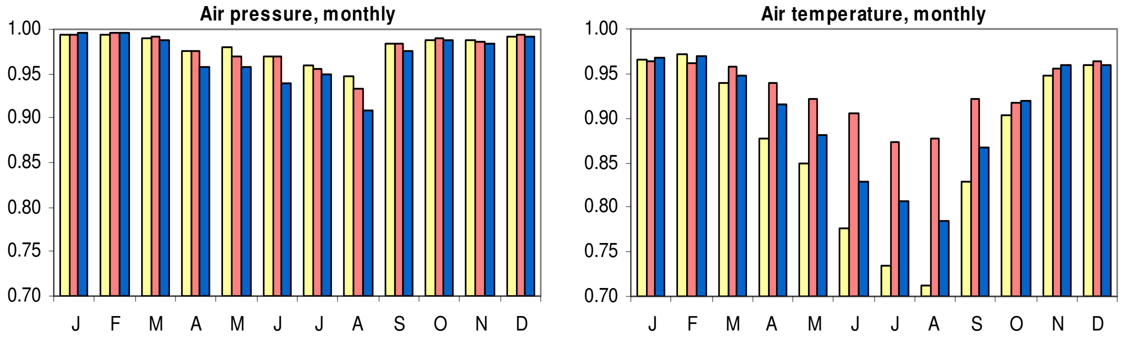


Fig. 3. Medians of correlation coefficients for all pairs of stations, for daily and monthly air temperature (50 values) and air pressure (6 values – without Prague-Klementinum at 1400 hours)

3. HOMOGENIZATION

Homogenization includes the following steps: detection, verification and possible correction of outliers (extreme values), creation of reference series, homogeneity testing (various homogeneity tests), determination of inhomogeneities in the light of test results and metadata, adjustment of inhomogeneities and filling in missing values (Fig. 4).

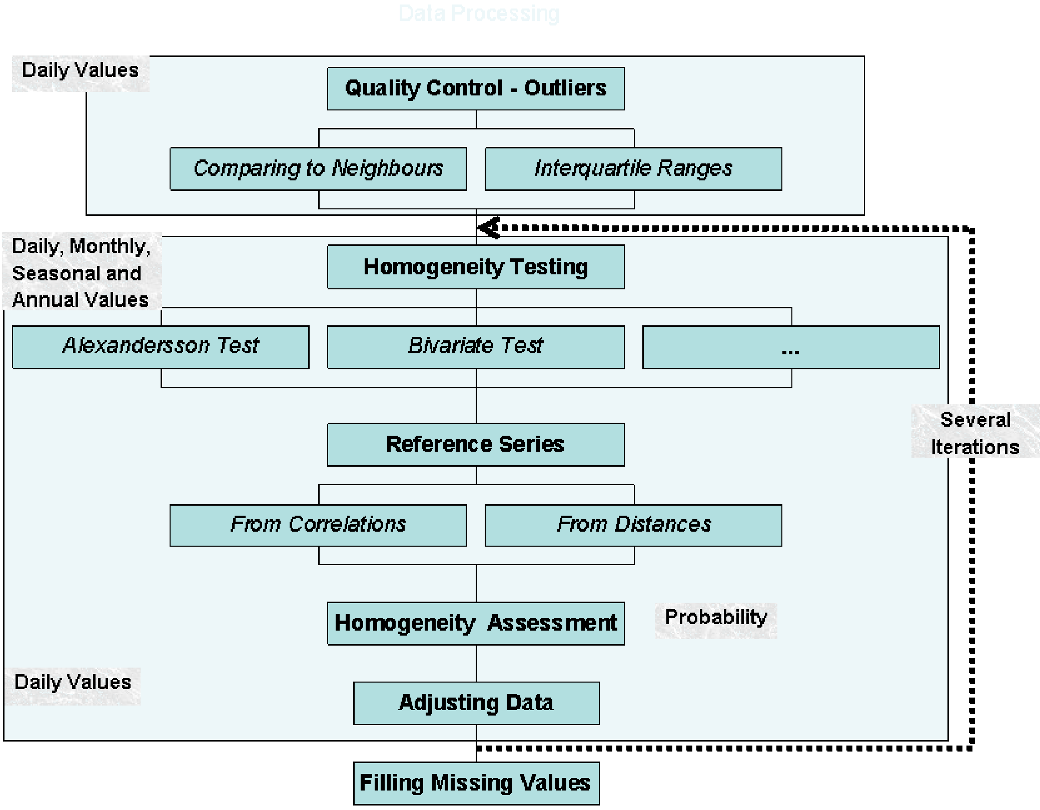


Fig. 4. Plan of the homogenization process

3.1 Outlier identification

Data quality control was carried out in two ways in this study: (i) by applying limits derived from interquartile ranges (either to individual series, i.e. absolutely or, better, to difference series between candidate and reference series, i.e. relatively), (ii) by comparing candidate station values to values from neighbour stations.

In comparisons with neighbour stations, the five best correlated series were selected (correlations calculated from series of first differences – see e.g. Peterson, 1998), the values of correlation coefficients being at least 0.50; no limit for distance or altitude difference has been applied. Only series with the same observation hours were selected. For the evaluation of outliers, various characteristics were considered. A count of statistically significant different neighbours (compared to candidate station) exceeding the confidence limit (0.95) was evaluated by means of difference series (neighbour minus candidate station), for each month individually. Cases in which more than 75% of neighbours differed significantly from the base station values were checked visually. To help in establishing the nature of the outliers, the values of neighbours were standardized with respect to candidate station average and standard deviation and a new (theoretical) value for the candidate station was also calculated – as a weighted average from the standardized values of the neighbours. Further, the coefficient of interquartile ranges (q75–q25) above q75 (or below q25) were evaluated (calculated from the standardized neighbour values), and applied to candidate station value. The reason for this was to assess similarity of neighbour values used with regard to test value: the more values of neighbours are similar, the higher is the value of the coefficient.

The final decision on removing outliers was based on the percentage of the count of significantly different neighbours, difference from "expected value", coefficient of interquartile range, and finally by visual (subjective) comparison of the standardized values of neighbours with the candidate station values. Fig. 5 shows an example of the output for decision-making about outliers.

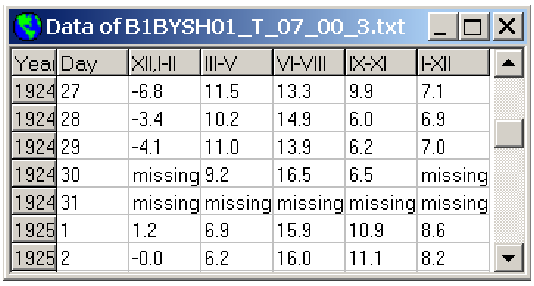
Id	Year	Month	Day	St_base	Remark	St_1	St_2	St_3	St_4	St_5	Exc_	Exc_	St_1_ba	St_2_ba	St_3_b	St_4_	St_5_	Expect_val
B1BYSH01_T_07:00				315.0	Altitudes,limits	282.0	203.0	560.0	215.0	199.0								
O1TESI01_T_07:00					st_1, distance	111.7												
O3PRER01_T_07:00					st_2, distance		29.0											
B2JIHL01_T_07:00					st_3, distance			237.1										
O2OLKL01_T_07:00_					st_4, distance				50.8									
B6VHOH01_T_07:00					st_5, distance					192.6								
B1BYSH01_T_07:00	1886	1	1	-13.6		-6.7	-12.8	-11.5	-10.0	-11.1	1	20.0	-6.9	-12.9	-11.3	-9.9	-13.1	-11.9
B1BYSH01_T_07:00	1886	1	2	-7.4		-2.4	-7.2	-1.0	-6.8	-10.6	1	20.0	-3.0	-7.3	-0.5	-6.7	-12.5	-7.0
B1BYSH01_T_07:00	1886	1	3	1.6		2.6	1.0	1.8	1.9	4.7			1.4	0.9	2.4	2.0	3.8	1.2
B1BYSH01_T_07:00	1886	1	4	2.0		3.4	0.9	1.5	2.2	1.1			2.2	0.8	2.1	2.3	-0.1	1.2
B1BYSH01_T_07:00	1886	1	5	4.4		4.4	3.7	4.5	2.0	0.1	1	20.0	3.1	3.6	5.2	2.1	-1.1	3.2
B1BYSH01_T_07:00	1886	1	6	1.6		2.6	0.6	0.9	0.6	3.7			1.4	0.5	1.4	0.7	2.7	0.6

Fig. 5. Example of output with auxiliary characteristics for quality control evaluation

In some cases, in which at least two neighbours were not available, interquartile ranges for each individual month of the candidate series were applied (i.e. absolutely) and the errors emerging were checked. This method has considerably inferior results in comparison with the relative method, but no other possibility existed for cases in the distant past.

3.2 Homogeneity test

As well as monthly, seasonal and annual averages, series of daily data were also tested. In this case we used all days of a particular month and further an aggregation of "seasons and year" calculated from the first days of all months, the second days, etc. (see Fig. 6). Although such "aggregate" series cannot be used for common time series analysis because the time is "cracked", it can be very useful for the purposes of finding discontinuity (seasonal to annual resolution), while original daily values, even when used only within particular months, can suffer from annual course (this is the case for air temperature rather than air pressure, mainly in winter) and normality is sometimes on the border of the 0.05 significance level. Using the aggregates over seasons and year leads to series for which normality is fulfilled without problems, and thanks to lower signal-to-noise ratio this approach is better for detecting real inhomogeneities in the series. Significant autocorrelations within a number of first lags (days) appear to present a larger problem and have to be further investigated. Series are more persistent in winter with stronger circulation effects, rather than in summer with its prevailing radiation factors.



Data of B1BYSH01_T_07_00_3.txt

Yea	Day	XII,I-II	III-V	VI-VIII	IX-XI	I-XII
1924	27	-6.8	11.5	13.3	9.9	7.1
1924	28	-3.4	10.2	14.9	6.0	6.9
1924	29	-4.1	11.0	13.9	6.2	7.0
1924	30	missing	9.2	16.5	6.5	missing
1924	31	missing	missing	missing	missing	missing
1925	1	1.2	6.9	15.9	10.9	8.6
1925	2	-0.0	6.2	16.0	11.1	8.2

Fig. 6. An example of using daily data for homogeneity testing

Several relative homogeneity tests (significance level 0.05) were used: the Alexandersson Standard Normal Homogeneity Test SNHT (Alexandersson, 1986, 1995), the Maronna and Yohai bivariate test (Potter, 1981), the Pettit test (Pettit, 1979), the t-test (Mitchell et al., 1966) and the Easterling and Peterson test (Easterling, Peterson, 1995). Tests were applied to 40-year sections of the series tested for monthly averages and 30-years series of daily data because the alternative hypothesis of the Alexandersson and bivariate tests assumes the presence of only one inhomogeneity in a series (we applied SNHT for a single shift). Series longer than 40 years were divided into several parts with an overlap of ten years (or five years for daily data). This is important in the light of tendencies to overestimation of detected inhomogeneities near the ends of series (see Alexandersson, 1995). Reference series were created separately with respect to each 40-year (30-year) parts of a candidate series (this means with its own selection of neighbours in each part). For daily data, 185 sections of series (of 49 original elements-terms-stations) were created and tested.

The use of series with durations of 40 and/or 30 years seems to be reasonable for homogeneity testing. Shorter series would not be so suitable from a statistical point of view, while, on the other hand, longer series usually contain more than one inhomogeneity (the typical duration of a period with one inhomogeneity does not usually exceed 30–40 years – see e.g. Auer et al., 2001).

To ensure that only one inhomogeneity detected by the Alexandersson or bivariate tests was present in a series, a further modification was introduced into the AnClim software. The series was divided at the position of a detected inhomogeneity and sections before and after it were tested separately. If no other inhomogeneity was found, we can rely on the results of the given test for the whole length of the series (especially the significance of a test statistic).

3.3 Reference series creation

Reference series were created in two ways: (i) an average from the best correlated stations, (ii) an average from nearest stations. Correlation coefficients used for station selection were calculated from the series of first differences, when inhomogeneities are manifested in the only value (see e.g. Alexandersson, Moberg, 1996; Peterson, 1998). Various types of reference series with analysis of their advantages and drawbacks have been discussed, for example, by Štěpánek (2005).

The values of correlation coefficients were not allowed to drop below 0.60 between neighbour stations (selection by means of correlation) and no distance or altitude limits were applied as additional conditions for air pressure and temperature. Weighted averages were calculated using correlations and/or reciprocal values of station distances as weights. Values of selected neighbour stations were standardized to candidate station average and standard deviation to avoid problems with biased reference series. This can often happen in the event of missing data in one of the neighbour series. The standardization was done for each particular month individually (also for daily data). No transformation of values has been applied to the data.

In the first stage, a list of proposed neighbour stations was obtained, which was subsequently checked and its approved version was then finally used for the reference series calculation.

3.4 Assessment of detected inhomogeneities

The main criterion for determining a year of inhomogeneity was the probability of the given inhomogeneity, i.e. the ratio between the count of detections for a given year from all tests for the given station (using all types of reference series, tests, daily, monthly, seasonal and annual series) and the count of all theoretically possible detections. The count of detections for groups of years was also taken into account (some inhomogeneities started in the course of the year and thus were manifested in at least two years). If metadata did not confirm the detected shift (in most cases), the percentage limit of all possible detections was taken higher and some other information (e.g. distribution of the given year within individual months or seasons, graphs of differences with reference series and some other characteristics) was required to decide whether the undocumented inhomogeneity could be regarded as "indubitably" proven and consequently corrected. For assessment of the inhomogeneities detected, the Real Precision Index (RPI – see Petrovic, 2004) may also be applied to find sections of series that exhibit change in the quality of measurements.

3.5 Adjustment of inhomogeneities

Adjustment of inhomogeneities detected was addressed by means of the reference series calculated from the average over the five stations with the highest correlation coefficients with the series being adjusted (correlations were calculated again from the series of first differences). The adjustment value was estimated as the difference between averages calculated from difference series between the tested and the reference series. The start of inhomogeneity was allocated to a particular month (where this was possible).

When dealing with daily data, there are several approaches to adjusting data for inhomogeneities detected. We may use either monthly adjustments which can be distributed into individual days (e.g., Mekis and Vincent, 2004) or we can calculate adjustments for daily data directly.

Using monthly data in this paper, the estimated individual monthly adjustments were smoothed by low-pass filter (weights applied to adjacent months were approximately 1, 2, 1) to suppress the influence of random errors in the series (the effect of smoothing results in a more realistic annual course for the adjustments, in line with what is better physically justified). The monthly adjustments obtained were then distributed (interpolated) among individual days and the final daily adjustments (again possibly smoothed to eliminate the edges of lines occurring each month) were then applied to data.

In the second case, the daily adjustments difference series (reference and tested) for each day of the year were used, taking 20 years before and after the change. Final daily adjustments were then smoothed using a low-pass filter for 60 days (to each side).

Various characteristics were analyzed before applying the adjustments: increment of correlation coefficients between candidate and reference series after adjustments, change of standard deviation in differences before and after the change, presence of linear trend, etc. In the event of any doubts, the adjustments were not applied.

3.6 Further considerations

The above-mentioned steps were performed in several iterations. At each iteration, more precise results were obtained. Missing values were filled in only after homogenization and adjustment of inhomogeneities in the series. The reason for this was that the new values were estimated from data not influenced by possible shifts in the series. Moreover, when missing data are filled in before homogenization, they may influence inhomogeneity detection in a negative way. The gaps were filled by means of linear regression between filled value series (dependent variable) and a reference series (independent variable), separately for each month. For assessing the quality of the process, various statistics were monitored, e.g. differences of averages and standard deviations in periods before and after the gap.

4. HOMOGENIZATION RESULTS

As has been shown above, the values of correlation coefficients for daily data (using each month individually) are comparable with values gained from monthly averages. The same holds true of correlations between tested and reference series. The medians of correlation coefficients for monthly air temperature range from 0.87 in the summer months to 0.98 in the winter months for individual observation hours; again the results at 1400 hours correlate the best. For daily data, the correlations for individual months range from 0.87 to 0.95. For air pressure, daily data correlates between 0.97 in summer and 0.99 in winter, monthly data between 0.94 and 0.99.

From these results, it follows that it is worth working with daily data in the course of homogenization, even if it is more demanding compared to “simple” monthly averages. By employing daily data we have longer series (28 to 31 times, depending on number of days in a particular month) and we can then also better detect shifts near the end of the series (not resolvable for monthly averages with breaks of less then five years to the end of series).

Fig. 7 gives count of inhomogeneities detected for daily and monthly series by the Alexandersson test with reference series created by means of correlations.

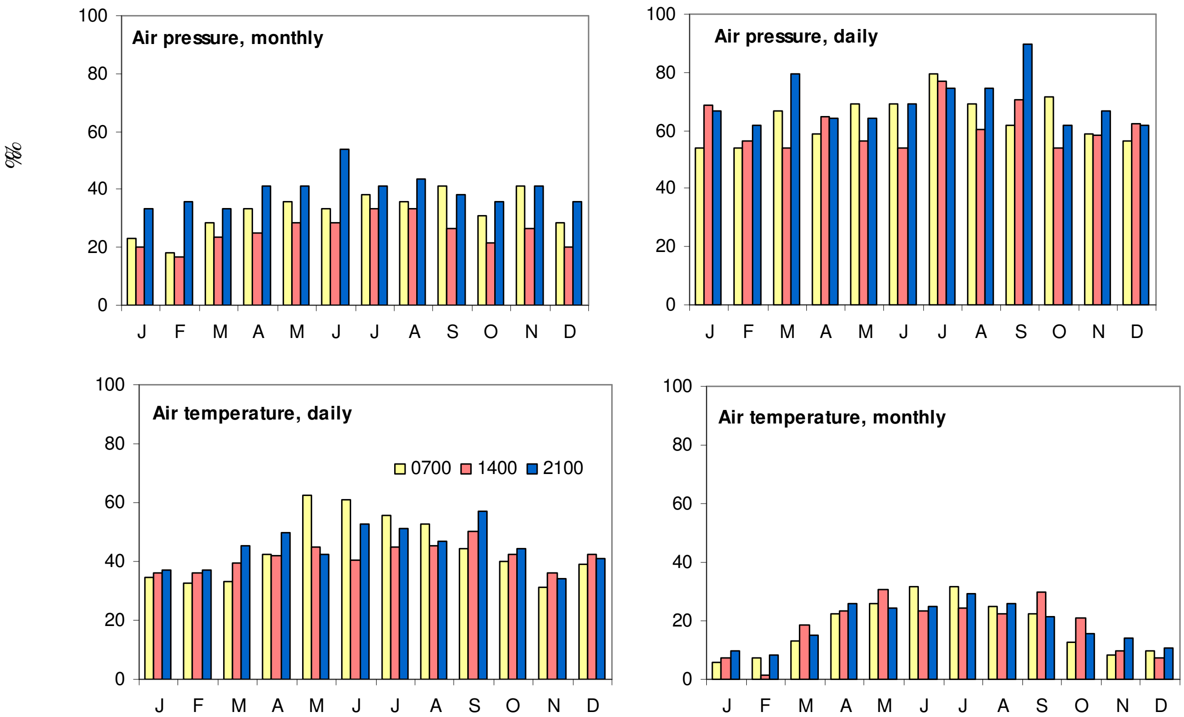


Fig. 7. Percentage of inhomogenities in air pressure and temperature series for daily data and monthly averages, detected by the Alexandersson test, related to the total number of series used

The annual course of numbers of inhomogeneities is evident from the figure, as are the differences between air pressure and temperature readings, as well as observation hours. The large difference between monthly- and daily-based detections is, among other things, due to the fact that in the Alexandersson test the series is divided into sections in the position of each detected break. Since the series contain more members, we are able to detect relatively more inhomogeneities (mainly in the shorter sections). In this sense, the numbers between daily and monthly series are not comparable. But the aim was to show that during homogenization we should try to use information that is as dense as possible, using daily data, individual observation hours, etc.

The advantages of using daily data mentioned above are apparent from the example in Fig. 8. In the event of missing values and breaks near the ends of series it is more difficult to detect inhomogeneities in the series if one works with only monthly data.

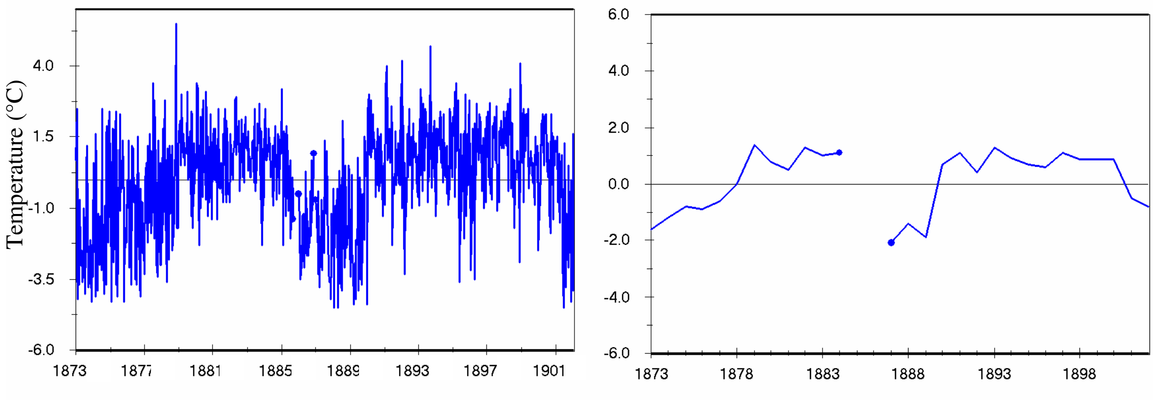


Fig. 8. Differences between tested and reference series for daily (left) and monthly (right) data for Brno, air temperature at 0700 (0600) hours, July 1873–1902

4.1 Homogenized pressure and temperature series of Brno

The creation of homogenized air pressure and temperature series for Brno covering the period 1848–2005 consists of several steps. First, the individual series for the different Brno stations (Brno stations before 1890 & Brno-Pisárky, Brno-Května, Brno-Tuřany – see Chapter 2) were homogenized according to the methodology described in Chapter 4. In the second step, a common compiled Brno series was developed by adjusting the individual parts. Starting from the recent observing station at Brno-Tuřany (reference station, 1958–2005), Brno-Květná data were adjusted to its measurements to obtain a series for the period 1923–2005. In the next step, the Brno-Pisárky station was adjusted to the combined Květná-Tuřany reference series to obtain a series for the period 1890–2005 (Fig. 9). This approach was applied separately for each observation time (0700, 1400 and 2100 hours).

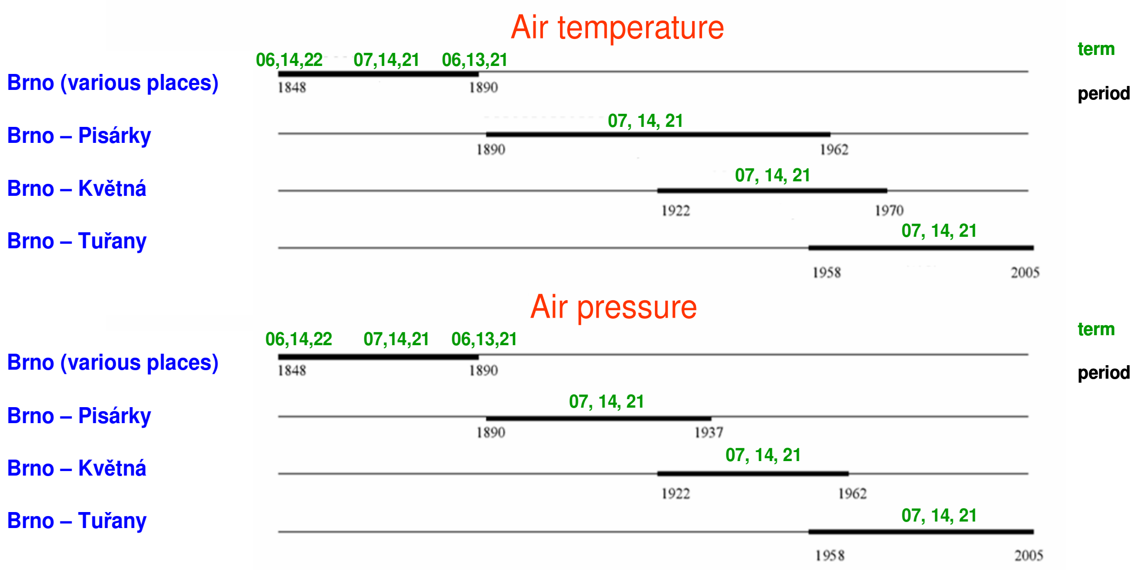


Fig. 9. Scheme of creation for the series compiled by combining measurements from several locations in Brno

The principle of combination used for the individual Brno stations is identical to that employed for adjustments of inhomogeneities applied to data in the course of homogenization. Two approaches may be selected: one that uses monthly averages or one that works directly from daily data. The only difference is that final offsets are not computed by comparing periods before and after the change; in this case we use the whole period in common (shortened to 20 years if it is longer). The overlap periods vary from 5 years (air pressure) or 13 years (air temperature) in the first round to 15 years (air pressure) or 20 years (air temperature) in the second round.

Fig. 10 gives an example of when final adjustment is obtained either from monthly averages or through direct use of daily data. It seems appropriate to calculate adjustments from daily values using a low-pass filter for 60 days, or, leading to the same results, using a low-pass filter for two months and subsequently distributing the smoothed monthly adjustments into daily values.

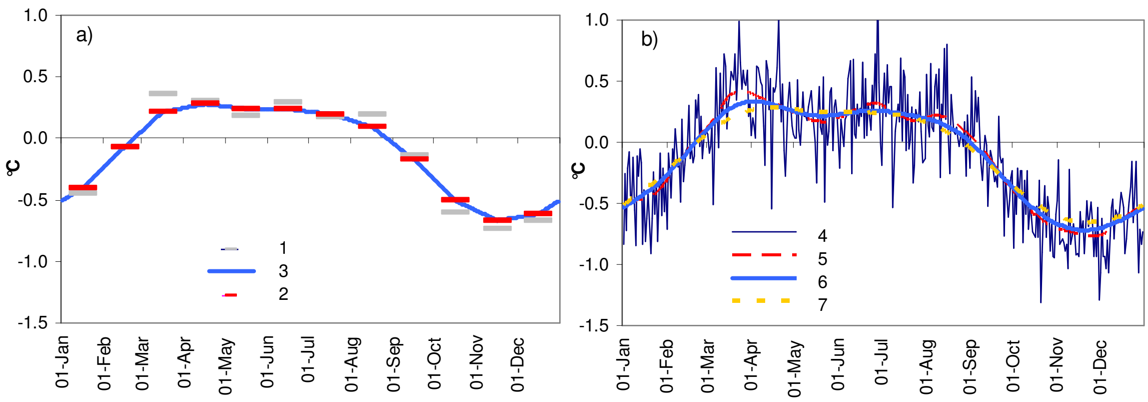


Fig. 10. Annual variations of adjustments applied to air temperature series at 1400 hours for Brno-Květná to the reference station Brno–Tuřany: a) monthly-based approach (1 – raw adjustments, 2 – smoothed adjustments, 3 – smoothed adjustments distributed into individual days), b) daily-based approach (4 – individual calendar day adjustments, 5 – daily adjustments smoothed by low-pass filter for 30 days, 6 – for 60 days, 7 – for 90 days)

The values measured at different observation hours exhibited quite different annual variations of adjustment, making it useful to work with them directly, and not just with calculated daily averages. For example, depending on the formula used for the calculation of daily averages, real inhomogeneities may be masked there.

A fully compiled series for the period 1848–2005 was again tested for homogeneity as a whole. Finally, homogenous Brno pressure and temperature series for 1848–2005 at 0700, 1400 and 2100 hours were obtained, from which corresponding daily and monthly averages were calculated. Fig. 11 shows fluctuations in annual averages both for the compiled homogeneous Brno series and the original series from the various places, in which it was derived.

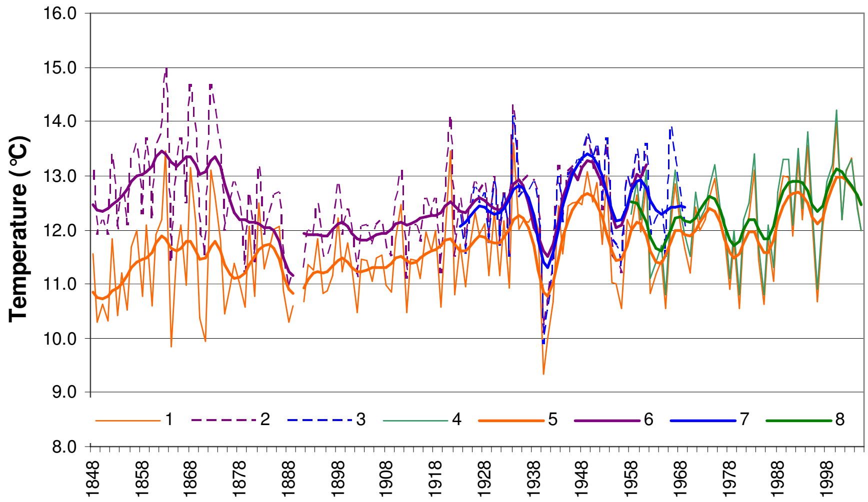


Fig. 11. Fluctuations of annual averages of air temperature series at 1400 hours (1 – compiled Brno series, 2 – Brno–various places and Brno-Pisárky, 3 – Brno-Květná, 4 – original Brno–Tuřany, 5, 6, 7, 8 –series smoothed by Gaussian low pass filter for 10 years)

5. CONCLUSIONS

This work was carried out in quest of a proper methodology for daily data homogenization and made an attempt to apply it subsequently to daily air pressure and temperature series for Brno in the period 1848–2005. Different methods for the homogenization of daily values were sought, and finally applied to find possible inhomogeneities and to obtain adjusted, homogeneous series. Although further investigation in this matter is required, progress so far may be summarized as:

(i) Two basic approaches, based on the homogenization of monthly series and projection of estimated monthly adjustments into a smoothed annual course of daily adjustments, or homogenization of daily values of individual months, estimating proper adjustments for each calendar day with smoothing adjustments, can be used.

(ii) The same final adjustments may be obtained from either monthly averages or through direct use of daily data. For the daily-values-based approach, it seems reasonable to smooth them with a low-pass filter for 60 days. The same results may be derived using a low-pass filter for two months (weights approximately 1:2:1) and subsequently distributing the smoothed monthly adjustments into daily values.

(iii) The values of the correlation coefficients between the candidate and reference series for daily data (working with each month individually) are comparable with values gained from monthly averages, although daily data are better in some months, monthly data in others. For this reason, a combination of both approaches in (i) is useful.

(iv) It is profitable to analyze series of individual observation hours because inhomogeneities manifest in different ways within their series – this is the case for the number of inhomogeneities detected, the value of change, the correlations between reference and tested series (and thus detectability of inhomogeneities) and other characteristics. Series of daily averages can serve as complementary information in the course of homogeneity test evaluation. For inhomogeneity assessment, we recommend the use of as much information as possible.

(v) The data processing in this work has been done by means of LoadData software (application for downloading data from central database, e.g. Oracle), ProClimDB software for processing whole datasets (finding outliers, combining series, creating reference series, preparing data for homogeneity testing, etc.) and AnClim software for homogeneity testing (http://www.klimahom.com/software). Further development of the software, e.g. connection with R software, is to be assumed.

Acknowledgements

The authors would like to acknowledge the financial support of the Grant Agency of the Czech Republic for project no. 205/05/0858. Sincere thanks should be extended to Ms Zuzana Pilařová, Institute of Geography, Masaryk University, Brno for digitizing and checking a great part of data used for analysis. Mr. Tony Long, Auchrae, Scotland, undertook the English editing.

REFERENCES

Alexandersson, A. (1986): A homogeneity test applied to precipitation data. Journal of Climatology, 6, 661–675.

Alexandersson, A. (1995): Homogenity testing, multiple breaks and trends. In: Proc. 6th Int. Meeting on Stat. Climatology, Galway, Ireland, 439–441.

Alexandersson, H., Moberg, A. (1996): Homogenization of Swedish temperature data. Part I: A homogeneity test for linear trends. In: Moberg, A: Temperature Variations in Sweden Since the 18th Century. The Department of Physical Geography, Stockholm University. Dissertation Series No. 5, 98 pp.

Auer, I., Böhm, R., Schöner, W. (2001): Austrian long-term climate 1767–2000, Multiple instrumental climate time series form Central Europe. Österreichische Beiträge zu Meteorologie und Geophysik, Heft 25, Wien, 155 pp.

Brandsma, T. (2000): Weather-type dependent homogenization of daily Zwanenburg/De Bilt temperature series. http://www.met.hu/omsz.php?almenu_id=omsz&pid=seminars&pri=6&mpx=1&sm0=0&tfi=brandsma

Brázdil, R. (1979): Historie měření srážek v Brně (History of precipitation measurements in Brno). Scripta Fac. Sci. Nat. UJEP Brunensis, 9, Geographia 2, 55–74.

Brázdil, R., Valášek, H., Macková, J. (2005): Meteorologická pozorování v Brně v první polovině 19. století. Historie počasí a hydrometeorologických extrémů (Meteorological Observations in Brno in the First Half of the Nineteenth Century. History of Weather and Hydrometeorological Extremes). Archiv města Brna, Brno, 452 pp.

Easterling, D. R., Peterson, T. C. (1995): A new method for detecting undocumented discontinuities in climatological time series. International Journal of Climatology, 15, 369–377.

Mekis, É., Vincent, L. (2004): New developments in the homogenization of precipitation and temperature series in Canada. In: Fourth seminar for homogenization and quality control in climatological databases (Budapest, Hungary, 6-10 October 2003), WCDMP-No. 56. WMO, Geneva, 39–62.

Mitchell, J. M., ed. (1966): Climatic Change. WMO, T.N. 79, Geneva, 80 pp.

Peterson, T. C. (1998): Homogeneity adjustments of in situ atmospheric climate data: a review. International Journal of Climatology, 18, 1493–1517.

Petrovic, P. (2004): Detecting of inhomogeneities in time series using Real Precision Metod. In: Fourth seminar for homogenization and quality control in climatological

databases (Budapest, Hungary, 6-10 October 2003), WCDMP-No. 56. WMO, Geneva, 79–88.
Pettit, A. N. (1979): A non-parametric approach to the change-point problem. App. Statist., 28, 2, 126–135.
Štěpánek, P., 2005: Variabilita teploty vzduchu na území České republiky v období přístrojových měření (Air Temperature Fluctuations in the Czech Republic in the Period of Instrumental Measurements). Ph.D. Thesis. Institute of Geography, Masaryk University, Brno, 136 pp.
Štěpánek, P. (2006a): AnClim - software for time series analysis. Institute of Geography, Masaryk University, Brno, 1.47 MB. http://www.klimahom.com/software/AnClim.html
Štěpánek, P. (2006b): ProClimDB – software for processing climatological datasets. CHMI, regional office Brno. http://www.klimahom.com/software/ProcData.html
Wijngaard, J. B. , Klein Tank, A. M. G. , Können, G. P. (2003): Homogenity of 20th century European daily temperature and precipitation series. Int. J. Climatol., 23**,** 679–692