PERGAMON

Physics and Chemistry of the Earth 27 (2002) 1063–1071

PHYSICS and CHEMISTRY of the EARTH

www.elsevier.com/locate/pce

Long term climate deviations: an alternative approach and application on the Palmer drought severity index in Hungary

L. Makra [a,*], Sz. Horváth [a], R. Pongrácz [b], J. Mika [c]

[a] *Department of Climatology and Landscape Ecology, University of Szeged, P.O. Box 653, H-6701 Szeged, Hungary*
[b] *Department of Meteorology, Eötvös Loránd University, H-1117 Budapest, Pázmány Péter sétány 1/A, Hungary*
[c] *Hungarian Meteorological Service, Budapest, P.O. Box 39, H-1675 Budapest, Hungary*

Received 10 December 2001; received in revised form 30 March 2002; accepted 17 July 2002

Abstract

A new statistical test (Makra-test), applicable for long time series, is introduced to identify extended sub-periods, namely, "breaks", average of which is significantly higher or lower than the mean of the entire time series. In order to apply this test, normal distribution of the time series being examined is a sufficient condition. In another case, if the number of elements increases in the time series, its distribution is near normal and the test can be applied, as well (according to the Central Limit Theorem). The method is demonstrated on monthly Palmer drought severity index (PDSI) data sets, computed for five stations of East Hungary in 1901–1999. Due to strongly recursive (auto-correlative) nature of PDSI every second month of the warm season (April, June, August and October) is analysed and treated as independent samples. Normality of the time series, which is a sufficient condition of the Makra-test, is validated by Kolmogorov–Smirnov test and χ^2-test. Analysis of the PDSI time series indicates that separate treatment of the months is important not only to ensure the normality, but also to consider the existing slight seasonal differences in standard deviation and skewness of the index in the East-Hungarian region. The Makra-test delimits one or more (maximum 4) significant sub-periods of the PDSI in every station and month (not considering the sub-intervals within the significant breaks, although most of them are also significant). Since the PDSI is based on monthly temperature and precipitation data that exhibit considerable inhomogeneities (Szentimry, 1999), the test is applied both for the original and the homogeneous time series. Effect of the inhomogeneity on long term variations of PDSI is strong: about the half of the significant breaks in the original time series disappear or become totally different from the time series based on homogenised data. All negative (dry) breaks of each month and station, however, occurred in more recent decades of the 20th century, according to both homogeneous and original series.

1. Introduction

Climate performs year to year variability, fluctuation on time scales of several years, and change on longer time scales. In order to maintain the current human activities and predict their future impacts, these changes and fluctuations should be investigated and analysed precisely. It is well known that in some places local climate parameters do not exhibit clear tendencies of global warming, since the effects of circulation are likely non-linear (Makra, 1998). Hence, besides careful application of trend analysis to identify long term monotonous tendencies, whether or not they correlate with the global ones, time series analysis should also incorporate other types of questions (statistical hypotheses).

The aim of the present study is to introduce a test on the existence of extended sub-periods, average of which is significantly higher or lower than the mean of entire time series. The method, hereafter called Makra-test, will be demonstrated on a few century-long time series characterising moisture anomaly (drought or wetness) conditions in Hungary.

Drought is a long period of meteorological anomaly characterised by extreme lack of precipitation and it is a normal issue of climate in Hungary. A drought event is different from other natural hazards since it is a slow-onset, insidious hazard that is often well-established before it is recognised as a threat, taking months or years to develop. In general, drying tendency is apparent in most of the stations in Hungary during the last 100 years (Pongrácz et al., 2000). Furthermore, frequency of

* Corresponding author. Tel.: +36-62-545000-3174; fax: +36-62-544-158.

E-mail address: makra@geo.u-szeged.hu (L. Makra).

1474-7065/02/$ - see front matter © 2002 Elsevier Science Ltd. All rights reserved.
PII: S1474-7065(02)00142-0

drought events tends to increase towards the end of the 20th century, especially considering moderate and severe droughts (Szinell et al., 1998).

It is common to distinguish four different types of droughts: meteorological, climatological, agricultural, and hydrological (water-supply) droughts (Wilhite and Glantz, 1985). Several drought indices have been defined (e.g., Faragó et al., 1989) and they have become common tools to measure the intensity and spatial extent of droughts. These drought indices comprise ratio or difference between precipitation and potential evapotranspiration (PE) for periods as long as they can generally characterise drought events well.

According to their characteristics drought indices can be categorised as follows: precipitation indices, water budget indices, soil moisture indices, hydrological and various aridity indices. In order to determine onset and strength of meteorological droughts, monthly values of the Palmer drought severity index (PDSI) are considered in this study. Compared to other traditional drought indices, PDSI can demonstrate several advantages: it is able to simulate moisture content of the soil month by month, and it is suitable to compare the severity of drought events at regions having totally different climate and seasons. Evaluation of summer PDSI times series for Europe can be found in Briffa et al. (1994).

The paper is structured as follows: Section 2, describes statistical methods of the Makra-test and the correction for inhomogeneity. Section 3, specifies basic concepts and main steps of the PDSI computation and briefly introduces the study region. Results of the Makra-test on the PDSI time series preceded by validation of its (sufficient) precondition, namely, the normality of the distribution are introduced by Section 4, also supplied by the effect of preliminary homogenisation of the input data. Finally, our experiences are briefly concluded in Section 5.

2. The statistical methods

2.1. The Makra-test

The applied statistical test, developed by Makra, is a method that is considered being a new interpretation of the two-sample test as it is discussed at the end of this section.

The basic question of this test is whether or not a significant difference can be found between the averages of an arbitrary sub-sample of a given time series and the whole sample (Makra et al., 2000a).

Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots, \xi_N$ represent independent random variables of normal distribution, with mean m. Let $E(\xi)$ be the expected value of ξ, while $D(\xi)$ the standard deviation of it.

Suppose that standard deviations of $\xi_i - s$ are identical and equal to σ. Now, choose an optional sub-sample of n elements from the given whole time series of N elements ($n < N$). Let

$$\overline{m} = \frac{\xi_1 + \cdots + \xi_n}{n} \quad \text{and} \quad \overline{M} = \frac{\xi_1 + \cdots + \xi_N}{N}, \tag{1}$$

where $n < N$.

Next, compose their difference: $\overline{M} - \overline{m}$. Then, after elementary steps, we get

$$\begin{aligned} \overline{M} - \overline{m} &= (\xi_1 + \cdots + \xi_n)\left(\frac{1}{N} - \frac{1}{n}\right) + \frac{\xi_{n+1} + \cdots + \xi_N}{N} \\ &= -\frac{N-n}{N}\frac{\xi_1 + \cdots + \xi_n}{n} + \frac{N-n}{N} \\ &\quad \times \frac{\xi_{n+1} + \cdots + \xi_N}{N-n}. \end{aligned} \tag{2}$$

Note, that this elementary transformation allows us to express the above difference as a sum of two statistically independent random variables:

$$\overline{M} - \overline{m} = \overline{Q} + \overline{R}, \tag{3}$$

where

$$\overline{Q} = -\frac{N-n}{N}\frac{\xi_1 + \cdots + \xi_n}{n}, \tag{4}$$

$$\overline{R} = \frac{N-n}{N}\frac{\xi_{n+1} + \cdots + \xi_N}{N-n}. \tag{5}$$

Our aim is to examine if $\overline{m}$ does not differ significantly from $\overline{M}$. In order to complete this, a 0-Hypothesis ($H_0^{(1)}$) can be set up, according to which: $E(\overline{M}) = E(\overline{m})$.

It is obvious that $E(\overline{M} - \overline{m}) = 0$.

Furthermore:

$$\begin{aligned} D^2(\overline{M} - \overline{m}) &= D^2(\overline{Q} + \overline{R}) \\ &= \left(-\frac{N-n}{N}\right)^2 \frac{1}{n^2} n\sigma^2 + \left(\frac{N-n}{N}\right)^2 \\ &\quad \times \frac{1}{(N-n)^2}(N-n)\sigma^2 = \frac{N-n}{Nn}\sigma^2. \end{aligned} \tag{6}$$

Hence, the hereby introduced random variable:

$$\mathrm{PS}^{(1)} = \frac{\overline{M} - \overline{m}}{\sqrt{\frac{N-n}{Nn}}\sigma} \tag{7}$$

can be characterised by standard normal distribution, $N(0;1)$.

This implies that having fixed the sample mean $\overline{M}$ and the standard deviation σ, test of the above 0-Hypothesis, concerning a given sub-sample mean $\overline{m}$, leads to the following comparison of $\mathrm{PS}^{(1)}$ and x_p,

$$P\left(\left|\frac{\overline{M} - \overline{m}}{\sqrt{\frac{N-n}{Nn}}\sigma}\right| > x_p\right) = p. \tag{8}$$

In Eq. (7) x_p is taken with p probability from the distribution function of the standard normal distribution, corresponding to a selected $0 < p \ll 1$ probability threshold.

If the absolute value of $\mathrm{PS}^{(1)}$ (Eq. (7)) is higher than x_p then $\overline{M}$ and $\overline{m}$ differ significantly. The 0-Hypothesis, according to which there is no significant difference between $\overline{M}$ and $\overline{m}$, can be rejected with the significance level p. (The significance-tests of Section 4.2. are carried out at $p = 0.01$ significance level.)

This is a sufficient condition, ensuring normal distribution of $\mathrm{PS}^{(1)}$ in Eq. (7), but it can be softened in case of very large N and n, considering the Central Limit Theorem. Stationarity and independence of the original distribution, however, are unavoidably necessary conditions of the Makra-test.

The Makra-test performs Eq. (8) for all possible subsamples with $n = 3, 4, \ldots, N-1$ elements of duration, starting from the $1\mathrm{st}, 2\mathrm{nd}, \ldots, (N-n)$th element of the time series. For example, in case of 99 data (years) this means 4752 repeated comparison of the sub-sample average and the overall mean. Detection of significant deviations also includes information on their duration, onset and end (Makra, 1999; Makra and Horváth, 1999; Makra et al., 2000b; Tar et al., 2001).

When performing Makra-test, the received significant breaks can be distinct or not. If they are not distinct, only one break is considered to be significant; namely, for which the value of the test statistics is the maximum. However, many test statistics among all, contrary to true 0-Hypothesis, can be much bigger, than the critical x_p value, belonging to the given significance level. Namely, distribution of maxima does not agree with normal distribution of the individual $\mathrm{PS}^{(1)}$ test statistics. In this way, correct detection of breaks needs further consideration (Szentimrey et al., 1992).

External similarity between the two-sample test and the main step of the Makra-test described by Eq. (8) is obvious. In the following it is presented that these two methods are identical, indeed.

When applying the Makra-test, our aim is to examine the following 0-Hypothesis:

$$H_0^{(1)} : E(\overline{m}) = E(\overline{M}). \tag{9}$$

The above 0-Hypothesis is equivalent with the following one:

$$H_0^{(2)} : E(\overline{m}_1) = E(\overline{m}_2). \tag{10}$$

We know that the $H_0^{(2)}$ 0-Hypothesis, in case of known standard deviation, can be checked by the $\mathrm{PS}^{(2)}$ probe statistics, where $\mathrm{PS}^{(2)}$ is as follows:

$$\mathrm{PS}^{(2)} = \frac{\overline{m}_2 - \overline{m}_1}{\sigma}\sqrt{\frac{n_1 n_2}{n_1 + n_2}}. \tag{11}$$

It can be proved that $\mathrm{PS}^{(2)}$ test statistics equals to $\mathrm{PS}^{(1)}$, introduced by Eq. (7), which is exactly the function of random variables of $H_0^{(1)}$ 0-Hypothesis:

$$\begin{aligned}\mathrm{PS}^{(1)} &= \frac{\overline{M} - \overline{m}}{\sqrt{\frac{N-n}{Nn}}\sigma} = \frac{\frac{n\overline{m}+(N-n)\overline{m}_2}{N} - \overline{m}}{\sqrt{\frac{N-n}{Nn}}\sigma} = \frac{\frac{N-n}{N}(\overline{m}_2 - \overline{m})}{\sqrt{\frac{N-n}{Nn}}\sigma} \\ &= \frac{\overline{m}_2 - \overline{m}}{\sqrt{\frac{N}{(N-n)n}}\sigma} = \frac{\overline{m}_2 - \overline{m}_1}{\sqrt{\frac{n_1+n_2}{n_1 n_2}}\sigma} = \frac{\overline{m}_2 - \overline{m}_1}{\sigma}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \mathrm{PS}^{(2)},\end{aligned} \tag{12}$$

where

$$\overline{m}_2 = \frac{\xi_{n+1} + \cdots + \xi_N}{N-n}, \quad \overline{m}_1 = \overline{m}, \quad n_1 = n, \quad n_2 = N - n.$$

(Note that in the elementary arithmetic steps of Eq. (2) similar operations were applied in inverse direction.)

Consequently, let $\mathrm{PS}^{(1)}$ be the probe statistics of the $H_0^{(1)}$ 0-Hypothesis. Characteristics of the statistical test (the Makra-test), received in this way, are identical with those of the known two-sample test, since

$$H_0^{(1)} \Longleftrightarrow H_0^{(2)} \quad \text{and} \quad \mathrm{PS}^{(1)} = \mathrm{PS}^{(2)}. \tag{13}$$

Therefore, the known two-sample $H_0^{(2)}$, $\mathrm{PS}^{(2)}$ test can be formulated in $H_0^{(1)}$, $\mathrm{PS}^{(1)}$ form, the Makra-test, as well. Advantage of the latter one is that $H_0^{(1)}$ 0-Hypothesis refers definitely to $\overline{m}$ and $\overline{M}$, and these variables are also found in $\mathrm{PS}^{(1)}$ (Eq. (7)).

Hence, the main step of the Makra-test, namely, comparison of the average of a selected sub-period and the mean of the entire time series, is equivalent to the two-sample test applied to the given sub-period and the rest of the whole sample.

In other words, application of Eq. (8) is possible to the non-independent samples, i.e., the given data set and its fixed sub-sample. The reasoning of such use is definitely the transformation into two non-overlapping subsamples.

The practical advantage of $\mathrm{PS}^{(1)}$ (Eq. (7)), used by the Makra-test, is that the procedure changes only three numbers ($\overline{m}$ and n—twice) in each step of the repeated application, while in case of $\mathrm{PS}^{(2)}$ (Eq. (11)) six numbers have to be modified.

2.2. Homogenisation

Examination of homogeneity and correction of temperature and precipitation data sets were performed using the MASH (multiple analysis for homogenisation) method, developed in the Hungarian Meteorological Service (Szentimrey, 1995, 1999). It is a relative homogeneity test procedure that does not assume homogeneous reference series. Possible break points and shifts can be detected and adjusted through mutual comparisons of time series within the same climatic area. The candidate series is chosen from the available time series and the remaining time series are considered as reference series. The role of time series changes step by step in the course of the procedure. Depending on the climatic elements, additive or multiplicative models are applied. The latter case can be transformed into the first one using logarithmic transformation.

Several difference series are constructed from the candidate and weighted reference series. The optimal weighting is determined by minimising the variance of the difference series, in order to increase the efficiency of the statistical tests. Providing that the candidate series is the only common series of all the difference series, break points detected in all the difference series can be attributed to the candidate series.

A new multiple break points detection procedure has been developed which takes the problem of significance and efficiency into account. The significance and the efficiency are formulated according to the conventional statistics related to type one and type two errors, respectively. This test obtains not only estimated break points and shift values, but the corresponding confidence intervals, as well. The time series can be adjusted by using the point and interval estimates.

Since a MASH program system—mentioned also in the recent collaborative survey on climate data homogeneity (Peterson et al., 1998)—has been developed for PC the application of this method was relatively easy.

3. The Palmer drought severity index and the region

3.1. PDSI

In the present study, drought is considered as a meteorological anomaly characterised by a prolonged and abnormal moisture deficiency (Palmer, 1965; Delezios et al., 1991). In order to determine onset and severity of meteorological droughts, one of the most commonly used drought index PDSI is evaluated in this paper. Monthly PDSI values have been calculated for five stations of the selected region (see Section 3.2) for the period 1901–1999.

PDSI indicates the severity of a wet or dry spell—the greater the absolute value the more severe the dry or the wet period. In general, monthly PDSI time series range between -9 and $+9$, specifically, severe and extreme conditions are characterised by absolute values greater than 4 and 6, respectively. These thresholds may vary among the geographic regions of the world, whereas the original attribution considered ± 4 to be the extremity threshold (Palmer, 1965). Furthermore, drought events occur in the case of negative PDSI values while positive values imply wet conditions.

Detailed procedure of computing the PDSI is described elsewhere (e.g., Palmer, 1965; Alley, 1984; Karl, 1986). PDSI is referred as an index of meteorological drought, the computation procedure considers monthly precipitation, evapotranspiration and soil moisture conditions, and these meteorological variables determine hydrological and agricultural drought.

In general, several methods can be used to calculate the PE, a key variable of the water balance and, also, of the PDSI computation procedure. Palmer (1965) applied the Thornthwaite formula (Thornthwaite, 1948), while later the Blaney–Criddle method provided better estimations (Alley, 1984).

In our analysis we calculated PDSI in both ways, whereas corn (maize) has been selected as the reference crop in the Blaney–Criddle method. Our results suggest that PDSI values are slightly smaller using the Thornthwaite formula than the Blaney–Criddle method. Furthermore, strong correlation (around 0.9) has been found at all stations between PDSI time series using these two approaches. Henceforward, we will focus on PDSI data sets based on the Blaney–Criddle method.

Basic concepts and steps of computation are as follows.

Step 1: *Hydrological accounting*. Computation of PDSI begins with a climatic water balance using series of monthly precipitation and temperature records. An empirical procedure is used to account for soil moisture storage by dividing the soil into two arbitrary layers. The upper layer is assumed to contain 25 mm of available moisture at field capacity. The loss from the underlying layer depends on the initial moisture content, as well as on the computed PE and the available water capacity (AWC) of the soil system. In the present calculations of PDSI, AWC values of 180–200 mm are used for the different soil types (Table 1). Runoff is assumed to occur, if and only if, both layers reach their combined moisture capacity, AWC. In addition to PE, three more potential terms are used and defined as follows: *potential re-charge* is the amount of moisture required to bring the soil to its water holding capacity. *Potential loss* is the amount of moisture that could be lost from the soil by evapotranspiration during a zero precipitation period. *Potential run-off* is defined as the difference between precipitation and potential recharge.

Step 2: *Climatic coefficients*. This is accomplished by simulating the water balance for the period of available weather records. Monthly coefficients are computed as proportions between climatic averages of actual vs. potential values of evaporation, recharge, runoff and loss, respectively.

Step 3: *CAFEC values*. The derived coefficients are used to determine the amount of precipitation (Π) required for the climatically appropriate for existing conditions (CAFEC), i.e., "normal" weather during each individual month.

Step 4: *Moisture anomaly index*. Difference between the actual and CAFEC precipitation is an indicator of water deficiency or surplus in that month and station, expressed as $D = P - \Pi$. These departures are converted into indices of moisture anomaly as $Z = K(j)D$, where

Table 1
Geographical co-ordinates, soil types and AWC values at the examined stations

Geographical co-ordinates			Settlement	Soil types	AWC (mm)
Latitude	Longitude	Height (m)			
47°33′ N	21°37′ E	123	Debrecen	Humous sandy soils	180
46°54′ N	19°45′ E	113	Kecskemét	Humous sandy soils	180
48°06′ N	20°38′ E	303	Miskolc	Meadow alluvial soils and alluvial meadow soils	200
47°58′ N	21°43′ E	107	Nyíregyháza	Brown forest soil (sandy brown forest soil with thin interstratified layers of colloid and sesquioxide accumulation)	180
46°15′ N	20°09′ E	79	Szeged	Meadow soils, solonetzic meadow soils, meadow chernozem (the term "meadow" is related to hydromorphic character)	190

$K(j)$ is a weighting factor for the month j, which also accounts for spatial variability of the departures (D).

Step 5: *Drought severity*. In the final step the Z-index time series are analysed to develop criteria for the beginning and ending of drought periods and an empirical formula for determining drought severity, such as

$$X_j = 0.897X_{j-1} + Z_j/3, \tag{14}$$

where Z_j is the moisture anomaly index and X_j is the value of PDSI for the jth month.

Eq. (14) indicates that PDSI of a given month strongly depends on the situation of the previous months and on the moisture anomaly of the actual month. As it is being presented in Section 4.1, the first term of (14) causes strong auto-correlation of PDSI in the region.

3.2. The selected region

Monthly PDSI values are analysed for century-long time series of five stations, available in the Hungarian catchment area of the Tisza River (Fig. 1). Energetic and hydrological processes of this region have already been analysed and simulated in separate studies (Mika et al., 1998, 2001). This large landscape has always been characterised by high proportionality of managed vegetation. Recently, 74% of the total geographical area (35,700 km^2) is cultivated. Corn, plant constant of which is applied in the above version of PDSI computation, is the most common agricultural plant, together with wheat, which, in turn, is less vulnerable to water balance anomalies, due to its early harvest in summer.

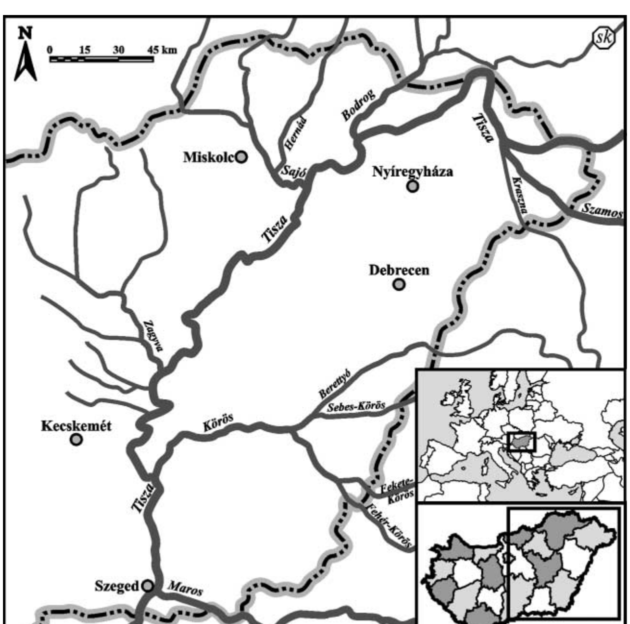

Fig. 1. Location of the five investigated stations in East Hungary.

Geographic coordinates, characteristic soil-types and AWC values of the five investigated stations are included in Table 1.

4. Results

4.1. Seasonal variations and normality of PDSI

In principle, PDSI is a non-seasonal characteristics of water availability. However, since the empirical constants of the calculation procedure (Section 3.1) are determined for other regions (mainly for continental USA), it is worth checking if statistical parameters of PDSI are really independent from month. Since the following analysis focuses on the growing season (of maize, according to the Blaney–Criddle formula of evapotranspiration). Table 2 summarises basic statistical parameters (mean, standard deviation and skewness) for four month (April, June, August and October) and the five stations (Debrecen, Miskolc, Nyíregyháza, Kecskemét and Szeged) between 1901–1999.

According to the two-sample test, there are no significant differences between the means of the data sets for consecutive months. (They differ from zero slightly, due to the fact that the reference period of computations was 1961–90, not the entire 99 years.) Standard deviations of time series of the selected five stations are similar according to the F-test. Further similarity can be detected in minimum standard deviation occurring in June, and maximum in August—it is quite unlikely that this is only by chance. So, we cannot speak about strict seasonal independence of PDSI values in East Hungary.

Table 2
Basic statistics of PDSI and results of normality tests, 1901–1999

PDSI	Debrecen				Miskolc				Nyíregyháza				Kecskemét				Szeged			
	Apr	Jun	Aug	Apr	Jun	Aug	Oct	Apr	Jun	Aug	Oct	Oct	Apr	Jun	Aug	Oct	Apr	Jun	Aug	Oct
Mean	−0.2	−0.1	−0.3	−0.2	−0.3	−0.4	−0.2	−0.4	−0.3	−0.2	−0.2	−0.1	−0.3	−0.4	−0.4	−0.3	−0.2	−0.3	−0.3	−0.1
Standard deviation	2.2	1.9	2.5	2.1	1.9	2.3	2.1	2.4	2.0	2.4	2.3	2.4	2.3	2.1	2.4	2.5	2.2	2.0	2.4	2.4
Skewness	0.0	0.1	0.4	−0.1	0.3	0.4	0.6	0.4	0.4	0.4	0.5	0.3	0.0	0.2	0.4	0.3	−0.1	0.3	0.3	0.4
z-value (K–S test)	0.6	0.9	1.1	0.4	0.8	0.9	0.7	1.0	0.9	1.0	0.7	0.8	0.5	0.9	0.7	0.9	0.5	1.2	1.0	1.0
k-value (χ^2-test)	7.8	17.2	22.5	1.7	8.39	7.2	3.9	7.7	7.8	7.7	5.1	6.1	1.1	8.2	4.8	9.3	4.0	12.1	9.8	14.5

(A strict common test of the seasonal differences would need special steps because of auto-correlation of PDSI, see below.)

Parameters of standard skewness are mainly positive (with some exception in April), indicating that generally more severe positive (very wet) anomalies may occur, compensated by more frequent, but less severe negative (dry) periods in the distribution.

Normality of the distribution is a sufficient condition of the Makra-test, hence, it was also investigated for the monthly time series of PDSI applying both the Kolmogorov–Smirnov (K–S) test and the χ^2-test. Normality is fulfilled in all $5(\text{stations}) \times 4(\text{months}) = 20$ cases. Namely, the z-values of the K–S test are lower than the 1 % probability threshold (1.63) and k-values the χ^2-test are also far less than the same threshold value (18.48 for $FG = 7$). Note, that without separation of monthly samples but kept all the 12 months together, Mika et al. (1994) found irregular (multi-modal), definitely non-Gaussian distribution of PDSI in Hungary and three other European countries.

Another argument on separation into monthly sub-samples is the strong auto-correlation of the time series, being a consequence of the recursive definition (Eq. (14) in Section 3.1) indicated in Table 3. The two- and even the four-month lag auto-correlation of PDSI is significant at the 1% probability level in all cases. These auto-correlation values do also somewhat depend on the season. Namely, values of two- or four-month transitions that span over the May–June precipitation maximum are lower than the others.

Table 3
Two- and four-month lag auto-correlation of PDSI

r, month	Debrecen	Kecskemét	Miskolc	Nyíregyháza	Szeged
Apr–Jun	0.597	0.586	0.692	0.602	0.517
Jun–Aug	0.871	0.759	0.697	0.787	0.856
Aug–Oct	0.711	0.643	0.863	0.729	0.704
Apr–Aug	0.463	0.389	0.416	0.448	0.392
Jun–Oct	0.594	0.454	0.589	0.537	0.644

4.2. Periods of significant breaks

The statistical test, described in Section 2, is applied to separate significantly drier or wetter sub-periods compared to the entire average. (If the additional assumption of the test is fulfilled, namely, no change in standard deviation occurs, then these periods also exhibit higher or lower inclination to dry or wet extremes.) The five stations, the four calendar months and the 99 year period are the same as in the previous section. PDSI time series with both original (non-homogenised) and homogenised series of temperature and precipitation are analysed. (This is what we exactly mean under non-homogenised, or homogenised PDSI.)

The idea of the Makra-test and the parallel results from evaluation of the two PDSI data sets are demonstrated for June at station Szeged, i.e., a case that is rich in breaks being sometimes very similar, and other times totally different in original vs. in homogenised time series (Fig. 2) e.g., the long positive anomaly period 1909–1944 of the original time series appears to be similar to the 1909–1941 period after homogenisation, with slightly modified average value. The 1946–1998 period performs significant dry breaks in both time series. On the other hand, break periods of 1938–42 and 1942–68 in the original time series are not present in the homogenised one, more precisely, they partly overlap with much longer breaks of the same sign appeared after homogenisation.

Tables 4 and 5 summarise periods and signs of the breaks in the original and the homogenised data sets (representing 19 figures similar to Fig. 2). Significant breaks are the most similar in case of non-homogenised and homogenised PDSI time series in August, while the largest differences occur in April.

The effect of homogenisation can be presented by the following comparisons. In the original data sets dry breaks occur more often concerning the number of breaks (22 vs. 14; i.e., 1.1 vs. 0.7 in average of the five stations and four months). Duration of the breaks demonstrates the same differences: 927 vs. 337 years, i.e., 46 vs. 17 years, compared to 20 cases; or 42 vs. 24 considering the above number of breaks. This asymmetry is in coincidence with the positive skewness of the

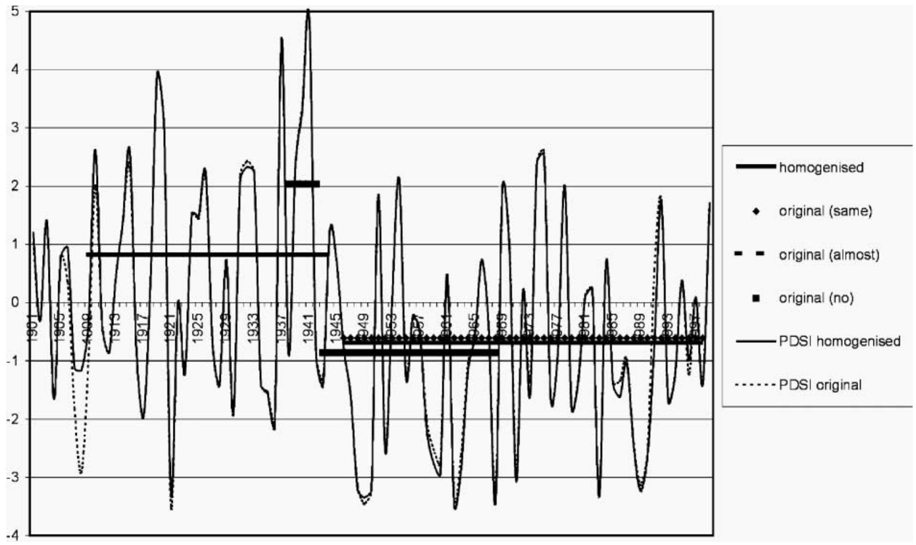

Fig. 2. Sub-periods with significantly different averages of PDSI from the mean of the entire data series, i.e., the breaks, Szeged, June, 1901–1999. Legends: same—significant sub-periods (breaks) that are **fully** identical in the original and homogenised series. Almost—significant sub-periods (breaks) that are *nearly* identical in the original and homogenised series. No—significant sub-periods (breaks) that are *not at all* identical in the original and homogenised series.

Table 4
Sub-periods with significantly different averages of PDSI from the mean of the entire original (non-homogenised) data set: the breaks, 1901–1999

Station	PDSI								
	April		June		August		October		All (+/−)
	Period	Year[a]	Period	Year[a]	Period	Year[a]	Period	Year[a]	
Debrecen	1912–1916	5 (+)	–	–	–	–	*1912–1944*	*33* (+)	***38*** (+)
	1925–1997	**73** (−)	**1958–1964**	**7** (−)	**1958–1964**	**7** (−)	*1945–1995*	*51* (−)	**138** (−)
Miskolc	–	–	–	–	–	–	1913–1915	3 (+)	**3** (+)
	1980–1996	17 (−)	1983–1997	15 (−)	1982–1994	13 (−)	1917–1997	80 (−)	***125*** (−)
Nyíregyháza	1901–1967	67 (+)	1913–1915	3 (+)	**1913–1915**	**3** (+)	1910–1952	43 (+)	*116* (+)
	1918–1996	**79** (−)	1954–1996	43 (−)	**1986–1994**	**9** (−)	1953–1995	43 (−)	
	–	–	–	–	–	–	*1921–1985*	*65* (−)	*239* (−)
Kecskemét	1906–1942	37 (+)	**1905–1915**	**11** (+)	–	–	**1905–1941**	**37** (+)	*85* (+)
	1925–1999	75 (−)	*1916–1994*	*79* (−)	**1982–1994**	**13** (−)	*1941–1997*	*57* (−)	*224* (−)
Szeged	*1910–1946*	*37* (+)	1938–1942	5 (+)	–	–	1925–1941	17 (+)	
	–	–	*1909–1944*	*36* (+)	–	–	–	–	*95* (+)
	1946–1998	53 (−)	1942–1968	27 (−)	1946–1964	19 (−)	**1942–1990**	**49** (−)	
	–	–	**1946–1998**	**53** (−)	–	–	–	–	*201* (−)
Positive	4 Periods	146 (+)	4 Periods	55 (+)	1 Period	3 (+)	5 Periods	133 (+)	(14) ***337***
Negative	5 Periods	297 (−)	6 Periods	224 (−)	5 Periods	61 (−)	6 Periods	345 (−)	(22) ***927***

Bold values denote significant sub-periods (breaks) that are **fully** identical in the original and homogenised series.
Italic values denote significant sub-periods (breaks) that are *nearly* identical in the original and homogenised series.
[a] Note: (−) significant dry period; (+) significant wet period.

distribution (Section 4.1), however, skewness coefficient treats each data individually and does not account for temporal coincidence.

In the homogenised data sets this asymmetry is decreasing: the number of dry breaks is 19 compared to the 17 wet breaks (0.95 vs. 0.85), i.e., 3 breaks changed their sign. Common duration of dry breaks increases just a little (935 years, i.e., 47 or 49 years, depending on the basis), but duration of wet breaks became 537 years (27 or 32 years, in average).

So, our results suggest that homogenisation yields longer breaks with no change in the overall numbers in the vegetation period. Furthermore, considerable shift can be detected in decreasing the asymmetry of the dominance of dry breaks.

Focusing on the homogenised time series (Table 5), the significant wet periods are detected in the first half of the century (between 1901 and the 1940s), while the dry breaks occurred in the second half (except for Debrecen with short summer breaks, only). All dry breaks

Table 5
The same as Table 4, but for the homogenised time series

Station	PDSI								
	April		June		August		October		All (+/−)
	Period	Year[a]	Period	Year[a]	Period	Year[a]	Period	Year[a]	
Debrecen	1906–1942	37 (+)	–	–	–	–	*1912–1944*	*33* (+)	***70*** (+)
	1925–1997	**73** (−)	**1958–1964**	**7 (−)**	**1958–1964**	**7 (−)**	*1945–1995*	*51* (−)	***138*** (−)
Miskolc	1910–1942	33 (+)	1901–1981	81 (+)	1937–1941	5 (+)	1912–1944	33 (+)	***152*** (+)
	1945–1997	53 (−)	1945–1997	53 (−)	1946–1997	53 (−)	1945–1995	51 (−)	***210*** (−)
Nyíregyháza	1914–1916	3 (+)	–	–	**1913–1915**	**3** (+)	1905–1941	37 (+)	***43*** (+)
	1918–1996	**79** (−)	–	–	**1986–1994**	**9** (−)	*1927–1995*	*69* (−)	***157*** (−)
Kecskemét	1905–1923	19 (+)	**1905–1915**	**11** (+)	1905–1927	23 (+)	**1905–1941**	**37** (+)	***167*** (+)
	–	–	–	–	1906–1949	77 (+)	–	–	–
	1943–1997	55 (−)	*1916–1998*	*83* (−)	**1982–1994**	**13** (−)	*1941–1994*	*54* (−)	***205*** (−)
Szeged	*1901–1945*	*45* (+)	*1909–1941*	*33* (+)	–	–	1915–1942	27 (+)	***105*** (+)
	1926–1998	73 (−)	**1946–1998**	**53** (−)	1945–1994	50 (−)	**1942–1990**	**49 (−)**	***225*** (−)
Positive	5 Periods	137 (+)	3 Periods	125 (+)	4 Periods	108 (+)	5 Periods	167 (+)	(17) ***537***
Negative	5 Periods	333 (−)	4 Periods	196 (−)	5 Periods	132 (−)	5 Periods	274 (−)	(19) ***935***

Bold values denote significant sub-periods (breaks) that are **fully** identical in the original and homogenised series.
Italic values denote significant sub-periods (breaks) that are *nearly* identical in the original and homogenised series.
[a] Note: (−) significant dry period; (+) significant wet period.

presented for April–August end in 1998, followed by an enormous wet period between autumn of 1998 and summer of 1999.

The difference between breaks of neighbouring stations that has not been totally eliminated by the homogenisation is largely explained by the true sporadic distribution of monthly precipitation and, to some extent, by small differences of soil characteristics being adapted to climate of much longer time periods.

The above characterised distribution of wet and dry breaks within the entire data sets is in coincidence with other findings suggesting that climate of the Great Hungarian Plain became drier in the recent decades (Molnár and Mika, 1997; Szinell et al., 1998; Makra, 1999; Pongrácz et al., 2000). Relationships of these PDSI anomalies and hemispherical temperature variations of the 20th century are presented in a separate paper (Horváth, in press).

5. Conclusions

To sum up, our results presented in this paper can be concluded as follows.

- Seasonal dependence of standard deviation, skewness and autocorrelation of PDSI is detected in East Hungary, despite its aimed non-seasonality for the region (and time period) of the original parameterisation performed by Palmer (1965).
- According to both the K–S test and the χ^2-test distribution of monthly PDSI data is normal for each of the five stations and four selected months between April and October.
- Hence, the Makra-test detecting sub-periods with significantly different averages from the mean of the entire data set can be performed for these PDSI series.
- Homogenisation of temperature and precipitation data leads to modified PDSI time series, that yielded longer duration of breaks with no change in the overall numbers in the vegetation period. Another major feature of the homogenised data is the considerably decreased asymmetry of the dry breaks' dominance.
- According to Makra-test, characteristic wet periods can be observed between 1901 and 1940s, which is followed by a significant dry period until the end of the century (between 1940–1990s). These results confirm other studies based on similar input data but different statistical methodology.

Acknowledgements

The authors thank T. Szentimrey for providing us with the MASH homogenisation program and also for his introductory advice on its use. Part of the computations was performed by using non-commercial software 'PALMER' -PC-version of the program computing the drought indices Z-index and PDSI, adapted from its original written by T. Karl, US NOAA, modified by N.R. Dalezios, University of Thessaly, Volos, Greece. The authors thank T. Szentimrey and L. Viharos for the statistical consultations and Z. Sümeghy for his contribution with digital mapping.

References

Alley, W.M., 1984. The Palmer drought severity index: limitations and assumptions. Journal of Climate and Applied Meteorology 23, 1100–1109.

Briffa, K.R., Jones, P.D., Hulme, M., 1994. Summer moisture variability across Europe, 1892–1991. An analysis based on the Palmer drought severity index. International Journal of Climatology 14, 475–506.

Dalezios, N.R, Papazafiriou, Z.G., Papamichail, D.M., Karakostas, T.S., 1991. Drought assessment for the potential of precipitation enhancement in northern Greece. Theoretical and Applied Climatology 44, 75–88.

Faragó, T., Kozma, E., Nemes, C., 1989. Drought indices in meteorology. Idöjárás 93, 45–60.

Horváth, Sz. Spatial and temporal characteristics of the Palmer drought severity index in the Hungarian catchment area of the Tisza river. Physics and Chemistry of the Earth, in press.

Karl, T.R., 1986. The sensitivity of the Palmer drought severity index and Palmer's Z index to their calibration coefficients including potential evapotranspiration. Journal of Climatology and Applied Meteorology 25, 77–86.

Makra, L., 1998. Mean sea level teleconnections on the northern hemisphere and the Earth. Climate and environmental change. International Geographical Union, Commission on Climatology. Proceedings 69–72. Evora, Portugal, 24–30th August 1998.

Makra, L., 1999. Concentration of contaminating gases and their connection with large-scale weather situations in Szeged. The Association of American Geographers, 95th Annual Meeting, Abstracts, 376. Honolulu, Hawaii, USA, 23–27th March 1999.

Makra, L., Horváth, Sz., 1999. Spatial and temporal characteristics of air quality status over a middle-sized urban area. In: de Dear, R.J., Potter, J.C. (Eds.), Proceedings of the 15th International Congress of Biometeorology (ICB'99) and the International Conference of on Urban Climatology (ICUC'99), Sydney, Australia, November 1999, ISBN 1 86408 543 6.

Makra, L., Tar, K., Horváth, Sz., 2000a. Some statistical characteristics of the wind energy over the Great Hungarian Plain. The International Journal of Ambient Energy 21/2, 85–96.

Makra, L., Horváth, Sz., Zempléni, A., Csiszár, V., Tar, K., Motika, G., Sümeghy, Z., Károssy, Cs., 2000b. Spatial and temporal characteristics of air quality status in southern Hungary. In: Proceedings of the 3rd European Conference on Applied Climatology, ECAC 2000, Pisa, Italy. October, 2000, CD Session VI.

Mika, J., Bartzokas, A., Dobrovolny, P., Brazdil, R., Niedzwiedz, T., Tsidarakis, L., Dalezios, R.N., 1994. On spatial and temporal variations of drought in selected regions of central and southeastern Europe. In: Brazdil, R., Kolar, M. (Eds.), Contemporary Climatology, 1994, IGU Symposium 15–20, Brno, Czech Republic, pp. 395–402.

Mika, J., Horváth, Sz., Fogarasi, J., Makra, L., 1998. Simulation of climate forcing mechanisms on the energy and water balance of a watershed. In: Proceedings of 19th Conference of the Danube Countries, Osijek, Croatia, 15–19, June 1998, pp. 331–342.

Mika, J., Horváth, Sz., Makra, L., 2001. Impact of documented land use changes on the surface albedo and evapotranspiration in a plain watershed. Physics and Chemistry of the Earth, Part B 26 (7–8), 601–606.

Molnár, K., Mika, J., 1997. Climate as a changing component of landscape: recent evidence and projections for Hungary. Zeitschrift für Geomorfologie 110, 185–195.

Palmer, W.C., 1965. Meteorological Drought. Research Paper, no. 45, US Weather Bureau, Washington, DC, p. 58.

Peterson, T.C., Easterling, D.R., Karl, T.R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E.J., Hanssen-Bauer, I., Alexanderson, H., Jones, P., Parker, D., 1998. Homogeneity adjustments of in situ atmospheric climata data: a review. International Journal of Climatology 18, 1493–1517.

Pongrácz, R., Bogardi, I., Duckstein, L., 2000. How do weather patterns and ENSO influence extreme hydrology in Hungary? In: Ijjas, I. (Ed.), Inter-Regional Conference on Environment—Water: Conference on Water Resources Management in the 21th Century with Particular Reference to Europe. Budapest, pp. 169–176.

Szentimrey, T., 1995. General problems of the estimations of inhomogeneities, optimal weighting of the reference stations. In: Proceedings of the 6th International Meeting on Statistical Climatology. Galway, Ireland, pp. 629–631.

Szentimrey, T., 1999. Multiple analysis of series for homogenization (MASH). In: Proceedings of the Second Seminar for Homogenization of Surface Climatological Data. Budapest, Hungary. WMO, WCDMP-No. 41, pp. 27–46.

Szentimrey, T., Faragó, T., Szalai, S., 1992. Window technique for climate trend analysis. Climate Dynamics 6, 127–134.

Szinell, C., Bussay, A., Szentimrey, T., 1998. Drought tendencies in Hungary. International Journal of Climatology 18, 1479–1491.

Tar, K., Makra, L., Horváth, Sz., Kircsi, A., 2001. Temporal change of some statistical characteristics of wind speed over the Great Hungarian Plain. Theoretical and Applied Climatology 69 (1–2), 69–76.

Thornthwaite, C.W., 1948. An approach towards a rational classification of climate. Geographical Reviews 38, 55–94.

Wilhite, D., Glantz, M., 1985. Understanding the drought phenomenon: the role of definitions. Water International 10, 111–120.